A theoretical framework for the study of spatial cognition

Maurizio Tirassa*, Antonella Carassa** & Giuliano Geminiani*

* Università di Torino — Centro di Scienza Cognitiva
via Lagrange, 3 — 10123 Torino (Italy)
email tirassa@psych.unito.it, geminian@psych.unito.it

** Università di Padova — Dipartimento di Psicologia Generale
via Venezia, 8 — 35131 Padova (Italy)
email carassa@psico.unipd.it

1. INTRODUCTION

The ultimate goal of this research is to contribute to the understanding of spatial cognition and its consequences on locomotion (that is, on the organization of the whole organism's movements in space). An organism's locomotion depends on what structure it superimposes on space itself and can therefore be understood as a form of interaction with a subjective environment, understandable in turn in terms of the organism's cognitive architecture.

As a first step, we will propose here a large-scale classification of the cognitive architectures possible, outlining the subjective structure that each of them superimposes on space and the relevant consequences on locomotion. Our classification differs from others that have been proposed for spatial cognition (e.g., Papi 1990) in that it does not build on the idea of *spatial behaviors*. We will argue, first, that cognition is better understood as interaction rather than behavior; and, second, that an organism's interactions can only be understood as generated and controlled by its cognitive architecture. This corresponds, in a sense, to the adoption of the organism's (rather than the observer's) subjective point of view.

The main division we will draw is between those architectures whose internal dynamics are entirely coupled to the dynamics in the world, and those that have at least some capability of decoupling. The latter correspond to representational architectures. Each class will be further decomposed according to a criterion of complexity. Our view of representation will not build upon a computationalist account of cognition.

2. ADAPTIVITY AND INTERACTION

Adaptivity is a living organism's capability of creating and maintaining a dynamic compatibility with its environment. The notion of adaptivity is circular, in that it involves neither the organism alone nor the world alone, but the interaction between the two. Each (type of) organism thus entertains specific (types of) interactions with a specific environmental niche.

The world has dynamics of its own which a living organism has to cope with if it is to survive and reproduce. From this point of view, adaptivity is an organism's capability of maintaining its own structure in the face of perturbing environmental dynamics. On the other hand, what is to be considered as environmental dynamics depends on what specific organism is considered. From this point of view, adaptivity is an organism's capability of "creating", or "viewing", a subjectively relevant set of environmental dynamics to cope with.

Adaptivity builds upon compatibility rather than correctness or optimality. In the case of a representational organism, for example, the point is not whether its representations faithfully mirror the objective reality of the outside world; nor has it any means to find out whether they do. What matters is instead whether its representations are *compatible* with reality, that is, if their dynamics result in the maintenance of the organism's capability of coping with the world. Chimpanzees are incapable of representing abstract causal links between objects or events and therefore of formulating theories of the world (Premack, Premack & Sperber eds. 1995); nonetheless, their mind is adapted to the subjective environment they live in. Humans are capable of theorizing (Geminiani, Carassa & Bara 1996), but not of representing the sonar profile of a moth like bats do, and so on. And, as we will argue, many species do not entertain representations at all and have simpler ways to maintain their compatibility with their environment.

Thus, while, in a sense, all species share the same "objective" world, each of them may also be said to live in a subjective one of its own, which, of course, may more or less resemble that of phylogenetically related species. Compatibility may then be viewed as a species' capability of capturing those variants and invariants in the world that are relevant for that species' interactions with it; and adaptivity may be viewed as the species' capability of maintaining compatibility.

Many different forms of adaptivity may be conceived of. Therefore, on the one hand, when describing nonhuman species, we should avoid the anthropocentric fallacy of conceiving of them as simply representing a greater or smaller subset of what our species is able to represent. On the other hand, just because, say, insects are unlikely to entertain representations doesn't mean that representations do not exist in more sophisticated species.

It is a consequence of this picture that it may be misleading to study adaptivity in terms of *behaviors*, if the term is taken to refer to factual descriptions of what organisms objectively do in an objectively defined world. Behavior is in the (representational) observer's eye only, not in the organism observed: what organisms do is not to behave, but to interact with their subjectively defined environment. It is more appropriate to study interaction in terms of the structure that generates and controls it. We call this structure the species' *cognitive architecture*.

