Robotics, Biological Grounding and the Fregean Tradition

Marti Hooijmans* **& Fred Keijzer**
(marti@hooijmans.net), (f.a.keijzer@rug.nl)

Faculty of Philosophy
Department of Theoretical Philosophy
University of Groningen
The Netherlands

Abstract

With the advent of dynamic, embodied and situated cognition, action and agency came to be thought of as the core of cognitive systems. In this context, robotics became an important way to study the behavioral kernel of cognition. However, in this paper we will ask whether robots, or what kind of robots, really satisfy the requirements necessary for this purpose. The only naturally occurring embodied agents are biological organisms, and these have many characteristics that robots do not. Biological organisms display much more sophisticated behavior than any current robot is capable of. Do these robots miss specific key characteristics, present in organisms, which they need in order to model full embodied agents? We will call this the biological grounding problem, analogous to the familiar symbol grounding problem. Interestingly, once this grounding issue is raised, a full biological interpretation of cognition turns out to be almost self-evident. As such a biological, or biogenic, approach also seems very plausible from a general, natural science point of view, the question arises why a strict biologically based interpretation of cognition remains so controversial in the cognitive sciences. In this paper, we try to explain this discrepancy by sketching the strong influence of formal, computational approaches that we describe as the Fregean tradition. In this Fregean tradition, many cognition-related concepts have developed a life of their own without the need of being tied down to a particular kind of material – in this case biological – system. Any limitation to biological systems seems overly strong in this perspective. To ease the tension between these two options, we will argue that the two are not mutually exclusive, but that each has its own domain: the biological and communicative domain respectively. The biological domain consists of issues relating to how agency arises and is constituted in the first place. The communicative domain deals with communication between agents and the symbol systems that have arisen for this purpose in the human case. We will finish by discussing a number of examples to illustrate the implications of this division of labor for the use of robots in dynamic, embodied and situated cognition, and argue for a differentiated approach where the biological domain is set free from the still dominant Fregean tradition.

1 Introduction

Are robots embodied? This question may seem strange at first blush. Robots are machines that are built to accomplish things in the world, say making cars. Having some kind of mechanical body, consisting of actuators and at least some sensors just seems part of the notion of robots: Only if there are body parts you can have a robot. However, when Tom Ziemke raised this question he had something slightly different in mind (Ziemke, 2001). Having a physical body and embodiment are not necessarily the same. The notion of embodiment has become important in the cognitive sciences as part of a movement that drew away from the idea that cognition is at heart a process consisting of internal symbol manipulation.

For example, Rodney Brooks (1999) has been very influential in pressing the point that it was necessary to shift attention from internal symbol manipulation to a view in which the connection of internal information processing with the external world became a central issue. For this purpose he introduced the notions of embodiment – intelligence derived from a specific bodily system involving

* Authors are in alphabetical order. We thank Daan Franken, Allard Tamminga and Marc van Duijn for their comments. This work was partly supported by the NWO research grant 016-038-301.

sensors and actuators – and situatedness – intelligence arises through interaction of an embodied agent with a particular environment (Brooks, 1999). It is now widely accepted that for any form of internal symbol manipulation to be considered cognition it must become embodied and situated in a way that connects it with happenings in the real world.

By becoming embodied and situated, the cognitive sciences could circumvent some major problems, such as for example the frame problem (McCarthy & Hayes, 1969) that derived from an overly strong reliance on knowledge representation. Brooks' famous argument to skip the representational bottleneck and let '*the world be its own model*' is the most well-known articulation of this idea. In the same way, embodied and situated models provided a way that seemed to solve what Harnad baptized as the symbol grounding problem (Harnad, 1990). In the original internal symbol manipulation approach, there were no direct links between the syntactic symbols and external circumstances. The only reason why meaning seemed to be present was through a meaningful interpretation by the humans who built and used such systems. The symbols therein remained ungrounded in the world. Harnad claimed that such grounding could be achieved – and the symbol grounding problem solved – by giving such internal systems their own connections to the outside world through sensors and actuators. Thus, by embracing embodiment and situatedness, which provided such connections, cognitive science became capable of dealing with this problem as well.

Brooks was of course not the only key figure in initiating this change in the perception of cognition[1] . One important addition that must be mentioned here is the field of dynamical systems theory, which deserves attention as a way to describe the continuous, ongoing character of the interactions between agents and their environment (Beer, 1995; van Gelder, 1998). In the following, we will refer to this general approach as DESC, standing for dynamic, embodied and situated cognition.

The move toward DESC pushed the field of robotics to the centre of theoretical interest for cognitive science, instead of the peripheral position it took in before. By building and studying robots one could – even must – come to grips with the agent-environment interaction issues that were of crucial importance for the understanding of cognition within DESC, in a way that could not be done by models of internal symbol manipulation, nor by connectionist networks. Of course, it was, and still is possible to simulate such interactions, but in addition there was also a need for the real thing as a way of discovering and dealing with the unforeseeable intricacies of real-life agent-environment interaction (Chiel & Beer, 1997). Thus, given that robots are physical systems that explicitly incorporate sensors and actuators, how could anyone in his right mind question whether robots are embodied or not?

The motivation to raise this issue comes from what may now, with hindsight, be called a new additional *biological* step in the development of DESC (Keijzer, 2001, 2006), or even a different biological approach altogether (Lyon, 2006a). One way to reach such a full biological view on cognition is by considering the *biological grounding problem*. Roughly, the problem here is that we cannot be very sure that the artificial agents that are being built, are really agents at all, as compared to biological organisms. When we see a six-legged robot we call it an insect-like robot, but as Caroline Chang put it: "*Just because a given robot has six legs, it is not a biorobotic insect*" (Chang, 2002, p. 1024). The similarity involved is extremely superficial, and is only a similarity for us. It does not need to reflect any fundamental similarities between the two systems. All cognition or agency might be imputed onto the system by us.

The point can be best introduced by turning back to the issue of symbol grounding, mentioned above. This problem can be described as the issue how a model *Y* can come to exhibit some property *x*, *meaning* in this case, which is exhibited by the real thing, that is, *us*. Harnad claimed that internal symbol processing models could acquire meaning by becoming embodied. Adding sensors and actuators that link the model to the external environment could ensure that the internal symbols acquired meaning independent of our own interpretations. One may quibble whether Harnad did enough to solve this problem, but there is also a more general point to make. Once one starts thinking about grounding issues – How can cognitive model *x* come to incorporate the cognitively necessary

---

1 For a genuine introduction and better overview of these developments see e.g. Anderson (2003); Clark (1997); Pfeifer & Scheier (2001)

property *Y*? – then it becomes difficult to decide when one has succeeded. This difficulty derives from the phrase 'cognitively necessary property *Y*.' How do we, as cognitive scientists, decide what are the necessary properties of any cognitive system?

Symbol grounding was about the need to incorporate meaning in cognitive models. Meaning is easily cast as a key ingredient of any kind of symbol manipulation, and thus, once it became clear that it had not been satisfactory dealt with in earlier models, the need to include it somehow became obvious (Dretske, 1981; Fodor, 1987). However, DESC introduced a new situation in which many new 'key ingredients' came into the cognitive picture, most notably of course dynamics, embodiment and situatedness. Intelligence was no longer a derivative of an internal cognitive system, but emerged out of the interactions between an agent and its environment. The cognitive system itself thus became spread out to include such interactions. However, with this extension of the cognitive domain arose a new source for grounding problems: the biological grounding problem.

So far, cognitive science tended to cast cognition as a general property that could exist in both natural and artificial forms, each of which could be cast as legitimate cases of cognitive systems. However, with the issue of grounding in mind, one needs to worry whether the artificial cognitive system does actually incorporate the necessary ingredients that exemplify the natural cases of cognitive systems in the form of biological agent-environment interactions. At the same time, human beings are very prone to make such assumptions (Scholl & Tremoulet, 2000; Mar & Macrae, 2007). To return to an example using 'insect-like' robots illustrating such assumptions, in his reaction to Brooks' paper "*Intelligence without representation*", Kirsh named his own paper "*Today the earwig, tomorrow man*? " and questioned whether this approach could scale up to the human case (Kirsh, 1991). At the same time, he seemed to think in a hand waving way that the earwig level had already been reached by these robots, which is of course absurd. Just as in the case of semantics, a human observer easily comes to a particular judgment of similarity, without a clear indication how or even whether this judgment is justified for the system under description. How much of the dynamical, bodily and environmental properties have to be included when modeling a cognitive agent? From this perspective, the question whether a robot is embodied suddenly becomes an open issue that is not solved by simply pointing at some robotic limbs or sensors.

The point is not that suddenly robots can no longer be considered to be sufficiently grounded compared to natural, biological cases of agent-environment interactions. The point is that we need good criteria to decide whether or not any particular robot is suitably grounded, and that this grounding can take many different forms (Sharkey & Ziemke, 2000; Ziemke, 2001). As Sharkey and Ziemke argue, the problem with biological grounding is that it remains unclear what criteria we are to apply to robotics to approach the biological cases, and also how much it matters if we don't.

