

Coordinating with the Future: the Anticipatory Nature of Representation*

Giovanni Pezzulo

Institute of Cognitive Sciences and Technologies - CNR

Via S. Martino della Battaglia, 44 - 00185 Roma, Italy

Tel: +39 6 44595206; Fax: +39 06 44595 243

giovanni.pezzulo@istc.cnr.it

Abstract

Humans and other animals are able not only to coordinate their actions with their current sensorimotor state, but also to imagine, plan and act in view of the future, and to realize distal goals. In this paper we discuss whether or not their future-oriented conducts imply (future-oriented) representations. We illustrate the role played by anticipatory mechanisms in natural and artificial agents, and we propose a notion of representation that is grounded in the agent's predictive capabilities. Therefore, we argue that the ability that characterizes and defines a true cognitive mind, as opposed to a merely adaptive system, is that of building representations of the non-existent, of what is not currently (yet) true or perceivable, of what is desired. A real mental activity begins when the organism is able to endogenously (i.e. not as the consequence of current perceptual stimuli) produce an internal representation of the world in order to select and guide its conduct goal-directed: *the mind serves to coordinate with the future*.

Keywords: Anticipation, Expectation, Internal Model, Prediction, Representation, Simulation, Goal

1 Introduction

Since the beginning of AI, intelligence has been conceived as the capacity to solve a problem by working on internal representations of problems, i.e. by acting upon internal, or mental, or small-scale models before acting in the world. The presence of internal, mental models is thus a leitmotiv in cognitive science (Johnson-Laird, 1983); and their role has been claimed to be mainly related to the future. According to Craik (1943, pag. 61), they allow generating imaginary experiences, mental simulation of external reality:

If the organism carries a small-scale model of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present and future, and in every way to react in a much fuller, safer, and more competent manner to the emergencies which face it.

In a similar way, Tolman (1932) discussed about *vicarious trial and error*, i.e. learning as if experience had really happened, as an essential trait of purposive behavior. Another very influential model in cybernetics and early cognitive science, the TOTE (test, operate, test, exit) model (Miller et al., 1960), implies an anticipated representation of the goal state to be matched against current perception in order to start (if there is mismatch) and stop (if there is match) behavior. From those initial speculations therefore, a countless number of AI systems for planning and reasoning have been based heavily on the concept of manipulating anticipatory, action-effect internal representations (often encoded by the designers) in order to achieve a

*This is a pre-print of a copyrighted paper that will appear in *Minds and Machines*, www.springerlink.com/content/100314/

goal (which is, again, often explicitly encoded). At the same time, the emphasis on reactive systems of the ‘new AI’ (Brooks, 1991) has lead many researchers to minimize the role of anticipatory behavior and future-oriented representations. As we will discuss, now that many of the arguments against traditional AI systems have been widely accepted, we can come back to the analysis of anticipatory behavior as one crucial aspect of natural cognition, and try to model how and why it emerges from reactive, sensorimotor engagement. In this new perspective, which we develop here, we think that anticipation can become really a key concept in the development of artificial cognitive systems and it can shed light on the nature of representation.

1.1 Anticipatory Systems

Internal models, anticipatory representations and goal-directed behavior have always been related in the analysis of cognitive systems. Cognitive agents are conceived as *anticipatory systems*; they are able to reason about the future and fulfill their goals thanks to their anticipatory representations. A popular definition of anticipatory system is provided by Rosen (1985):

A system containing a predictive model of itself and/or its environment, which allows it to change state at an instant in accord with the model’s predictions pertaining to a latter instant.

The point here is that in an anticipatory system the sensorimotor state is insufficient for determining behavior, but typically an *anticipation* of the future is also needed. Many cognitive functions in the individual and social domains have been claimed to be essentially anticipatory, such as visual attention (Mehta and Schaal, 2002), motor control (Wolpert et al., 1995; Wolpert and Kawato, 1998), imagery (Jeannerod, 1994, 2001; Kosslyn and Sussman, 1994), and mind reading (Frith and Frith, 2006). In the last years a new paradigm, which we could call an *action-based* or *motor-based theory* of cognition, is increasingly gaining relevance. Cognition is *for doing*, *not for thinking* and accordingly the basic concepts of cognitive science such as representations are being redefined. As suggested by the discovery of mirror neurons (Rizzolatti et al., 1996) among others, representations are mainly action-oriented and deeply based on the motor apparatus. *Simulative* theories of cognition (Cotterill, 1998; Hesslow, 2002) claim that representing is engaging in simulated interaction with the environment by means of *internal predictive models* which permit to predict the consequences of own actions as well as the dynamics of the environment.