3. COGNITIVE ARCHITECTURES AND NEUROBIOLOGY

The cognitive architecture of an organism is the functional architecture of its nervous system, that is, a high-level description of the properties and of the aspects of the functioning of its nervous system that are relevant for and causally generate its interactions with the world. In the case of a representational species, by definition, these coincide with its mind.[1]

There are two reasons why we focus on cognitive architecture rather than on neurobiology alone. The first is that the study of the nervous centers that deal with, say, landmarks in a certain species needs a parallel study of what a landmark is to that species, which is in turn part of the study of that species' cognitive architecture. The second is that, in an evolutionary perspective, what is selected for or against is not a nervous system as such (except, of course, for what concerns susceptibility to pathological events like diseases or traumata), but the adaptivity of the interaction that it is able to generate; and this regards again the cognitive, rather than the strictly neurological, architecture of a species.

---

[1] In this perspective, the term *cognition* and its derivatives should have been kept for the species that possess a mind, that is, those that entertain representations. On the other hand, the term is widely used in the relevant literature (e.g., Maturana & Varela 1980) and most alternatives are not less ambiguous from a philosophical point of view (*control architecture*, for instance, is typically used in autonomous robotics and would have therefore been inappropriate to the discussion of biological entities, let alone representational ones).

In the next sections we will sketch some types of cognitive architectures, that is, some different types of solutions to the problem, faced by every active species, of how to adaptively generate and control the interaction with the environment. For the moment, let us consider one property of cognitive architectures, namely, their innateness.

An organism should not be viewed as just cast into the world, a stranger in a strange land: it has instead to be born prepared for the interaction with the niche it will find itself in. In lower organisms, whose life span is too brief and nervous system too simple to allow for individual differentiation or learning, this may mean that each architectural component has to be completely developed from start. In general terms, however, it means rather that the possible modifications that an architecture may physiologically undergo are implicitly defined in the architecture itself (Barkow, Cosmides & Tooby eds. 1992; Cosmides & Tooby 1994; Lorenz 1965). Some architectures may be more rigid and some more flexible, so to let each individual follow its own developmental trajectory, according to the particular interactions it has with the environment; but, in any case, the space of possible developmental trajectories is intrinsic to the initial state of the architecture and is therefore a species-level property, and an adaptive one.

The innate endowment of a species thus determines not the specific interactions that each of its members will entertain with the world, but the whole space of that species' possible interactions with its subjective environment; the complexity of such space varies in accordance with the complexity of the species' architecture. *Learning* is not a natural kind, but the innate capability of a cognitive architecture of undergoing specific types of modifications, possibly triggered in part by specific types of interactions with the subjectively defined environment.

To say that cognitive architectures are innate also means that they, like most biological traits, are the product of evolution. Although adaptivity may be viewed as the property of a whole species as well as of each of its members, natural selection ultimately operates upon the slight individual variations existing between the latter. The phylogeny of cognitive architectures is therefore a side effect of the differences in the innate endowments of the individuals that make up a species (due to statistical differences in the species' genetic pool as well as to mutation, recombination, etc.), plus the differences in their respective reproductive success.

To resume: the living organisms we are interested in are those that engage in active interactions with their subjectively construed environment by way of self-organization (Maturana & Varela 1980; Varela, Thompson & Rosch 1991). The structure which governs an organism's interaction with the environment, thus maintaining that organism's adaptivity, is its innate cognitive architecture. The cognitive architecture of a species defines the subjective structure its members will superimpose on the world. Our position so far may therefore be described as a Kantian version of constructivism.

In the next sections, we will sketch some types of cognitive architectures, that is, some ways in which an organism's internal dynamics may co-evolve with the subjectively relevant dynamics in the environment so to generate an adaptive interaction. Each type will superimpose a specific type of subjective structure on space, which will have relevant consequences on locomotion. The main division we will draw is between those architectures whose internal dynamics are entirely coupled to the dynamics in the world, and those that have at least some capability of decoupling their internal dynamics from the external ones. The latter correspond to representational architectures. From a neurobiological point of view, we expect this division to mirror the division between species whose nervous system has no proencephalic differentiation and those whose nervous system has at least some. These two main classes will be further decomposed into subclasses according to a criterion of complexity.