From here on, the debate can be waged in several ways. First, one might ignore biology altogether and just build robots. This is perfectly legitimate, but does not address the cognitive issues focused on here. Second, one might search for general modeling criteria that specify when a robot forms a good model of some biological behavior (Webb, 2001). The problem that we see here is its methodological focus that on its own does not sufficiently address the issue which biological aspects are cognitively relevant, a point to which we return in the next section. Third, one might oppose the whole idea of taking biology as the paradigm case of cognition and remain committed to a much more general interpretation of cognition, e.g. as computation or information processing. We will discuss the motivation behind this option below in the section on the Fregean tradition. Fourth, and this will be the approach that we will develop first, one can try to deal with this (possible) shift in perspective concerning cognition. In particular, the question arises: *How can, and to what extend must, biological grounding be achieved when robots are used as models of cognition as conceived within DESC?*

The overall approach in this paper will be to develop a characterization of two general tendencies within the sciences dealing with cognitive phenomena. First, we will sketch the biological turn in the study of cognition and provide a short overview of the biological grounding problems. Secondly, we will provide a sketch of what we call the *Fregean tradition*, a major intellectual development within the 20 century that provides the intellectual roots and background of not only cognitivism, but also to a large extend that of DESC itself. The Fregean tradition can be characterized by its tendency to focus on formal systems of decontextualized symbols and its aim to investigate and develop such systems. DESC may distance itself from formal symbol systems to some extent, but its defenders are often raised within this tradition and they still interpret cognitive phenomena in a general, abstracted way, rather than as a phenomenon which as an empirical fact exists *only* in biological organisms. Of course, biological systems are acknowledged as realizers, but primarily in a secondary and grudging way. Biology has to be forced onto cognition, and to overcome being considered a matter of mere and essentially arbitrary implementation details. The old discussion on connectionist networks and their *possible* relevance nicely illustrates the point. Thus, we claim that an approach that starts out with biological cases constitutes a genuinely different development which goes beyond what has been done so far under the label of DESC. All the same, this does not discredit the Fregean tradition as an important intellectual achievement with many merits. The problem rather becomes how the two *relate* to one another, and how this relationship bears on *testing* cognitive theories by artificial means, such as in robotics.

Our suggestion to ease the tension between the two is by distinguishing two separate domains or tiers of cognitive phenomena, being on the one hand a *biological tier* and on the other a *communicative tier*. The communicative tier deals with the issue how communication between agents is achieved. The existence of agents is here assumed without detailed claims concerning their nature or makeup, as the problems addressed are related to how communication between such agents takes place. We claim that the Fregean tradition has its legitimate background in such a communicative context and has many applications here. The biological tier on the other hand relates to the issue what requirements are fulfilled in biological organisms, most notably those centered on sensorimotor capabilities, in a way that makes these organisms bona fide agents in the first place. Elsewhere, one of us used the phrase *animality* to refer to this yet to be articulated biological organization (Keijzer, 2006). The problem here is that while we already have an extensive 'Fregean' conceptual framework, ultimately derived from our own communicative and symbolic practices, we have much less of a biological conceptual framework tailored to deal with the multitude of processes and structures involved in generating biological cases of agency. This makes it very seductive to bring the Fregean tradition to bear on these issues too. However, by clearly differentiating between these two general domains, we provide an argument why the Fregean tradition would be overstepping its bounds here. Instead, animality – the biological organization involved – should become the target of a biologically based account of cognition.

The remainder of this paper follows the structure outlined above. The next two sections will discuss the switch from symbol grounding to biological grounding and the turn to a full biological interpretation of cognition. The subsequent section will introduce the Fregean tradition and sketch some of its characteristics and implications. Then we will discuss the two tier view on two different cognitive domains and argue that the first provides a natural home for work in the Fregean tradition and the second a different, conceptually very open empirical space that needs to be filled in by future work on the biological roots of cognition. We will round of by discussing how this feeds back to robotics.

2 From symbol grounding to biological grounding

What we here call the biological grounding problems developed out of a different and more generally appreciated grounding problem, that of symbol grounding. The latter can be traced all the way back to Searle's influential Chinese room thought experiment. This thought experiment was aimed at *computationalism* within AI, a stance that Searle dubbed *strong AI*: "*the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer really is a mind*" (Searle, 1980, p. 417). Thus, according to strong AI, if an AI agent would master syntax, semantics would follow, and with it would come understanding and intentionality. Searle used his Chinese room thought experiment to argue that even when there was a full syntactical mimicking of the outward characteristics of a cognitive agent, understanding and meaning would still be absent. What was missing, according to Searle, were the special characteristics of the human biological brain, which were necessary for understanding and intentionality as present in humans. This insistence on the relevance of the particular biological makeup of the human brain was subsequently severely criticized for many years because cognitive scientists and AI people thought he completely missed the point of the multiple realizability of cognitive phenomena.

The Chinese room made it very clear that the connection between symbolic representations, in particular explicit symbolic representations within an explicit symbol sign system, and the meaning of these symbolic representations was not a trivial matter. It became evident that it was unclear where the assignment of meaning within a symbol system came from, or by whom it was used: either the symbol system making up an artificial intelligence, or rather the scientist who developed the system. Whether one did agree or not with Searle, an explanation had to be given that clarified how purely syntactic symbols systems could get (or possibly even had already) their own semantics, independent from that of any outside observer.

While Searle's main aim with the Chinese room argument was to argue against computationalism within cognitive science, Harnad (1990) later used Searle's argument to formulate the *symbol grounding problem*. He drew the following conclusion from Searle:

> The symbols and the symbol manipulation, being all based on shape rather than meaning, are systematically interpretable as having meaning — that, after all, is what it is to be a symbol system, according to our definition. But the interpretation will not be intrinsic to the symbol system itself: It will be parasitic on the fact that the symbols have meaning for us, in exactly the same way that the meanings of the symbols in a book are not intrinsic, but derive from the meanings in our heads. Hence, if the meanings of symbols in a symbol system are extrinsic, rather than intrinsic like the meanings in our heads, then they are not a viable model for the meanings in our heads: Cognition cannot be just symbol manipulation. (Harnad, 1990, p.339)

However, while Searle lay the blame with the internal process of symbol manipulation itself, which was somehow too limited, Harnad came to a different conclusion. The problem was not symbol manipulation itself, but the lack of suitable connections between such symbol systems and the real world. This was what he called the symbol grounding problem: How can the symbol system become connected to things outside this symbol system. His proposed solution was by providing such systems with their own sensors that fed into connectionist networks with their own pattern recognition and associative capabilities, thus providing a source of representations more directly linked to sensory states. Without such linkages, and as long as the symbol system remained an engineered entity, the only meaning present in such a system would be the one provided by external agents, that is by us.

The proposed theoretical solution for the symbol grounding problem by Harnad is not trivial to implement in practice though. Within DESC, situated robotics seems to address the grounding problem from the start by using actual robot bodies that are situated in the world and so linking the internal cognitive system to the outside. Following Brooks we may call this the *physical grounding hypothesis* (Brooks, 1991). However, Sharkey and Ziemke, also sympathetic to DESC, argued that the grounding problem is not limited to problems concerning syntax and semantics (Sharkey & Ziemke, 2000). They advocated a strongly biological, integrated view on cognition and on solving the symbol grounding problem. They also claim that grounding is a general issue that involves more than only linking the symbolic realm to the external world. In their view, there are basic problems with the idea that any physical body with sensors and actuators suffices to 'hook' internal representations to the world.

Why would physical grounding with robots not be enough to counter the symbol grounding problem? First, it is unlikely to be the whole story regarding meaning and cognition (Sharkey & Ziemke, 2000). Harnad's diagnosis of the symbol grounding problem did only partially address Searle's worries. Searle himself advocates a strong biological outlook on cognition, making much of the particular physical make up of the brain. Mere physical grounding does not deal with this issue. In addition, Sharkey and Ziemke worry that current day computer based robotics technology is not sufficient for genuine embodiment. It might be that there are significant differences between the sensorimotor organization of animals and that of robots (Keijzer, 1998). Physical embodiment might not be enough if the different components are merely put together in a robot, without any intrinsic integration, except that the robot operates sufficiently according to the standards of the designer. The situation is very similar to the original symbol grounding problem where a symbol system was engineered in such a way that it satisfied its designers as a mechanism that manipulated symbolic entities, meaningful to us, on the basis of their syntactic form. Within DESC, robots are engineered in a way that leads to meaningful actions in the eyes of human observers, but not for the robot itself. As Sharkey and Ziemke note about robot action: "*The meaning of a robot's 'actions' is also in the* *observer's world and not in the 'robot world'. The robot's behaviour has meaning only for the observer.*" (Sharkey & Ziemke, 2000, p.326). For example, consider a robot that is designed to reach a specific kind of goal in its environment such as a recharging station. To what extend can this be a model of a biological agent that is hungry and in search of food which it finds by using particular cues in its environment? The problem is that we cannot tell. The general similarity is sufficient for many researchers in DESC-style robotics to pursue this kind of research, as such a set up seems intuitively similar to the kinds of problems faced by biological organisms. But *this similarity is a similarity for us*, and derives from a *very specific interpretation* of both the artificial and the biological agent. What criteria do we really have for dismissing the huge differences in physical make-up, evolutionary background, and sensorimotor complexity? How can we develop clear ideas concerning the validity of such a robotic approach? In this way, they extend the original grounding problem to the biological grounding of artificial actions.

