Embodiment is meaningless without adequate neural dynamics

Luc Berthouze and Adriaan Tijsseling
Cognitive Neuroinformatics Group; Neuroscience Research Institute
AIST Tsukuba Central 2; Umezono 1-1-1
Tsukuba 305-8568, Japan

Abstract

Traditionally, cognition has been considered as a "mental processes" only domain. Recently, however, there is a growing consensus that cognition should be "embodied", i.e. it emerges from physical interaction with the world through a body with given perceptual and motoric abilities. Terms like "emergence", "enaction", "grounding", and "situatedness" are often used, but little attention is being paid to actually understanding the neural dynamics correlates of an emergence of cognition. Nor is hardly being investigated how the structure of the body-environment coupling is perceived and manipulated by our brain. It is as if talking about neural dynamics would somehow throw us backwards to the old cognitivist approach. In this position paper we present a balanced view, in which we try to keep things in their respective place.

Introduction

Traditionally, cognition has been defined as the collection of mental behavioral capacities (e.g. language and problem solving). As such, research into these kinds of capacities has focused on the underlying mental processes. Simply stated, any computational modeling experiment would take a set of stimuli to use as input to either an algorithm, a program or a set of rules, and then analyze the output. A "good" performance will lead the researcher to interpret real cognitive behavior in terms of the devised model. This approach may have led to several interesting findings (Goldstone's work on perceptual learning (1998), Kruschke's categorization model (1993)), but so far computational models have not matured beyond being rather simple toy models.

The most serious problem they suffer is which has been coined "symbol grounding" by Harnad (1990). It is one thing to have a sophisticated computer model that can recursively parse complex sentences, but it is another thing to have a program that actually understands the meaning of these sentences. According to Harnad (1995), the model's internal representations (be it categories, memories, or propositions) have to be grounded, either directly or indirectly, in sensorimotor interactions with the environment (this environment can really be anything such as the world we live in or a virtual reality scene).

To solve the grounding problem, it has been suggested by various researchers (e.g. Brooks, 1990; Varela et al., 1991; Pfeifer & Scheier, 2000) that a cognitive system has to have a physical body, and be situated in an environment from which it gets sensory information and in which it can perform actions and interact with other sentient beings. This interaction between body and environment is assumed to be structured and subject to evolution and the system's own morphological growth.

This does, unfortunately, still not address the problem of grounding meaning. It is assumed that the system is now possible to "capture" the meaning of a perceived stimulus, e.g. a red cube. But why would this be the case? We have a body, an environment, and some structured interaction between them. What we do not know is how this meaning is supposed to emerge from the transient collective activity of the system's internal set of neurons. There is some kind of structure in the interaction of an agent with a red object, and therefore, presumably, there is some kind of structure in the state of the neurons reflecting the ongoing sensorimotor activity when perceiving and manipulating the red object. Does this mean that as a result the meaning "red object" just emerges in the system? Not at all, because we still have no knowledge about the structure of the system's internal dynamics and the nature of the mutual interaction between body, environment, and internal dynamics.

It might even be the case that meaning is an epiphenomenon, something that arises as a side-effect of the system being able to act on what it has learned and perceive the results of its own actions. In other words, perhaps the agent is able to act on a red object, just because it happened to develop the appropriate stimulus-action response for it. Whatever meaning may be, before we can even address questions related to this, it is a necessity to first address the role of the internal dynamics on the problem of embodiment.

What is embodiment?

We define embodiment as the dynamic interactive coupling between brain, body and environment. By coupling, we mean that the body is subjected to the physical laws of the environment and that body and environment (including agents living in it) are mutually connected. Environmental changes affect the body and vice-versa.

There are at least two aspects of embodiment. Morphology has been the focus in research. It refers to the structure of the system's body and it involves many factors such as, the number of degrees of freedom, the elasticity of skin, the distribution and location of sensors (a lower resolution of tactile sensors in the back vs. a very fine resolution at the fingertips), the synergy of limbs, reflexes (explained in a later section), redundancies between degrees of freedom (horizontal motions of the neck and eyes, left and right arms, capacity to use the elbow to replace the hand in some tasks, etc.). As an illustration, the human hand contains several postural configurations as a result from having opposable thumbs, and as a consequence, many various manipulations are possible, independent of the neural circuitry that controls it. Wallace, Weeks & Kelso (1990) show the many functions of hand as special-purpose device:

Passive	Carrier, scooper, pusher
Percussive	Tapper, clapper, pounder
Expressive	Pointer, communicator
Exploratory	Groomer, toucher
Prehensile	Reacher, grasper

From a developmental point of view, body growth also affects the neural circuitry. Thelen & Smith (1994) conducted experiments showing the effect of growth-related changes in body fat and muscle strength on infants' stepping speed on a treadmill. Such changes are shown to require a "rescaling" of the system.