4. COUPLED ARCHITECTURES

The internal dynamics of these cognitive architectures are entirely coupled to the external ones. These organisms have no internal model of their environment and are therefore only capable of external cognition: to them, the world is the only possible model of itself.

Since concepts are the active constructions of a representational mind which superimposes its own a priori categories on the world, coupled architectures have no concepts of any sort. This implies that their subjective environment does not build upon the existence of *objects*. To say

that coupled architectures have no object-based construction of space refers to something far more primitive than object permanence. As we will argue later, the latter term refers to an organism's capability of realizing that objects exist even when they are out of immediate perception, and therefore of recognizing them as being the same in different presentations. The level logically antecedent to object permanence is object impermanence: the difference, however, relates to the type of representation entertained by an organism, rather than to whether that organism entertains representations of any sort. Coupled architectures, instead, have no representations at all, so that the point here is not whether objects are or are not permanent, but simply whether they exist.

4.1 Reflex-based architectures

The simplest types of coupled architectures are only composed of reflexes. The internal dynamics of a reflex-based architecture depend exclusively on the external ones. Its interactions may be viewed as a set of fixed stimulus/response patterns, and are therefore completely driven by the environment: there are no internal states relevant to the interaction, except for local modulation of reflexes via activation, habituation, or crossed inhibition.

Since the coexistence of more than a (comparatively) small number of stimulus/response patterns would create inextricable problems of coordination and integration, the overall architecture of a reflex-based organism will necessarily be simple. The subjectively relevant environmental dynamics, and therefore the organism's interactions, will be correspondingly simple.

The subjective space of a reflex-based organism will also be correspondingly limited, consisting of the small set of stimuli that it is sensitive to. These may include taxis and other simple forms of trail following, and the avoidance of aversive stimuli. In practice, therefore, space has no proper *structure*, in the literal sense of the term, to these organisms.

4.2 Affordance-based architectures

The organisms that belong to this second class of coupled architectures have internal states that play a role in their interactions with the environment; we borrow the term *affordance* from Gibson (1977) to refer to them.

Although the internal dynamics of an affordance-based architecture are still entirely coupled to the environmental ones, the picture becomes far more complex than was with reflex-based organisms. The coupling here is flexible, in that the internal states contribute in determining what environmental dynamics are currently the most relevant, among the several available at each moment. Thus, an individual who is looking for prey and one who is looking for mate will react to different affordances; and both will have to be able to adjust their internal dynamics if required by the external ones (e.g., if a predator is detected). These architectures may thus be described as dynamically ascribing a comparative weight to each affordance available, according to the current internal state, and then reacting to the balance of weights that has thus been created. Of course, the criterion with which these weights are allocated is part of the architecture itself.

The subjective space of these organisms is composed of all the affordances that are available at each moment. It has therefore a proper structure, although a non-objectual one, because the affordances are spatially oriented with respect to the organism's egocentric positioning and because they vary in attractiveness or aversiveness.

The interactions of these types of organisms with their subjective environment may therefore be conceived of as a complex and continuously changing balance between the affordances available. The complexity of this balance may vary greatly from species to species. Some species can stabilize specific types of affordance, so to let them govern the interaction over a certain interval of time; this interval may be longer or shorter, thus making the stabilization more or less permanent. The honeybee, for example, is capable of permanently fixing the flight trajectory that leads from the hive to an interesting source of food; the desert ant impermanently keeps track of the direction that will lead it back to the nest at the end of the current cycle of exploration; and the housefly is incapable of stabilizing its affordances at all.

What is interesting, in describing the interactions of all these different species in terms of a dynamic balance between affordances, is that there is no need to ascribe special representational or quasi-representational capabilities to honeybees or to desert ants with respect to houseflies; it

suffices for an explanation of the differences between these species that the former be capable of assigning a permanent or impermanent relative weight to certain environmental affordances. Since the nervous systems of all these insects are roughly similar, any other solution would be implausible from a neurobiological point of view.

5. DECOUPLED ARCHITECTURES

The architectures that belong to this second main class are those whose internal dynamics are decoupled from the external ones; that is, those that entertain representations. These should correspond roughly to the species whose nervous system includes proencephalic structures.