Barbara Webb's work on biorobotics is important here (Webb, 2001). She asks the methodological question how biological behavior can be modeled by the use of robots and uses, for example, seven dimensions on which robotic models can be judged. This is essential to move beyond a "*vague justification in terms of intuitive similarities between robots and animals*" (Webb, 2001, p.1033). However, the biological grounding issue arises at a point before a more detailed methodology gets its grip. Before one can systematically apply particular criteria that enable a systematic evaluation of a particular robotic model these criteria have to be made clear. The point is that the aim here is not a particular biological behavior that is subsequently modeled more or less adequately. The aim is to clarify which biological aspects are to be considered cognitively relevant, where cognition is interpreted widely as in DESC.

Returning to Sharkey and Ziemke, it becomes very clear that there is a lot to choose here, with a wide variety of options for including, or excluding biological details. In several texts, Sharkey & Ziemke (2000, 2001) and Ziemke (1999, 2001) provide a number of slightly varying taxonomies of embodiment and grounding. To give an idea of the options when it comes to symbol grounding and biological grounding, we will sketch an overview, drawn from their work, but with our own, slightly different, categorization and labels.

Physical grounding: This is Harnad's first attempt of grounding the semantics of an artificial cognitive system. In his first proposal he focused on sensory linkages and neural nets to enable the requirement that the system made its own categorization of the world. He acknowledged that motor control would be important too but did not include it in his original proposal, giving the impression at least that it was not essential. The danger here is that one engineers a system, which only mimics our own semantics and the relation with external features as perceived by us, without an account why the system would actually embody meaning.

Robotic grounding: Here the physical grounding of cognition is achieved by linking it to a robot having sensors and actuators. This might still be done in a very classical way, such as in old Shakey (Nilsson, 1984), where the inner cognitive system remains central, and the sensory and motor components are additional hookups for this central intelligence. Ziemke notes in this context that this only ensures that an agent is hooked to its environment, but by itself this doesn't ground the internal mechanisms nor the behavior. Both internal mechanism and behavior are engineered by its builders and reflect their own interests and meaning rather than that of the robot built. The thing is made to play-act as an agent for us and it is not evident why it can be said to act for itself.

Organismoid grounding: Here the shift would be to DESC proper with its stress on organism-like robots or animats. Sensorimotor linkages between agents and their environment are here the starting point for building such artificial agents rather than an inner and separable cognitive core. Brooks' subsumption-based robots provide a clear example of this organismoid grounding. The difference with the above is that the inner meaningful entities here are not, or at least less, a direct derivative of the insights of the researcher. Instead, they are derived in a process of trial and error as useful internal states that are intrinsically connected to the behavioral functions the robot has to perform. However, this raises the question whether the robot isn't depending too much on its engineer for choosing and evaluating these behavioral functions.

Grounding by learning: The most basic case within organismoid grounding is where an agent always reacts to the environment in the same way. However, in biology, grounding is not achieved once and for all. A biological agent adapts to changes in its environment and this can be considered as an important aspect of being grounded. By making an organismoid agent capable of learning, such a system could better adapt its inner states and processes to external circumstances, and so grounding would be enhanced. However, internal parameters can be consistently modified by learning, but the control structure itself remain a given; reflecting engineering decisions instead of being intrinsic to the system itself.

Evolutionary grounding: Within an organismoid grounding approach, another possible addition would be not (only) to ground the internal parameters of control structures that regulate a robots behavior, but the control structures itself, as is being done within evolutionary robotics. In this case, the entire control structure is to be grounded in the environment by a process of simulated evolutionary adaptation. Two problems relevant for grounding remain. First, the genetic algorithm remains a matter of the engineers decisions as to what are the relevant features for selection. Second, the robot's body itself still consists of engineered parts like stock sensors, a computer, actuators and so on. Building these from raw materials and assembling them in new ways goes beyond the robot's powers.

Organismic grounding: Grounding issues are a matter of questioning whether artificial cognitive agents or aspects thereof are sufficiently encompassing to incorporate aspects that are present in real agents. Real agents are agents that do not owe their existence to our own actions. In practice, this means biological agents or organisms. A big difference between artifacts and biological systems is that the former are always engineered, and what Maturana & Varela (1980) called *allopoietic*: a system which is put together from the outside. Maturana and Varela contrast allopoietic to *autopoiesis*, or self-construction. In their view, the presence of autopoiesis is the essence of living systems. An autopoietic system is a system that is its own constructor. An autopoietic system build its own structure from given molecular components, and maintains this structure counteracting all kinds of disruptive happenings coming from without. Once one looks at any of the options described above, they seem fundamentally inadequate as they miss this basic self-producing aspect and always remain tied to the interests and preconceptions of the designers and engineers involved. As a remedy one can push for the constraint of organismic grounding which implies that only living organisms can be cognitive (Sharkey & Ziemke, 2000, 2001) and (Ziemke, 1999, 2001).

Uexküllian grounding: This form of grounding within organismic grounding more generally derives from what Sharkey and Ziemke call Uexküllian embodiment. In this case the extra assumption is made that biological grounding implies experiential grounding. Von Uexküll was a well-known biologist working in the beginning of the 20 century and became well known for his *Umwelt* theory which claimed that each biological creature had its own experiential space, which differed for different species. "*Von Uexküll criticized the mechanistic doctrine 'that all living beings are mere machine' for the reason that it overlooked the organism's subjective nature, which integrates the organism's components into a purposeful whole*" (Sharkey & Ziemke, 2000, p.320). This specific form of organismic grounding is also the position where the enactive approach of writers like Varela, Thompson and Noë should be placed.

This overview is certainly not a fully satisfactory characterization of the different options. Nor are all of these options mutually exclusive or sufficient to cover all possibilities. Our main aim for presenting this overview is to give an impression of the situation that arises once these grounding issues are taken seriously: Once one starts asking questions whether the characteristics of a certain model are actually sufficient to generate the cognitive feature that one deems important – such as semantics or agent-environment interaction – then *there are no clear criteria that tell us where to stop*. As the sequence above suggests, one can add more and more biological features that constrain the cognitive feature more and more, but new or at least different grounding issues pop up again and again until one ends up with a full biological grounding as a requirement for cognition. For all intermediate cases, the question can always be asked whether they actually satisfy the requirements necessary for instantiating an independent case of cognition rather than something that makes us think it is cognitive.

3 From biological grounding to biology

Grounding problems arise within cognitive science because one starts out with a specific cognitive feature and then builds a model that ought to suffice as a realizer of this cognitive feature. However, as long as we do not have very clear criteria as to what makes something 'cognitive' there will always remain doubt as to whether the model will become a case of cognition itself. The important point here is that our ideas as to what cognition might be are not clear at all.

Within cognitive science there is a strong tendency to ascribe cognition when we can interpret a system in terms like perceiving, remembering etc. Dennett (1987b) introduced the notion of an *intentional stance* to describe this way of having intentional – mental or cognitive – systems. Taking an intentional stance amounts to treating a system as a rational agent and figuring out what its beliefs and desires are likely to be. In this way, we can often understand and predict what such a system will do without needing to know about its detailed physical makeup. The intentional stance provides a way to combine our mental or intentional vocabulary with a mechanistic understanding of cognitive phenomena. One can use the mental vocabulary to predict and explain particular, cognitive, systems, while the stress on it being a *stance* makes it perfectly clear that this is merely a different description of a physical system. There is no risk of an unaccounted for ghost in the machine.

Cognition thus conceived is a useful and pragmatic way of demarcating the cognitive domain. Nevertheless, it comes at great theoretical costs. Most notably, we will argue, is that it obstructs a clear linkage between the cognitive domain and particular kinds of material systems.

Suppose that systems are to be considered cognitive or not on the basis of applying an intentional stance: If such a stance is applied successfully, then the system counts as a cognitive one. However, as we are free to apply the intentional stance to whatever we want – be it a falling stone, thermostat, animal or human being – and given that the notion of successful appliance is open to many different interpretations, this is a very unconstrained way of demarcating a cognitive domain. In particular, it amounts to a way of having cognition which remains independent of any particular material organization of the so described systems. Thus the view arises that, even though an AI program runs on a computer and not 'on a brain,' it still may be considered cognitive.