The second aspect has not been closely investigated by cognitive scientists. It refers to the fact that having a morphology casts constraints (necessary transformations) on the functioning or development of the neural circuitry that controls it (e.g. the redirection in blind mole rat of auditory input via lateral geniculate nucleus to the visual cortex causing it to function as an "auditory cortex" (Doron & Wollberg, 1994)). From a behavioral point of view, this is nicely demonstrated with an experiment by Good (1999). Good showed that binaural recordings performed by blind people and sighted people yield considerably different results when played back to other blind people. The recording done by sighted people was shown to be "flat", whereas recordings from blind people captured depth information. This was caused by the fact that, while sighted people look straight ahead, blind people actually slightly rotate their head while walking, so they receive differential auditory information.

Several other experimental evidences in neuroscience push us to a more balanced view of the respective roles of morphology (as anatomical structure) and neural dynamics. Kitazawa, Kimura & Uka (1997) show that even Japanese macaques (macaca fuscata) show complete deadaptation to prisms after 4 or 5 trials in a pointing task. In a prism adaptation experiment, the subject is fitted with prismatic lenses which change its visual perception of the environment so that for a same target position, a lateral displacement will be perceived by the subject, thus requiring an end-effector correction. More significantly, Ingram et al. (2000) study deafferented subjects in visuomotor adaptation tasks and conclude that proprioception is not an absolute requirement for adaptation to occur: Deafferented subjects were able to adapt, mainly through visual attention.

We conclude that body structure alone is not sufficient to understand a complex concept as embodiment. The neural circuitry is a not lesser part of the coupling between body and environment.

Hypotheses on the development of cognition

We make the following hypotheses on embodied sensorimotor development: (1) Coordinated behavior would emerge from interaction with the world, in particular from spontaneous motion (or self-exploration), through the body, inborn reflexes, and acquired behavior. This interaction constitutes a closed loop between environment, the agent and other agents. It starts with being born with a given body morphology and innate reflexes, immediately engaging in complex sensorimotor interaction with the world, acquiring novel behaviors along the way, which in turn help in further exploration of the world. (2) Categorization of acquired coordinated behavior (Thelen and Smith, 1994) results in categorical responses to novel sensorimotor patterns obtained from the system's interaction with the world. This enables the system to learn about basic sensorimotor couplings (body couplings and body-environment couplings). Goldfield (1995) makes the following proposition: *Exploratory activity reveals stable regions in dynamic geometries*. Infants explore the dynamics of their own action and may discover information specific to stable regions in the high-dimensional spaces of possible actions. With further exploration, they can refine those couplings and hence this would lead to the emergence of stable sensorimotor configurations over time. For example, consider the grab reflex of newborns. There are many ways to trigger this grab reflex, but they all induce the one grab reflex category. However, with increased interaction with the environment – grabbing various types and kinds of objects, dropping objects, moving objects – the infant acquires new sensorimotor couplings built upon the grab reflex. And with learning each new behavior, combinations of previously learned sensorimotor categories can lead to the emergence of more complex behavior. Ultimately, basic sensorimotor categories provide useful functionality by creating compact, bounded chunks that may then eventually be combined into higher-order categories through e.g. language (Harnad, 1996; Tijsseling, 1998). (3) Entrainment: Dynamically, perception and action would have a natural tendency to approach one of the potential attractors, whose low-dimensional representations are the categories above. A typical resulting behavior would be the *circular* repetitive actions of Piaget (1945). As suggested by Goldfield (1995), performing an act or perceiving that same act would reveal attractor dynamics and parameterizations which are potentially available to be detected by the perceptual organs. The attractors put the actor/perceiver into the basin of attraction for a class of activity (i.e. emergence of imitation). Exploration of the attractor dynamics then makes it possible to better approximate a stable or optimal form of the act.

The above three factors contribute to the scaffolding process during cognitive development. With scaffolding we refer to the support given to a system in order for it to construct and extend skills to higher levels of competence. Scaffolding is similar to scaffolding around a building. It can be taken away after the need for it has ended. For example, after acquiring coordinated sensorimotor skills from grabbing, moving, or throwing objects, the grab reflex tunes out of the system.