The concept of representation lies at the very heart of cognitive science, but it comes in different acceptances; it is therefore necessary to explain briefly what we mean by it.

We reject the computationalist framework, according to which mental representations are pieces of information internally stored in some predefined formal code (Newell & Simon 1976). This position is nowadays philosophically and psychologically unacceptable (Bruner 1990; Edelman 1992; Harnad 1990; Nagel 1986; Putnam 1988; Searle 1980, 1992), if only because of its rather controvertible consequence that whatever physical object undergoes internal changes due to world events could then be considered representational — including autonomous robots à la Brooks (Vera & Simon 1993) as well as thermostats and computers. On the other hand, the perceived failure of symbolic accounts of cognition has lately led many researchers to completely reject the very idea of representation (e.g., Brooks 1991) while, at the same time, keeping a view of cognition based on computational functionalism. We view these positions as the two horns of one dilemma (Tirassa 1999a), and want to avoid both.

To entertain a representation is instead to be in a certain semantic, or *intentional* (Searle 1983), relationship with the world (including one's own body and, for a few species, even one's own representations). This is the (largely unexplained) material property of certain types of physical objects, namely, certain highly sophisticated types of nervous systems. Like all the material properties of a physical object, representations will have a causal role in some of the interactions that that object has with the world: they are thus at the same time (part of the) causes and effects of the organism's interactions with the world (Tirassa 1999b). It is therefore more correct to talk of mind/body rather than of mind (or body) alone. (For further discussion of the problem of representations in cognitive ethology, see Allen & Bekoff 1997; Griffin 1978; Prato Previde, Colombetti, Poli & Cenami Spada 1992).

Representational architectures are, by definition, decoupled from the external world. *Decoupled* here does neither mean that representational organisms have internal dynamics that are abstract entities independent on the world, as is typical of classical artificial intelligence, nor that they live in a world of Platonic symbols or that environmental contingencies and opportunities are not crucial to them, a position that would simply be solipsistic. The point is instead that representational architectures have internal models of their subjective environment. These models are based on concepts, that is, on subjective ontologies resulting from largely innate categories superimposed on the world.

As regards the structure of space, the subjective environments of representational architectures will build upon the existence of *objects*. In this section we will discuss the three possible main subclasses of these architectures.

5.1 Deictic architectures

Deictic architectures are the simplest representational architectures. A deictic architecture (the term is borrowed from Agre & Chapman 1990) only represents what it can currently perceive. Objects are thus impermanent to it: they only exist insofar as they can be perceived. This means that it can have no token/type distinction: each object represented is at the same time a prototype of the whole class which it belongs to and a specific instance thereof. An object to these architectures is therefore represented as the whole set of interactions that it is (subjectively) possible to entertain with it.

This may resemble the case of affordance-based architectures. The difference is that an affordance (in our acceptation of the term) is not an object proper, but simply the potential for a certain action. The action that an affordance calls for will possibly be executed according to its relative weight as compared with the other affordances available. To entertain a deictic representation means instead to view an object as the experience of the possible interactions that concern that particular world entity. Thus, it is not that the concept of affordance applies to nonrepresentational architectures and not to representational ones; it is only that it means something very different in the two cases (which is why we have restricted our use of the term to the former case).

The idea of a deictic representation may also seem to apply to any representational architecture, at least under a certain acceptance of the concept (e.g., Glenberg 1997; Millikan 1998). From this point of view, the main difference between a deictic architecture and a more sophisticated one is that the former has no object permanence, that is, that it is incapable of singling an object out and possibly labeling it as an individual entity, and thereby of realizing that it exists even when it is out of immediate perception. This makes a great difference in the subjective structure of the environment.

As regards the subjective structure of space, a deictic organism will only interact with the space it can currently perceive; but, differently from what happens in a coupled architecture, it will represent that space as a region wherein proper objects exist (which is also why we have started to use the term *perception* only here). It will therefore be able to plan a trajectory in this region, deciding in advance what path to follow (according to criteria such as distance or dangerousness), what obstacles to avoid and how, and so on. This capability, although confined to the space that is currently perceived, allows nonetheless highly sophisticated interactions, at least as compared to those that are possible to lower-level architectures.