But there is also an opposing intuition: There ought to be something about the systems themselves that makes them cognitive or not: Humans and falling stones are just too different. The issue also comes to the fore in research on animal cognition, where it is a major research effort to establish whether particular animals can be deemed cognitive or not. Thus, there is clearly more to be said about delineating cognitive systems than applying an intentional stance. According to this intuition, we are not free to postulate cognition wherever we want.

A generally sensible way to proceed here would be to investigate the physical means that might account for any differences between genuine cognitive systems – that is systems that can be targeted as cognitive beyond reasonable doubt – and systems of which we think that we merely ascribe cognition to them. However, an intentional stance deemphasizes the means by which an agent achieves its goals. Dennett once introduced the revealing phrase 'wise wiring' (Dennett, 1987a), which illustrates the point. The phrase 'wiring' refers to an arbitrary set of connections that are in themselves neither systematic nor very important as long as it produces the required 'wise' result. This choice of words implies that the physical means involved are not specific or special but merely involves an *ordinary* system that just happens – as if by coincidence – to produce particular results. Thus, if one would stress the importance of the wiring, then this would count as a dismissal of the need for cognition, rather than its articulation. In addition, tying cognition to particular kinds of systems will always exclude systems that are currently taken to belong to the cognitive domain as derived from the intentional stance. Of course, most of those exclusions would be according to the intent of such an endeavor, but border disputes would nevertheless arise, which might seem to discredit the very project.

These opposing tendencies lead to a double bind when it comes to answering the question of what cognition is. The intentional stance provides an intuitively plausible cognitive domain, but remains too unspecific to be the whole story. At the same time, the force of the intentional stance criterion makes it almost impossible to delineate cognitive systems in a more specific way, because it is not the 'wiring' that counts, but its being wise.

To conclude, an intentional stance sets up a *conceptual* dichotomy between intentional and mechanical systems, but does so without a corresponding dichotomy between different kinds of material systems, and even prohibiting any material distinction to be the relevant one. At the same time, there are good reasons to suppose that there must be more specific physical, organizational or dynamical aspects to cognitive systems that ought to set them apart as a particular kind of material system. Thus, taking cognition as something that can be simply ascribed by taking an intentional stance may suffice for a cognitive science in the short run, but it also obstructs raising the question what cognition could be on a more fundamental material level.

When we now return to the biological grounding problems, the issue becomes how we can clarify the criteria that we are to use to demarcate a cognitive domain. What is the weight we must ascribe to biological grounding? An obvious and clear way to cut through the indecisiveness arising from the different criteria described as different grounding options, is to skip the intentional ascription criterion altogether (Keijzer, 2006). Instead of detailed discussion which biological grounding problems are more serious than others and for what reasons, the whole situation would be reversed. Biology would become the starting point for thinking about cognition, and the problem would be rather to investigate how cognitive phenomena arise within biological organisms and in such a more piecemeal way become clear on what they might amount to. It would become a mere side issue whether one could ever have something that can be deemed cognitive outside the domain of life.

This idea of an intrinsic link between life and cognition has acquired some attention in recent years (for example Keijzer, 2001, ; Sharkey & Ziemke, 2000, ; Stewart, 1996,), often inspired by the work of Maturana & Varela (1980). An important upshot of this idea is that biological organisms – which includes humans obviously – are the only truly paradigmatic cases exhibiting cognition, and thus the proper focus of the cognitive sciences. On this count, Lyon recently made a strong case for what she calls a *biogenic* approach of cognition (Lyon, 2006a, 2006b), which she opposed to an anthropogenic approach.

A biogenic approach is an approach to cognition that takes the facts of biology as its *starting point*. In her view, cognition is a biological process, or set of biological processes, which – like circulation, respiration and elimination – is best understood in the context of living organization. An investigator adopting a biogenic approach assumes that the principles of biological organization present the most productive route to a general understanding of the principles of cognition. This would for example involve a close study of simple model organisms like *Caenorhabditis elegans*, *Aglantha digitale* and even bacteria. Here one finds the most basic and presumably accessible cases of cognition, such as basic forms of behavior, the molecular or neural control thereof and the best chance – it would seem – to investigate how nervous systems operate at a whole organism level.

Lyon contrasts her biogenic approach to an anthropogenic one, where the human mind provides both starting point and benchmark of cognition. An anthropogenic approach would also include biological facts, for instance from neuroscience and evolution, but these would not provide the foundation or articulation of cognition itself. They would be secondary constraints to be satisfied by an account of human cognition. It is important to stress that both biogenic and anthropogenic are conceived as methodological biases concerning places to start investigations and to define the fundamental characteristics of cognitive phenomena. Both biogenic and anthropogenic approaches would ultimately cover all aspects of cognition, but whereas an anthropogenic approach would start out with the complexities of the human case, a biogenic approach would first develop models of much more basic forms of cognition. When phrased in this way, the cards of a biogenic approach become very strong, as this is the standard procedure in the natural sciences more generally. How much would we now know about human genetics, if it wasn't for studies on simple model systems? And why would the case be different when it comes to human cognition?

One obvious hurdle for accepting the force of this argument has been taken away by DESC. Classical and also connectionist cognitive science could maintain that only humans exhibited 'real' cognition beyond any reasonable doubt. However, work on the embodiment and situatedness of cognition now makes this much less evident. According to DESC the much more widely distributed phenomenon of organism-environment interactions ought to be cast as the center of cognition, which is a clear step toward a focus on basic model systems.

So, a biogenic approach to cognition has much to offer. Its orientation on biology and basic cases make it fit in with the other natural sciences, and by starting out with biology it provides a clear guiding line as to how to deal with biological grounding problems. When the problem of demarcating cognition is to be solved by abandoning the intentional stance as a criterion for demarcation, and instead to tie cognition to a particular kind of system, then a biogenic approach would be a very promising way to make such a project work.

Nevertheless, there are strong tendencies within the cognitive sciences which would oppose such a biological foundation for cognition. Many cognitive scientists would hold that cognition is not to be limited as just suggested. For example, when it comes to embodiment, Andy Clark recently argued that there is no need for any special interpretation of the body. Instead, he gives it a computational interpretation and holds that the body is a stable platform whose features can be relied upon in the computation of certain information-processing solutions. This is, in his view, "*what, at least for all cognitive scientific purposes, the body is*" (Clark, 2006). In addition, rather than making cognition a specific phenomenon, cognition is here cast as a general computational or information processing phenomenon, which can be spread out over computational processes within and without the body. In this view, any specific linkage to biology would seem inappropriate, and Clark even names it *biochauvinism* at one point (Clark, 2005).

What to make of all this? In the following we will try to say something about why the exact opposite of a biogenic cognitive science is so intuitively plausible for many cognitive scientists. Why does cognition seem to be primarily a general information processing affair, which does definitely occur in biological organisms, but which are not at all necessarily tied to them? We will try to develop the outlines of an answer to this question by sketching what we have named the *Fregean tradition*: A major intellectual tradition of the 20^{th} century which provides in our view the natural conceptual home for current views concerning a free-floating interpretation of cognitive phenomena.

4 The Fregean tradition

We introduce the notion *Fregean tradition* to denote a widespread body of ideas which forms the basis or background for many different fields and disciplines such as formal logic, analytical philosophy, linguistics, computer science, philosophy of mind, cognitive science, artificial intelligence, situated robotics, and DESC more generally (Hooijmans, 2006).

The Fregean tradition owes its name to the German philosopher and mathematician Gottlob Frege (1848-1925), although we want to stress that many others contributed to its origination and development as well. Frege introduced three important philosophical innovations that became very influential in the 20 century. First, there was his development of *formal symbolic logic*. Introduced in his 1879 *Begriffsschrift* (Frege, 1879; Beaney, 1997)[2] , formal logic provided a general symbolist backdrop that would make possible and shape many sciences such as computer science, linguistics, cognitive science and artificial intelligence. Second, there was his *theory of meaning*. One important idea in his theory of meaning is the distinction between *sense* and *reference* (Beaney, 1996; Beaney, 1997, p. 151-182), which dissociated the inner aspects of meaning from their external reference. Another important idea concerning meaning was Frege's context principle; meaning can not be assigned to individual words but only to sentences as a whole, because only sentences have a truth value. Individual words he viewed as subjective percepts, signifiers in a private sensory reality. Third, He developed the principle of *compositionality*, which states that the meaning of a proposition consists of the sum of its constituents and the rules to bind those constituents (Mendelsohn, 1996). This principle is very important in logic and might just be the most important principle of linguistics (Parsons, 1996). Frege's impact was for a long time indirect, working mainly through his early admirer Bertrand Russell who together with Alfred North Whitehead wrote the widely influential *Principia Mathematica* (Russell & Whitehead, 1913), and through the work of Wittgenstein, whose early (Wittgenstein, 1922) as well as later (Wittgenstein, 1953, st. 22) is to a large extent a reaction to Fregean ideas.