Even though scaffolding should provide the system with resources to develop a structured interaction with the system, we should not forget that the above relies heavily on the presence and the delicate balance of stability and plasticity in the system (see the stability-plasticity dilemma (Grossberg, 1982; Murre, Phaf & Wolters, 1998), usually ascribed to neural dynamics.

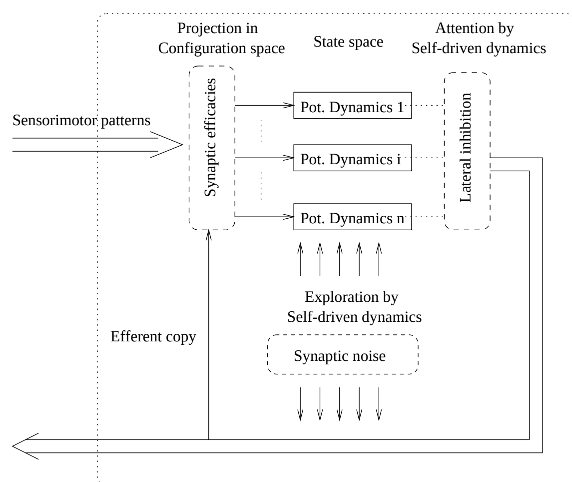
Two major question remains. Why is it that the system can actually develop? and why doesn't it just degenerate or collapse (as it happens in many dynamical systems)? The system should be adaptive to changes in interaction (and/or morphological growth), sensitive to instabilities in interaction dynamics while preserving and maintaining the integrity of the pre-existing set of behaviors and building an increasingly complex repertoire of behavior. Stability should emerge at the behavior level in order to have a coordinated set of behaviors, but at the micro-level, the system should perhaps be in a permanent instability, in order for it to stray around trivial states.

To summarize, even if we have a proper scaffolding development process going, it does still not guarantee the system actually develops and in the right direction. Research on morphology and the structure of the interaction between body and environment easily makes the internal dynamics of the system move out of focus. Imagine trying to construct a nice humanoid robot without a "brain", only the necessary internal wiring process sensory information and control motoric actions, for example, the AI-life insectoid robots. Presumably, people will try to interact with such a robot and perhaps even ascribe some intentional mental states to it, but will eventually habituate soon it ceases activity. It could be improved by adding spontaneous motion, but one should ask oneself, will the system build upwards from just that? It is what is in between the robot's ears, so to speak, that is also contributing to the development of cognition. The combination of and interaction between body, internal dynamics, environment as a whole determines what the system will and can learn. The relation between body, environment, and internal dynamics should be regarded as mutual: They provide each other with the necessary constraints to (1) balance stability and plasticity, (2) map the high-dimensional sensory and motoric information onto a far lower dimensionality, (3) extract the invariants from the environment to construct the necessary sensorimotor categories and build upwards from this to more complex categories and behavior.

Implications on neural dynamics

The above considerations are mainly behavioral. From a behavioral perspective, developing means that the system is acquiring skills (the meaning of which is attributed to it by external observers), which are then used as building blocks in the construction of a more elaborate repertoire of behaviors and skills. Let us call this the viewpoint of the external observer. Next, let us consider the viewpoint of the internal observer instead: What, within the brain-body compound, is attributing "meaning" (if any) to the behavioral repertoire?

Before we answer this question, let us stress that we have no intention to re-introduce the notion of a homunculus. We are not talking about some entity inside the brain that is making "remote observations" on its activity. If there is any such observation, then it is involved in the dynamics of "doing" (i.e. the entire cycle from processing sensory information to producing motoric commands). In other words, we adhere to the principle that, even in the brain, perception and action are inseparable. They form a perfectly meshed whole, in which the one cannot exist without the other. This notion of an "internal observer" is taken in the quantum-mechanic sense (Wheeler, 1983). Consequently, how do we interpret the stability-plasticity dilemma in terms of the above notion of an "internal observer"? One of the authors (Berthouze, 2000) introduced the filter hypothesis as depicted below.



The filter hypothesis is an attempt to attribute the notion of an internal observer, which by filtering, categorizing and generating action, is involved in the dynamics by *observing* it. It bridges external interactive dynamics and internal dynamics by entrainment dynamics.