In principle, two particularly simple forms of learning are possible in a deictic architecture, one consisting in the acquisition of a novel way to cope with a deictic object, and the other in the addition of a new deictic object to the architecture's subjective ontology, possibly as a specialization of a previously existing one. Thus, if an animal is capable of creating a new deictic object (say, *my mate* in a monogamous species), it may be able to interact with it in ways that would be specific to that particular individual, while at the same time never being able to realize that it is an individual object (because there is no such thing as an individual object in the animal's subjective ontology).

The capability of forming a deictic object which happens to be "objectively" composed of only one instance also allows for the creation of a nest. Although, in order to go back to it, the organism has either to perceive it or to resort to simple modalities of orientation like trail following, it is however a (deictic) object or place, and therefore something very different from the nest of an affordance-based organism.

5.2 Base-level representational architectures

The logically successive step is an architecture capable of object permanence. This corresponds to the possess of concepts in the proper sense of the term, that is, as types of entities to which different tokens, or individuals, may belong.

This makes true learning possible, whereby a novel individual may be added to an existing class or a new class may be created, possibly as a specialization of one that already exists. The new class, differently from what happens in deictic architectures, would be a true class in its own respect, that is, it would comprise a proper token/type distinction, a set of individual entities that belong to it to a possibly variable degree, and so on. In any type of architecture, of course, learning can only take place within the space of possibilities generated by the innate endowment of the architecture.

As regards the subjective structure of space, base-level representational organisms are the first to possess a map of the territory they inhabit. A region to them may be divided into subregions, or zones; each region or zone comprises individual paths and individual objects that may be used as landmarks. A landmark (or, in general, an object) to these organisms needs not be a physical piece of matter in the material sense of the term; it may also be a landscape, a skyline, a socially shared mark like an odor, and so on.

To a base-level representational architecture, the nest is therefore an individual place, reachable via a network of individual paths that are characterized in their turn by landmarks. Other individual places may include sources of food or water, dangerous zones, and so on.

5.3 Metarepresentational architectures

The final step is metacognition, that is, an organism's capability of representing its own representations. As far as we know, the human species is likely to be the only one on this planet to have such capabilities, that could be shared, in the best case, by a few other primates (see Premack, Premack & Sperber eds. 1995).

Metacognition dramatically changes a species' interactions with its subjective environment. It makes it possible to attach abstract labels (that is, symbols) to existing entities, to imagine non-existing entities and to treat them as if they were real, to use symbols for referential purposes or as place-markers, and so on.[2] This makes it possible in turn to formulate theories about the world, to reason formally, to reuse in a certain type of interaction the features of the world that were relevant to a different type of interaction, and to communicate in a mentalist way with conspecifics.

As regards the subjective structure of space, a metarepresentational species is capable of creating abstract regions or zones with abstract borders and landmarks and, most important, of entertaining survey maps. We do not conceive of survey maps as allocentric: since a representation is, by definition, someone's subjective point of view, they can only be egocentric. Survey maps result instead from the pretense to be dislocated in a different position (say, one kilometer above the city) and watching the world from that perspective. This allows to draw spatial inferences and therefore to plan in advance a path in a known as well as in a partially unknown region. It must be remembered, however, that these plans are always, by necessity, partial: they are not recipes for action to be followed blindly, but guides for action to be further specified in the interaction with the real world (Tirassa 1997).

Metacognitive organisms have two further capabilities. First, given the appropriate cognitive tools, they can externalize their representations by means of drawings or language. Second, given the ability to understand the representations entertained by a conspecific and to affect them in a desired way, they may communicate spatial information to one another with the aid of these externalized tools.

6. FUTURE DEVELOPMENTS

The classification we have proposed here has a very high level of abstraction and currently relies fundamentally on analytical considerations. The next step of our research will be to derive, for each class of architectures described, a description of the types of interaction that it may generate with regard to the control of space and locomotion. Subsequently, we will check the empirical validity of our analysis by looking for confirming or disconfirming evidence from neurobiology, ethology and autonomous robotics. This is also likely to lead to a refinement of the classification.