---

2 The English translation, including this reprint in Beaney (1997), is usually referred to as "*Begriffschrift*". We will use this name from here on as well.

The term *Fregean tradition* should be loosely interpreted as a historical tradition that started with Frege's work – or rather derived a big impetus from it – but which was later on developed in many different ways, which went far beyond what Frege envisioned himself. What made the work of Frege important to the extent that one might actually speak about a *Fregean* tradition? Why not go back to Boole, Kant, Leibniz, Descartes, or even Plato? Our main reason is that Frege's ideas set in motion a chain of conceptual and institutional developments that led to the establishment of whole new fields of science and philosophy that played a key role in the 20th century's development of the cognitive sciences and artificial intelligence (Reed, 1997). Some authors, for instance Anderson (2003), have good arguments to point to Descartes as the proper founding father. Anderson argues that the main characteristics of *cognitivism*, being *representation*, *formalism* and *rule-based transformation*, can be traced back to Descartes. However, while Descartes' influence turned the mind into something that remained outside the scientific domain of physical systems, the theoretical underpinning in the work of Frege brought them back as formal, but not strictly immaterial symbol systems, and it is only after Frege, and the ensuing flurry of work on logic, analytic philosophy and computation that these concepts acquired their modern, specific interpretation, as well as their huge scientific influence. From Frege onwards, there is a continuously ongoing body of work that via analytical philosophy, logic, computation theory, led to early cognitive science and artificial intelligence (see figure *1*).

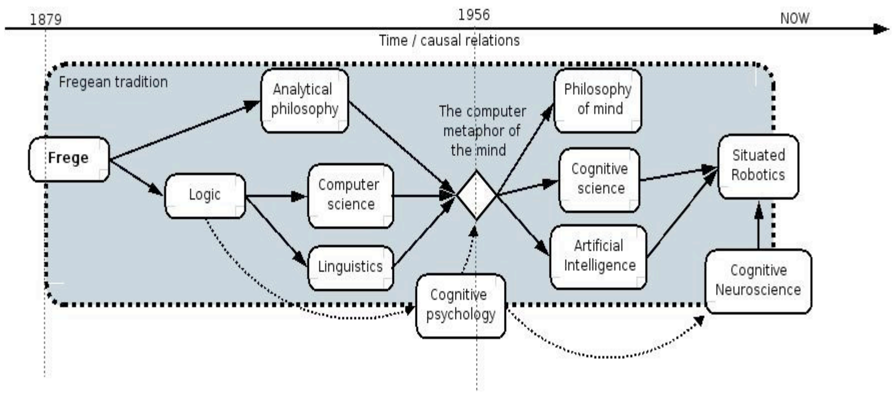

Figure 1: A diagram of the Fregean tradition. The time line at the top is a very rough guide indicating two mayor years. 1879 was when the Begriffschrift was published, 1956 was the year the Dartmouth Summer Research Conference on Artificial Intelligence was held. Generally the diagram should be read as a flowchart; the arrows signify what discipline caused the mayor development of another. That means the boxes signify the *first onset* of a certain science.

We want to stress two points in particular that make Frege a key figure for later developments in the cognitive sciences. First, Frege made huge contributions to the formalization of logic and meaning, turning it into a domain of its own which could be, and was developed further without any need for a strong empirical grounding. From here on fields like logic and analytical philosophy became to a large extent self-sufficient formal domains. At the same time, these fields remained tied to central human features like reasoning and meaning, in a way that was different from, say, mathematics more generally. Second, Frege was very much opposed to what he called psychologism. Psychologism is any approach which confuses formal logic with empirical facts about how people think (Boden, 2006). Formal logic is a normative enterprise which aims to develop laws of logic, while psychology aims to find out how people actually think, which may be valid or not. An effect of his anti-psychologism was that he maintained a firm distance between his formal treatment of topics like reasoning and meaning, and empirical, psychological studies of reasoning. For these two reasons, Frege can be cast as the one who provided both the means, as well as a motivation for a free-moving formal approach, which wrestled philosophy free from psychology.

The ecological psychologist Edward Reed (1997, p. 187-194) argued, that modern philosophy would not even have existed without the explicit division of logic and psychology that notably Frege argued for. According to Reed, it remains to be seen whether psychology and philosophy would ever have separated paths if it would not have been for the development of formal logic, and if Frege would not have been so determinedly anti-psychologistic. This anti-psychologism could also be found with Russell and Husserl, who can be viewed as the founders of modern academic philosophy, while both were heavily influenced by Frege. So in short, if it would not have been for Frege's ideas, it might just be so that analytical philosophy (Dummett, 1991; Kenny, 1995) and logic (Schirn, 1996) would not have developed in the form it did.

Of course, one major player still needs to be introduced, the computer. What Frege's work to a large extent initiated was the development of a field that investigated laws of formal reasoning. One of the upshots of this research was the development of the modern digital computer, with Turing (1950) and von Neumann (1945) as the the central names. However, the development of the computer did not only feed on the Fregean tradition but also fed back on it, giving it an enormous impetus by making the formal operations on propositions the core of a machine that followed their execution, independent of any direct human intervention. One might say that formal operations became *self-sustaining* within a digital computer. The story about the rise of the cognitive sciences from here on is well-known and has been told already much better than we can accomplish here by Gardner (1984) and recently by Boden (2006). As we all know, the computer metaphor of the mind and a general information processing interpretation of cognition became the generally accepted story concerning the nature of mind.

Within these subsequent developments, Frege's anti-psychologism did not hold, and there was a lot of influence from psychology on the development of the cognitive sciences more generally. However, when psychology turned away from behaviorism and became 'cognitive', it also became Fregeanized to an important degree. Frege's notion of a formal domain, that could be investigated irrespective of the details of its human and neural implementation became a mainstay of cognitive psychology, providing the ideas as to what probably happened in the brain, rather than the other way round. Chomsky's focus on an abstract language competence rather than actual language execution is a nice example of the importance of formal ideas for linguistics (Chomsky, 1957). Another example of a concept coming from the Fregean tradition would be the abstract notion of information processing, which comes originally from Shannon's theory of communication, or information theory (Shannon & Weaver, 1949). Notions from information theory, such as encoding, decoding, input, output, signal transduction, have become fundamental features within modern cognitive psychology, and point out an important influence from a formal theory of human communication on cognitive science, which investigates the operations of human cognition.

The upshot is that, once the Fregean tradition got underway, many different ways of modeling topics like reasoning, thought, knowledge and cognition developed. What makes these topics all part of a single Fregean tradition is the way in which they all are treated as domains in their own right, which of course must be embodied somehow, but the essence of which is not captured by the detailed characteristics of any embodiment or realizer. The key task would be the *naturalization* of a cognitive domain, but not an all out *naturalistic* approach that starts out with the particulars of a specific kind of natural systems, as for example a biogenic approach does. Even when the strictly formal starting point becomes modified, such as in neural networks or neuroscience, many still tend to adhere to the notion of a general cognitive domain which could be exemplified in various systems, only one of them being the human case. Making the additional claim that it is something about the specific biological make up of organisms that makes them cognitive remains very much a minority view, which sounds counter-intuitive to anyone raised within the wide Fregean tradition.

We also want to stress that DESC itself is still mostly operating within this Fregean tradition. Our reasons follow the same line of thought that Lyon made with her differentiation between anthropogenic and biogenic approaches to cognition (Lyon, 2006a). In both cases the difference lies with the starting position, being either the human case or biology more general. When it comes to DESC, it aims to understand cognition better by bringing in all kinds of extras that were missing from classic computational accounts of cognition, such as embodiment and the situatedness of intelligent action. In other words, it aims to solve the grounding problems that arise when one starts out with an abstracted – even if not human-based – notion of cognition. DESC is still very different from a full biogenic approach that takes biological systems as its paradigm case.

To round off, we have introduced the Fregean tradition as a very general and important domain of thought which extends over quite a number of scientific and philosophical fields. Its origin and center lies in a focus on a formal symbolic domain construed as a compositional system whose atoms are explicit symbolic representations or representation like elements (van Emde Boas & Janssen, 1979). This idea has two extremely important links that made it very influential for cognitive science. First, such a system was developed as a formalized and objectified form of human conscious thought, even if only in an idealized and normative way. Second, such a system could be mechanized within a computer and became dissociated from direct human intervention. This central assumption of a symbolic realm is very strong and provides a starting point for conceptualizing cognition even when one diverges from classic computationalism, such as in work on neural networks.

The notion of a Fregean tradition provides a backdrop framework for cognitive science in a historical and conceptual way and puts the conceptual framework of cognitive science in a broader context. We think the Fregean tradition constitutes an extremely important domain of modern thought and science and also an important backdrop for understanding cognition. However, we also think that a Fregean approach on its own misses essential and fundamental aspects of cognitive phenomena, a point that we hope became clear in our discussion on biological grounding. In the following section, we propose a way that might help to overcome this dilemma.