It comprises of several dynamics. One dynamics is the spontaneous exploration with which the system acquires a structured lower-complexity sensorimotor space, shaped by its natural body dynamics and its reflexes. Because the incoming flow of sensorimotor patterns is structured, spatio-temporal regularities can be extracted and synaptic efficacies adapted. The synaptic configuration, in turn, projects (filters) the incoming flow into a new space in which multiple coordinative structures can coexist. Goldfield (1994) also suggests that coordinative structures with attractor dynamics emerge from spontaneous dynamics. As a consequence of producing variable trajectories under specific constraints, the microscopic components of actions become assembled into coordinative structures which exhibit attractor dynamics (a process of self-organization). Collective variables exhibit entrainment, whereby the system enters preferred stable states and exhibit abrupt phase transitions, both being characteristics of spontaneous activity in infants.

Within the environment-body-brain coupling, the system's sensorimotor space will not only be shaped by reflexes and acquired behavior but also by internal self-driven dynamics. An example of self-driven dynamics are the poissonian spikes in cortical areas (Canepari et al.,1997; Collins et al.,1995; Wiesenfeld & Moss, 1995). In cortical areas, the system can adopt stable recurring states of activity (e.g. synchronized firing), lasting up to several seconds, with stationary probability of transitions among them. This mode of synchronized firing is not consistent with an integration of excitations by cells between synchronized firing events. Instead, it seems to be triggered as a threshold phenomenon by coincidence of random inputs. This creates instability in the sensorimotor space, such that a search around local stable states is maintained. This will help the system to avoid dead-stable or trivial states.

Noise, which is normally considered an obstacle to information processing, can actually be treated as an enhancer of information through its resonance with external signals (Tijsseling & Berthouze, 2001). It might then be seen as a catalyst for "attentional processes" (Berthouze, 2000). External behavior is never exactly the same. But similarity can be defined if one focuses on some invariant features. *Attention* may in part be just an attribution by an external observer to the dynamics of the system. Without selective filtering by the weights, the system cannot achieve prototypical response to novel patterns, being unable to focus on a particular feature.

Another dynamics is the categorization discussed earlier. During category learning, attractors in the state-space of the network will emerge. A particular sensorimotor coupling will be pulled into an attractor by the filtering mechanism defined by the system's neural wiring. Categories can then be defined as singularities in the continuous flow of sensorimotor data. As such, categories are not static, off-line representations but transient states in the state space. They are activated as the system is involved in it (similar to point attractor dynamics in memory models (Tsuda, 1992)). The category is cued by similar instances, which can be a sensory stimulus, some self-driven dynamic exploration, or a performed action.

If both the adaptation and the dynamics itself have similar time scales, then the overall behavior becomes non-trivial and emergent. Because as the neural system learns, it results in changes of the whole dynamics, which in turn is learned by the neural system. Thus, the problem of stability versus plasticity is considered at both micro (the neural level) and macro level (the behavioral level). Stability should emerge at the behavior level in order to have a coordinated set of behaviors, but at the micro-level, the system should perhaps be in a permanent instability, in order for it to stray around trivial states.

Concluding remarks

We discussed the scaffolding process that occurs during development and how it can be constrained by, on particular, morphology and the structure of the environment. Then we argued that within the environment-body-brain coupling, the role of the internal dynamics cannot be forgotten. After all, the meaning of acquired categories strongly depends on how the internal dynamics functions. But then, the structured development and the dynamic cognitive processes brings us to the question how stability and plasticity can be balanced. Moreover, the development processes that expand a system's knowledge, experience, and sensorimotor couplings, lead us to wonder why such a system would not collapse at some point. Why does the addition of, for example, a new category not make the entire or a large part of the category system instable or render it unusable?

We don't know how. But we are convinced the notion of complexity is a crucial factor in addressing this issue. Both physical and neural complexities must be equally taken into account (even if past work might have wrongly emphasized on the sole importance of neural complexity). The complexity of the body has various implications. At the behavioral level, it may support richer interaction dynamics. While making local control routines more difficult, it actually helps structure the flow of sensorimotor patterns, therefore reducing the dimensionality of the sensorimotor space that has to be learned. Hence, a more complex body does not necessarily make control more difficult but it definitely pushes for new dynamic-based, adaptive control paradigm (which is much closer to biological reality).

The structure of the body determines what can and cannot be learned by the system. A Martian might have a different kind of perceptual apparatus and perceive the same stimulus in a radically different way. In other words, the physical structure of the stimulus is the same, but the way it is reconstructed within a cognitive system, or which aspects of information are extracted from it by the system, might differ across species (Tijsseling, 1998).