The ultimate goal of this research is to build a taxonomy of active organisms based on three columns, namely, *cognitive architecture*, *neurobiology and embodiment*, and *interaction*, that be analytically and empirically consistent on each level.

---

[2] Let us remark that *to be capable of using symbols* is not the same thing as *being a symbol system* as postulated in classical cognitivism (see also the discussion at the beginning of this section).

REFERENCES

Agre, P.E., Chapman, D. (1990) What are plans for? *Robotics and Autonomous Systems* 6: 17-34.

Allen, C., Bekoff, M. (1997) *Species of mind*. Cambridge, MA: MIT Press.

Barkow, J.H., Cosmides, L., Tooby, J., eds. (1992) *The adapted mind*. New York and Oxford: Oxford University Press.

Brooks, R.A. (1991) Intelligence without representation. *Artificial Intelligence* 47: 139-159.

Bruner, J.S. (1990) *Acts of meaning*. Cambridge: Cambridge University Press.

Cosmides, L., Tooby, J. (1994) Beyond intuition and instinct blindness: toward an evolutionarily rigorous cognitive science. *Cognition* 50: 41-77.

Edelman, G.M. (1992) *Bright air, brilliant fire: on the matter of the mind*. New York: Basic Books.

Geminiani, G., Carassa, A., Bara, B.G. (1996) Causality by contact. In: *Mental models in cognitive sciences. Essays in honor of Phil Johnson-Laird*, eds. J. Oakhill & A. Garnham. Hove, UK: Erlbaum/Taylor & Francis.

Gibson, J.J. (1977) The theory of affordances. In: *Perceiving, acting, and knowing*, eds. R.E. Shaw & J. Bransford. Hillsdale, NJ: Erlbaum.

Glenberg, A.M. (1997) What memory is for. *Behavioral and Brain Sciences* 20: 1-55.

Griffin, D.R. (1978) Prospects for a cognitive ethology. *Behavioral and Brain Sciences* 4: 527-538.

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

Lorenz, K. (1965) *Evolution and modification of behavior*. Chicago and London: The University of Chicago Press.

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

Millikan, R.G. (1998) A common structure for concepts of individuals, stuffs, and real kinds: more Mama, more milk, and more mouse. *Behavioral and Brain Sciences* 21: 55-100.

Nagel, T. (1986) *The view from nowhere*. Oxford: Oxford University Press.

Newell, A., Simon, H.A. (1976) Computer science as empirical enquiry: symbols and search. *Communications of the Association for Computing Machinery* 19: 113-126. Reprinted in: *Mind design*, ed. J. Haugeland. Cambridge, MA: MIT Press, 1981.

Papi, F. (1990) Homing phenomena: mechanisms and classification. *Ethology Ecology & Evolution* 2: 3-10.

Prato Previde, E., Colombetti, M., Poli, M., Cenami Spada, E. (1992) The mind of organisms: some issues about animal cognition. *International Journal of Comparative Psychology* 6: 79-119.

Premack, A.J., Premack, D., Sperber, D., eds. (1995) *Causal cognition: a multidisciplinary debate*. Oxford: Clarendon Press.

Putnam, H. (1988) *Representation and reality*. Cambridge, MA: MIT Press.

Searle, J.R. (1980) Minds, brains, and programs. *Behavioral and Brain Sciences* 3: 417-424.

Searle, J.R. (1983) *Intentionality: an essay in the philosophy of mind*. Cambridge: Cambridge University Press.

Searle, J.R. (1992) *The rediscovery of the mind*. Cambridge, MA: MIT Press.

Tirassa, M. (1997) Mental states in communication. *Proceedings of the 2nd European Conference on Cognitive Science*, Manchester, UK.

Tirassa, M. (1999a) Taking the trivial doctrine seriously: functionalism, eliminativism, and materialism. *Behavioral and Brain Sciences* 22: 851-852.

Tirassa, M. (1999b) Communicative competence and the architecture of the mind/brain. *Brain and Language* 68: 419-441.

Varela, F.J., Thompson, E., Rosch, E. (1991) *The embodied mind. Cognitive science and human experience*. Cambridge, MA: MIT Press.

Vera, A.H., Simon, H.A. (1993) Situated action: a symbolic interpretation. *Cognitive Science* 17: 7-133.