5 A two tier view on cognition

Does a biogenic approach discard too much? It may seem rather extreme to limit the cognitive domain to biological organisms, disqualifying artificial instances of cognition. On the other hand, can one simply dismiss the arguments related to biological grounding and the straightforward scientific sense of a biogenic approach? How to choose between these two options? In the following, we will try to limit the tension between the two by superimposing a very general two tiered classification of cognitive phenomena (Hooijmans, 2006). On the one hand there is a *biological tier*, which relates to the biological foundation of organismal behavior, and the means by which biological organisms become bona fide agents. On the other hand, there is a *communicative tier*, which relates to the communication between agents – agents whose existence is merely assumed – and where the focal point centers on the characteristics of the symbol systems used for communication. While the first tier seems to be best dealt with by a full biological and even biogenic approach, the same need not apply for the second tier. As all cognitive scientists know, symbol systems can acquire a life of their own and develop in ways which are not necessarily constrained by their biological origins. Thus, the second tier seems well-fitted to, and already well-explored by research within the Fregean tradition. Let us look at these two tiers in slightly more detail.

The biological tier: As said, this domain relates to the biological foundation of organismal behavior, and the means by which biological organisms become bona fide agents. We want to stress that the biological tier is not necessarily limited to biological systems, excluding artificial cases for all times. The point is that, without a solid account of cognition, we must first develop such a solid account on the basis of systems that are unequivocally cases of cognition and which are not merely made to look like cognitive systems to us. In line with DESC, we claim that the core issue to be investigated would be the development of sensorimotor systems and their subsequent increasing complexity in the course of evolution. From a classic cognitive perspective this may lay outside the domain of cognition itself, but DESC made at least clear that at this more fundamental level, many processes and structures are operative which are essential components of explanations for high-level cognition. For a biogenic approach these sensorimotor processes would simply be the easier studied basic forms of higher-level cognitive ones. One way of developing this domain is by differentiating it from agency and by focusing on the material and functional setup that leads to the applicability of agentive interpretations. One can refer to this domain with the concept of *animality*, which can be described as "*the sensorimotor organization by means of which animals modify environmental conditions*" Keijzer. However, we should also stress that this focus on sensorimotor organization is only one option, where others remain possible. Lyon (2006b) for example, talks of *biogenic family traits* and tends to favor a more open ended approach with two main constraints: cognition has formed in an evolutionary process and complex cognitive capacities descend from simpler forms in a continuous way.

The communicative tier: This domain consists of communicative behavior, which we understand broadly as interaction between biological or artificial agents where information is transferred. In our view, this domain is also the home ground of the Fregean tradition. Within the context of (human) communication notions like representations and symbols self-evidently come to the fore. Of course, communication is much wider than human communication, and much of it happens without intricate symbol systems, but in the human case it did give rise to language and its intricate structures. Such structures do acquire an existence of their own, and can be studied in their own right, particularly when written down or objectified in other ways. In addition, one may expect an impact of such structures on biological agents and in particular humans, which would provide a middle ground between the two tiers, but leaving intact the biological tier as a separate territory

What is the benefit of making this distinction? First, it sets basic organismal behavior apart as a *key issue* for cognitive science and as the fundamental starting point for thinking about and investigating cognition. Second, by stressing communication as a different domain, room is created for work within the Fregean tradition which focuses on symbols, symbol systems and meaning not tied to individual subjects. This space is created without jeopardizing the biological foundation of cognition, as it arises originally from this biological context, but still can acquire a certain independence. When it comes to human language and other sufficiently complex symbol systems, these can acquire to a significant degree a life of their own, independent from any individual agent, and can be used without reference to such agents. This allows the study of such systems independent from particular agents, and also their active development in increasingly complex ways, as occurs in the different forms of logic and analytic philosophy.

What is the relation between these two tiers? This all depends on the kind of species that is under investigation. As long as one deals with species that do not use intricate symbol systems, cognition can be studied within this biological domain alone. The human case, at least, seems different as our cognitive functioning is heavily influenced by the symbolic environment in which we are raised and in which we operate. It is here that the two domains become intertwined and human cognition can even be seen as an amalgam of both. The importance of external symbols for developing and changing our thought has been stressed recently by quite a number of authors, such as Dennett (1991, 1997) who talks about language as software of the brain, Clark (1997) who stresses that language is the ultimate artifact and a major enabler of thinking, and Deacon (1997) in whose eyes language is co-evolutionary catalyst that shapes human symbolic capacities. All these authors see language as a tool that enhances human thought and cognition.

The important implication of the two tier view, however, would be that this specifically human symbol dependent form of cognition is an addition to a pre-existing biological cognitive domain, not something that can be separated from it, like the external symbol systems themselves. The two tier view provides a criticism of the tendency to take these symbol-dependent aspects of cognition as the fundamental aspect of cognition, as has occurred under the influence of the Fregean tradition. An example of this influence is provided by the notion of representation. While initially the notion of representation referred to external symbols it later became seen as something internal to the cognitive system. As Luc Steels described it: "*It is only more recently under the influence of logic and early A.I. work that representations have become viewed as something purely internal.*" (Steels, 2003, p. 2389). So the definition of representation has shifted from something external, to an ingredient within cognitive agents. For a two tier view however, the external status of representation would remain primary and the question would become how handling of external representations could change and develop internal cognitive processes. Whether this would involve the literal internalization of representations would be an open issue, not a basic assumption.

How will the biological and communicative domains ultimately fit together? Obviously, the matter is an empirical issue, but it seems that up to now research within the Fregean tradition has been very dominant and invaded what we call the biological cognitive domain, interpreting all of cognition in terms that ought to remain restricted to the much more circumscribed domain of (human) communication. For example, taking up the issue of representation – even suggesting that one might do without representations in some cognitive domains where the notion of representation is ill-defined – standardly leads to strong criticisms. Thus, by delineating these two domains, we hope to set up a better distinction between problems where a symbolic and representational approach seems justified –

the communicative domain – and problems that should be studied in their own terms without an a priori Fregean-based interpretation. Staying within a Fregean domain ultimately leads to grounding issues that cannot be solved by Fregean means.

6 Back to robotics

We started out by questioning the embodiment of robots and from here drifted into issues concerning biological grounding, the Fregean tradition and communication. In the previous section, this led to the development of a two tier view on cognition. In this final section, we want to sketch how these theoretical considerations feed back on the domain of robotics that provided our starting point. From our perspective, situated robotics, as a paradigmatic example of DESC, constitutes an amalgam of biological and Fregean aspects. In the following, we will give a short discussion as to how these two aspects could be teased apart and each clarified in its own way by applying this distinction between the biological and communicative domains. We will argue that robots can be used as valuable research instruments in both.

6.1 The biological tier

The biological tier deals with issues relating to the biological origins and aspects of cognition resulting in the the existence of biological agents. To stress once more, this does not mean that cognition is necessarily limited to biology, but that cognition as a natural phenomenon – that is cognition as far as we know it and up to now – is limited to biological cases. Thus, biological cases are to be studied now in order to get clear on cognition, and afterwards one can decide whether cognition is plausibly present in other than biological systems.

What are the relevant implications of this biological starting point for the use of robots? The first thing to note is that all implications will be held hostage to whatever theory of cognition will be developed within a strict biological context. It will also be important to note that such a theory should not depend on our tendency to ascribe an intentional vocabulary to a wide range of entities, including biological ones (Mar & Macrae, 2007; Scholl & Tremoulet, 2000). A biological theory of cognition should derive its notion of cognition from the kind of system or organization under description. The problem, of course, is that this work hasn't been done yet, or at least only in a very general and yet to be developed way (Lyon, 2006a; Maturana & Varela, 1980; Moreno et al., 1997; Stewart, 1996; van Duijn et al., 2006). Thus, the honest answer to the question at the beginning of this section is that we do not know what the implications are. Having said that, it seems also clear that a number of plausible implications can be formulated.

First, the focus should be on solving problems that do arise within this biological context, and more specifically the sensorimotor phenomena that take place. This would provide a criticism of, for example, work by van Dartel who developed several situated robotic models, with the aim to underline and develop a notion of *situated representation*, taking up the challenge of Haselager, de Groot & van Rappard (2003) to operationalize the notion of representation and make it useful within the context of situated robotics. However, while the notion of representation has a natural home in the communicative domain, it is not obvious why this concept needs to be introduced in the biological domain, without a specific problem that requires this particular kind of solution.

Second, it would be irrelevant – and rather a drawback than an advantage – whether the robot looks or acts like a recognizable agent to us. We are by nature equipped with the "tendency to innately, automatically, and spontaneously view a broad variety of different targets as holding goals and mental states" (Mar & Macrae, 2007, p.117-118). This tendency probably plays a key role in our own social behavior (Gallese, 2007), but it constitutes a masking device when it comes to getting a good view on the processes underlying the entity that evokes this ascription of intentionality in us (Keijzer, 2006). Thus, to be of use in the biological domain, robots should specifically target aspects of biological organization non-intentionally described but cognitively relevant.