As for the complexity of the neural system, we suggest to shift the focus of research onto the "structure" of the complexity of the neural system to be developed/studied in the sense of the difference between EOC (edge-of-chaos) systems and HOT (highly-optimized-tolerance) systems (Doyle, 1998). While there is no question that the human brain which is a very complex structure, can exhibit complex dynamics, even a simple cell-assembly of three chaotic neurons can give rise to complex (chaotic) dynamics (and will be characterized as a

EOC system). At the behavioral level, we will observe high instability which might not be very compatible with the nominal simplicity (and coherence) of human behaviors. Meanwhile, a very complex system will not necessarily produce complex behaviors but exhibit properties which are highly desirable and present in human system: Nominally simple yet high performance in an uncertain environment (but for which it has been designed by evolution), operating at densities (depending on the time scale considered, it might not always be true) above the standard critical point but vulnerable to design flaws and unanticipated changes in the external conditions. Indeed, a change in the morphology of the brain or in the body can collapse the stability of the agent's behavior.

References

Berthouze, L. (2000). *Robot Learning - An Interdisciplinary Approach*, chapter Bootstrapping the developmental process: The filter hypothesis, pages 8–30. World Scientific.

Brooks, R. (1990). Elephants don't play chess. *Robotics and Autonomous Systems*, 6:1–16.

Canepari, M., Bove, M., Maeda, Capello, M., & Kawana, A. (1997). Experimental analysis of neuronal dynamics in cultured cortical networks and transitions between different patterns of activity. *Biological Cybernetics*, 77:153–162.

Collins, J., Chow, C., & Imhoff, T. (1995). Stochastic resonance without tuning. *Nature*, 376(20):236–238.

Doron, N., & Wollberg, Z. (1994). Cross-modal neuroplasticity in the blind mole rate *Spalaz ehrenbergi*. *NeuroReport*, 5:2697–2701.

Doyle, J. (1998). *Complexity and Robustness*. Tokyo University, Oral communication.

Goldfield, E.C. (1995). *Emergent forms: Origins and Early Development of Human Action and Perception*. New York: Oxford University Press.

Goldstone, R.L. (1998). Perceptual learning. *Annual Review of Psychology*, 49:585–612.

Good, D. (1999). *Humane Interfaces: Methods and Practice in Cognitive Technology*, chapter On why the blind leading the blind is a good idea. Amsterdam: North Holland.

Grossberg, S. (1982). *Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control*. Boston, MA: Reidel Press.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42:335–346.

Harnad, S. (1995). *The Artificial Life Route to Artificial Intelligence: Building Situated Embodied Agents*, chapter Grounding symbolic capacity in robotic capacity, pages 277–286. New Haven: Lawrence Erlbaum.

Harnad, S. (1996). *Communicating Meaning: Evolution and Development of Language*, chapter The origin of words: A psychophysical hypothesis, pages 27–44. NJ: Erlbaum.

Ingram, H.A., van Donkelaar, P., Cole, J., Vercher J-L., Gauthier, G.M. & Miall, R.C. (2000). The role of proprioception and attention in a visuomotor adaptation task. *Experimental Brain Research*, 132:114–126.

Kitazawa, S., Kimura, T., & Uka, T. (1993). Prism adaptation of reaching movements: Specificity for the velocity of reaching. *Journal of Neuroscience*, 17:1481–1492.

Kruschke, J.K. (1993). Human category learning: Implications for backpropagation models. *Connection Science*, 5:3–36.

Murre, J.M.J., Phaf, R.H., & Wolters, G. (1992). CALM: Categorizing and Learning Module. *Neural Networks*, 5:55–82.

Piaget, J. (1945). *La Formation du Symbole chez l'Enfant*. Paris: Delachaux et Niestl.

Pfeifer, R., & Scheier, C. (2000). *Understanding Intelligence*. Cambridge, MA: MIT Press.

Thelen, E. & Smith, L. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.

Tijsseling, A. (1998). *Connectionist models of categorization: A dynamical approach to cognition*. Doctoral dissertation, Southampton University, Southampton, England.

Tijsseling, A. & Berthouze, L. (2001). A neural network for temporal sequential information. Submitted to the *8th International Conference on Neural Information Processing*.

Tsuda, I. (1992). Dynamic link of memory – chaotic memory map in nonequilibrium neural networks. *Neural Networks*, 5:313–326.

Varela, F., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Wiesenfeld, K. & Moss, F. (1995). Stochastic resonance and the benefits of noise: From ice ages to crayfish and squids. *Nature*, 373(5):33–36.