What might those be? Going back to the requirement of organismic grounding (See section 2), a fundamental issue would be the autopoietic – self-constructing – characteristics of living systems, which at a metabolic level are continuously producing themselves, in this way remaining in existence over time. As Maturana & Varela (1980) took this process as being cognitive, work on fundamental self-construction – say with nano-robots – would be a key target for robotic research in the biological domain. However, there is also a tendency to interpret cognition in a more exclusive way, for example by restricting it to macroscopic – in comparison to the molecular processes of metabolism – sensorimotor phenomena (Cotterill, 2001; van Duijn et al., 2006). In this case, modeling the ways in which animals have acquired large-scale mobility – when compared to their single-cell building blocks – through the use of muscle sheets, nervous systems and perceptual feedback is a clear target for robot research. Another important domain would be for developing models for the way in which the external link between movement and sensors is used in such processes is also a definite target (Lungarella & Sporns, 2006). Finally, and more conventionally, would be the use of robots to gain a better understanding of the kinetics and dynamics of bodily movement and its coordination (e.g. Kuniyoshi et al., 2004,).

The above should give an impression of what might be important usages of robots within the biological domain. More extensive overviews of relevant robotic approaches can be found elsewhere (for example, Harvey, Di Paolo, Tuci, Wood and Quinn (2005), and Pfeifer, Iida and Bongard (2005)). What we hope to have indicated is that while within the biological tier the research issues derive from biological systems, this still leaves an important role for robotics.

6.2 The communicative tier

The communicative tier deals with phenomena involving communication, and the role that symbols play therein. The major difference with the biological tier is that in communication there is an external medium involved which is structured in a particular way, and which can be detached – in principle – from the communicating agents. The central example would be human language, which consists of externally available acoustic forms that can be written down or otherwise preserved. In addition, many other symbol systems have now been developed in this inter-human context, such as in logic and mathematics. Such independent symbol systems constitute the natural home of the Fregean tradition. Fregean-based forms of cognitive modeling can even be seen as the enterprise to make symbol systems jump through all kinds of hoops that are in some way derived from things that humans can do. While such work is important to get a grip on the kinds of things symbol systems can be made to do, it remains unclear to what extent this reflects or relates to human cognitive processing. As said above, human cognition seems to be an amalgam of both tiers where the biological one remains the great unknown. Given this situation, and the resulting complexities in the combination of the biological and communicative tiers, what are plausible targets for robotic studies within the communicative tier?

There are two self-evident domains in this context: human-robot communication and robot-robot communication. We discuss an example of both. Work on human-robot communication shows that the Fregean tradition, and its concepts of representation, comes naturally in a communicative context. Next to that, work by Steels and Vogt illustrates that communication can be validly viewed as a relatively autonomous domain with respect to the biological tier.

Roy et al. (Roy, Hsiao & Mavridis, 2004; Roy & Mukherjee, 2005) developed several robotic models and systems dealing with human-robot linguistic communication. One of them is a robotic system called *Ripley*. Ripley just looks like a single arm that can grip objects. However it can also recognize speech, interpreting its meaning by using visual information. The authors claim that adding this visual context grounds the speech recognizer in vision. An important concept in the Ripley robot is *mental imagery*, as a way to provide a stabilized conceptualization of the world which is not egocentric – strictly depending on the particular perspective of the system or agent – but allocentric – that is in world-related terms independent of any agent, (e.g. see Klatzky, 1998). According to Roy et al., mental imagery plays a key role in this encoding process:

> Language refers to the stabilized conceptualization of the world provided by mental models and imagery – we do not talk of objects as being in motion when we know that the apparent motion was caused by our own movements (Roy et al., 2004, p. 2).

Mental imagery, as understood here, can give rise to a conceptualization of the world that distinguishes between the influence of the robot's own actions and independent changes in the world, and then use this in action. Language needs to encode these conceptualizations. Roy et al. uphold that it is impossible to directly ground language in egocentric visual representations. A mediating allocentric *representational layer* is required to provide a stable view of the environment that can switch between first and third person view as to enable a speaker to imagine a listeners point of view and vice versa. This shared mental imagery between listener and hearer is a precondition for communication, according to Roy et al. In this way, this form of robotic research comes very close to the work of people like Glenberg and Kaschak (2002), Glenberg and Robertson (2000, 1999), and Deacon (1997). They also uphold that an intermediate representational layer is needed between visual data and language, that can objectify the visual data in such a degree that communication about it is possible. For present purposes, Roy et al.'s work illustrates both the comparative independence of the communicative domain, as well as the natural fit of Fregean ideas in this context.

It might be that the work of Roy et al. remains too close to human communication to warrant the conclusion that an allocentric representational layer is required for communication. A way to scrutinize this claim further is by turning to robot-robot communication. In addition to any symbolic structures involved in communication, communication has also a performative dimension, relating to whether communication is successful or not. This performative dimension implies that the system and its interaction with the environment constrain and provide criteria for success in a self-organizing manner (Steels & Vogt, 1997). One can thus focus on this interaction process and the requirements for successful communication that can be derived from them. This criterion can be studied by using robots engaging in successful communication with one another.

Luc Steels and Paul Vogt (1997) developed a robotic system – called the Language games experiment – where several robotic agents together created a proto-language in an autonomous way. This experiment had as goal to create a multi-robot environment where physically embodied robots could evolve and learn a rudimentary language within their own robot community, through so called *adaptive language games*.

One of the hypothesis of Steels and Vogt is that "*communication, if not full-fledged language, is a necessary stepping stone towards cognitive intelligence*" (Steels & Vogt, 1997, p.1). They ask the questions how to ground the evolving language in the sensorimotor reality of the robots, and, in connection, how meaning originates. They think the origin of meaning is a matter of "*how the distinctions that the robots lexicalize may arise in the first place*" (ibid.). In other words, how can a set of perceptual categories, a grounded ontology, arise in an agent, and how can this individually grounded ontology lead to a shared vocabulary in a group of agents.

Such a self-organizing, distributed, situated language game is a framework which provides a performative measure along which robots can evolve a rudimentary language. A framework like this is a good way to get insight in how a language like system could relate to a situated, sensorimotor physical base, even if this physical base is not fully biological in the sense as we have argued for above. Also, the fact that a rudimentary language can arise from a few simple constraints – without a thorough and fully understood embedding of its underlying sensorimotor system with the world – shows the relative self-sufficient nature of symbol systems.

All in all, within a communicative context there seem to be very good reasons to use representations that can be shared by different individuals. It will be more difficult to decide to what extent this is a precondition for communication, as the work of Roy et al. seems to imply, or rather a consequence of more elementary forms of communication.

7 Conclusion

Robots seem to be obvious candidates for investigating the wider array of cognitive features that are brought to the surface by dynamical, embodied and situated cognition. The problem that we discussed was that it may seem obvious that robots provide satisfactory models of embodied cognition, but that on closer inspection this becomes questionable. We discussed the biological grounding problem as brought forward by Sharkey and Ziemke (2000) and argued that when biological grounding is taken seriously, the best way to go would be a full biological interpretation of cognition. However, it is also clear that much of cognitive science operates in comparative independence of this biological foundation of cognitive phenomena. To explain this situation we sketched a Fregean tradition in which cognitive science and its sister disciplines are embedded, and which has its own, strongly independent interpretation of the cognitive domain. To circumvent a standoff between these two positions, we then argued that these positions correspond with two different cognitive domains, a communicative one, centered on symbol systems, and a sensorimotor one, centered on the biological organization required for generating full biological agents, capable of using any symbol system in the first place. The main problem with this latter domain is that the Fregean tradition has so far been extremely dominant and well-developed in the sciences dealing with cognition with the result that it has expanded into the biological sensorimotor domain as well. The two tiers view should help to become aware that the latter is a case of Fregean overstretch which should and can be remedied by treating the biological domain of cognition as a domain in its own right.

References

Anderson, M.L. (2003). Embodied cognition: a field guide. *Artificial intelligence*, *149*, 91–130.

Beaney, M. (1996). *Frege: Making sense*. London: Duckworth.

Beaney, M. (Ed.) (1997). *The Frege reader*. Oxford: Blackwell.

Beer, R. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, *72(1-2)*, 173–215.

Boden, M. (2006). *Mind as machine: A history of Cognitive Science*. Oxford: Oxford University press.

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, *47*, 139–159.

Brooks, R. (1999). *Cambrian intelligence: The early history of the new AI*. Cambridge, MA: MIT Press.

Chang, C. (2002). Biorobotics researcher: To be or not to be? *Behavioral and Brain Sciences*, *24(06)*, 1054.

Chiel, H. & Beer, R. (1997). The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences*, *20(12),* 553–557.

Chomsky, N. (1957). *Syntactic structures*. New York: Walter de Gruyter.

Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Clark, A. (2005). Intrinsic content, active memory and the extended mind. *Analysis*, *65(285),* 1–11.

Clark, A. (Forthcoming 2006). Pressing the flesh: A tension in the study of the embodied, embedded mind? *Philosophy and Phenomenological Research*.

Cotterill, R. (2001). Cooperation of the basal ganglia, cerebellum, sensory cerebrum and hippocampus: possible implications for cognition, consciousness, intelligence and creativity. *Progress in Neurobiology*, *64(1)*, 1–33.

Deacon, T. (1997). *The symbolic species: The co-evolution of language and the human brain*. London: Penguin books.

Dennett, D. (1987a). Eliminate the middletoad! *Behavioral and Brain Sciences*, *10*, 372–374.

Dennett, D. (1987b). *The intentional stance*. Cambridge, MA: MIT press.

Dennett, D. (1991). *Consciousness explained*. Boston, MA: Little, Brown and co.

Dennett, D. (1997). *Kinds of minds: Toward an understanding of consciousness*. New York: Basic Books.

Dretske, F. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT Press.

Dummett, M. (1991). *Frege and other philosophers*. Oxford: Oxford university press.

Fodor, J. A. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT press.

Frege, G. (1879). Begriffschrift. Reprinted in M. Beaney (Ed.) (1997). *The Frege reader*. Oxford: Blackwell.

Gallese, V. (2007). Before and below 'theory of mind': Embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B*, *362*, 659 – 669.

Gardner, H. (1984). *The mind's new science: A history of the cognitive revolution*. New York: Basic

Books.

Glenberg, A. M. & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9(3)*, 558–565.

Glenberg, A. M. & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, *28*, 1–26.

Glenberg, A. M. & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language*, *43(3)*, 379–401.

Harnad, S. (1990). The Symbol Grounding Problem. *Physica D*, 335–346.

Harvey, I., Di Paolo, E., Tuci, E., Wood, R., & Quinn, M. (2005). Evolutionary robotics: A new scientific tool for studying cognition. *Artificial Life*, *11(1-2)*, 79–98.

Haselager, P., de Groot, A., & van Rappard, H. (2003). Representationalism vs. anti-representationalism: A debate for the sake of appearance. *Philosophical psychology*, *16(1)*, 5–24.

Hooijmans, M. (2006). Robotics and representation: Explorations in terminology. Master's thesis, University of Groningen, Groningen.

Keijzer, F. A. (1998). Some armchair worries about wheeled behavior. In R. Pfeifer, B. Blumberg, J-A. Meyer & S.W. Wilson (Eds.), *From animals to animats 5* (pp. 13-21). Cambridge, MA: MIT Press.

Keijzer, F. A. (2001). *Representation and behavior*. Cambridge, MA: MIT press.

Keijzer, F. A. (2006). Differentiating animality from agency: Towards a foundation for cognition. In R. Sun, N. M. (Ed.), *Proceedings of CogSci/ICCS 2006*, (pp. 1593–1598)., Alpha, NJ. Sheridan Printing.

Kenny, A. (1995). *Frege: An introduction to the founder of modern analytic philosophy*. Oxford: Blackwell.

Kirsh, D. (1991). Today the earwig, tomorrow man? *Artificial Intelligence*, *47(1)*, 161–184.

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In C. Freksa, C. Habel, & K. F. Wender (Eds.), *Spatial cognition* (pp. 1-17). Berlin: Springer.

Kuniyoshi, Y., Ohmura, Y., Terada, K., Nagakubo, A., Eitoku, S., & Yamamoto, T. (2004). Embodied basis of invariant features in execution and perception of whole body dynamic actions: Knacks and focuses of roll-and-rise motion. *Robotics and Autonomous Systems*, *48(4)*, 189–201.

Lungarella, M. & Sporns, O. (2006). Mapping information flow in sensorimotor networks. *PLoS Computational Biology*, *2(10)*, 1301–1312.

Lyon, P. (2006a). *The agent in the organism*. PhD thesis, Australian National University, Canberra.

Lyon, P. (2006b). The biogenic approach to cognition. *Cognitive Processing*, *7(1)*, 11–29.

Mar, R. & Macrae, C. (2007). Triggering the intentional stance. *Novartis Foundation Symposium*, *278*, 111–20.

Maturana, H. R. & Varela, F. J. (1980). *Autopoiesis and cognition: The realisation of the living*. Dordrecht: Reidel.

McCarthy, J. & Hayes, P. (1969). Some philosophical problems from the standpoint of artificial intelligence. *Machine Intelligence*, *4*, 463–502.

Mendelsohn, R. L. (1996). Frege's treatment of indirect reference. In M. Schirn (Ed.), *Frege: Importance and legacy* (pp. 410–437). New York: Walter de Gruyter.

Moreno, A., Umerez, J., & Ibanez, J. (1997). Cognition and life: The autonomy of cognition. *Brain & Cognition*, *34*(1), 107–129.

Nilsson, N. (1984). Shakey the Robot *SRI AI Center Technical Note 323*.

Parsons, T. (1996). Fregean theories of truth and meaning. In M. Schirn (Ed.), *Frege: Importance and legacy* (pp. 371–409). New York: Walter de Gruyter.

Pfeifer, R., Iida, F., & Bongard, J. (2005). New robotics: Design principles for intelligent systems.

Artificial Life, *11(1-2)*, 99–120.

Pfeifer, R. & Scheier, C. (2001). *Understanding Intelligence*. Cambridge, MA: MIT press.

Reed, E. S. (1997). *From soul to mind: The emergence of psychology from Erasmus Darwin to William James*. New Haven, CT: Yale University Press.

Roy, D., Hsiao, K., & Mavridis, N. (2004). Mental imagery for a conversational robot. *IEEE Transactions on Systems Man and Cybernetics Part B (Cybernetics)*, *34(3)*, 1374–1383.

Roy, D., & Mukherjee, N. (2005). Towards situated speech understanding: Visual context priming of language models. *Computer Speech and Language*, *19(2)*, 227–248.

Russell, B., & Whitehead, A. N. (1913). *Principia mathematica*. Cambridge: Cambridge university press.

Schirn, M. (Ed.). (1996). *Frege: Importance and legacy*. New York: Walter de Gruyter.

Scholl, B. & Tremoulet, P. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, *4(8)*, 299–309.

Searle, J. (1980). Minds, brains and programs. *The Behavioral and Brain Sciences*, *3*, 417–457.

Shannon, C. & Weaver, W. (1949). *Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.

Sharkey, N. & Ziemke, T. (2000). Life, mind, and robots: The ins and outs of embodied cognition. In S. Wermter & R. Sun (Eds.), *Hybrid Neural Systems* (pp. 313–332). Berlin: Springer.

Sharkey, N. & Ziemke, T. (2001). Mechanistic versus phenomenal embodiment: Can robot embodiment lead to strong AI? *Cognitive Systems Research*, *2(4)*, 251–262.

Steels, L. (2003). Intelligence with representation. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, *361(1811)*, 2381–2395.

Steels, L. & Vogt, P. (1997). Grounding adaptive language games in robotic agents. *Proceedings of the Fourth European Conference on Artificial Life*, 474–482.

Stewart, J. (1996). Cognition = life: Implications for higher-level cognition. *Behavioural processes*, *35*, 311–326.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59(236)*, 433–460.

van Dartel, M. F. (2005). *Situated representation*. PhD thesis, University of Maastricht, Maastricht.

van Duijn, M., Keijzer, F., & Franken, D. (2006). Principles of minimal cognition: Casting cognition as sensorimotor coordination. *Adaptive Behavior*, *14(2)*, 157–170.

van Emde Boas, P. & Janssen, T. M. V. (1979). The impact of Frege's principle of compositionality for the semantics of programming and natural language. Technical report, Department of Mathematics, University of Amsterdam.

van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and brain sciences*, *21*, 615–665.

von Neumann, J. (1945). First Draft of a Report on the EDVAC. Technical report, University of Pennsylvania.

Webb, B. (2001). Can robots make good models of biological behaviour? *Behavioral and Brain Sciences*, *24(06)*, 1033–1050.

Wittgenstein, L. (1922). *Tractatus logico-philosophicus*. London: Routledge.

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.

Ziemke, T. (1999). Rethinking grounding. In A. Riegler, M. Peschl & A. von Stein (Eds.), *Understanding representation in the cognitive sciences: does representation need reality?* (pp. 177–190). New York: Plenum.

Ziemke, T. (2001). Are robots embodied? In C. Balkenius, J. Zlatev, C. Brezeal, K. Dautenhahn & H. Kozima (Eds.), *Proceedings of the First International Workshop on Epigenetic Robotics, volume 85* (pp. 75–93). Lund: Lund University Cognitive Studies.