1.2 Anticipatory Behavior and Representations

In philosophy the *productive* perspective of cognition originates from Kant’s idea (Kant, 1998) that, although our representations are rooted on empirical ground, we produce them by means of our cognitive apparatus instead of passively processing the structure of the world. Sloman (2006) takes a similar perspective and describes how a cognitive system can not only deal with affordances (Gibson, 1979), but also with *vicarious affordances*: affordances for oneself in the past (such as missed opportunities) or future (planning to exploit to-be-created opportunities). Elaborating Piaget’s constructivism (Piaget, 1954) and the central notion of *schemas*, Bickhard (2001) proposes an interactionist hypothesis about the emergence of representations: against homuncular theories in which there is a ‘spectator’ of representations, they are conceived as *ways for setting up indications of further interactive potentialities*; and those indications constitute representational content. Thus representation is of future potentialities of interaction. Dennett’s *multiple drafts model* (Dennett, 1991) also adresses the idea of drawing and manipulating passive representations and describes sensing as producing multiple concurrent interpretations on the basis of current goals. Barsalou (1999) and Grush (2004) provide unitary accounts of all these phenomena proposing theories of cognition relying on *simulation* and *emulation*.

These ideas are also exploited in a computational perspective. For example, in cognitive science and cognitive robotics (Arbib, 1992; Drescher, 1991; Neisser, 1976; Pezzulo and Calvi, 2006a, 2007; Roy, 2005; Wolpert and Kawato, 1998) mechanisms such as *schemas* or *coupled forward-inverse models* are used, which (1) do not passively process all stimuli but produce perceptual anticipations, (2) orient attention only toward stimuli which are relevant for action, and (3) do not respond to environmental stimuli but to expectations

produced by internal models. Schemas can also be arranged hierarchically (e.g., Demiris and Khadhour, 2005; Haruno et al., 2003; Pezzulo et al., 2005); in that case top-down influences canalize the cognitive processing in accordance with expectations generated at the higher level, although the system remains also responsive to stimuli-driven, bottom-up dynamics.

Putting attention to the anticipatory aspects of human and animal conduct is necessary in order to explain the variety of mental phenomena that can produce behavior which is not immediately driven by what is here-and-now in the environment, but *is about* and *for* the future. Anticipatory mechanisms are required not only for advanced cognitive capabilities such as imagination and mind reading, but also for adaptive behavior of simple animal species. According to Cisek and Kalaska (2001), “Early animals could have gotten along with simple trial-and-error behavior and reinforcement learning, but more effective behavior requires the development of neural systems to predict the sensory consequences of potential actions and to use these predictions to select the action that leads to the most desirable results.”

There are in fact many examples of apparently simple animal conducts which have an anticipatory component. For example, Schomaker (2004) describes how the prey-catching behavior of the *jumping spider* implies an estimate of future states of the world, and Colombo and Graziano (1994) report of pigeons putting attention to expected stimuli. Increasingly sophisticated behaviors are possible by means of anticipation: for example rats can preventively compare alternative strategies in navigating T-mazes (Hesslow, 2002); and monkeys are able to avoid the more immediate affordances of the environment and to invent novel ways to use a stick for resolving complex problems (Dunker, 1935). These studies (among others) demonstrate that in order to select and control their actions appropriately several animals are able to anticipate their (sensory) effects and to evaluate them as desirable or not.

The idea that the stimulus-response paradigm is unable to explain most animal behavior is certainly not novel Dewey (1896). For example, James (1890) and more recently Hommel et al. (2001) suggested that the brain learns and exploits an *ideomotor code* (ACTION \rightarrow EFFECT and not STIMULUS \rightarrow RESPONSE for action selection, planning, and control. Perhaps the most remarkable consequence of this functional organization of action is that it allows realizing an ‘inversion’, from ACTION \rightarrow EFFECT to EFFECT \rightarrow ACTION, and then to select and trigger actions depending on their (desired) effects. This permits to act goal-directed instead than just reactively¹.

From anticipation to goals and goal-directedness Not all anticipatory processes serve for deriving knowledge (e.g., about the environment, or the consequence of one’s own actions). Anticipation serves also to autonomously generate increasingly distal *goals*: anticipatory representations that select and guide action. In order to coordinate purposively several of its actions in time, a cognitive agent has to be able to conceive and represent (implicitly or explicitly) its future, including its objectives, and to maintain its predictions and representations accurate.

Goals are of primary importance for our analysis since they relate anticipatory mechanisms, representations, and future-oriented conducts. One idea that we put forward in this paper (see also Pezzulo and Castelfranchi, 2007) is that representation originates from anticipatory mechanisms for the sake of action control; this capability is then exapted, and goal (state) representations are produced that permit detachment from the current sensorimotor flow and ultimately the purposive realization of distal objectives. Thanks to anticipation, cognitive agents can break the boundaries of their sensorimotor interactions by formulating novel goals to achieve outside those immediately dictated by their environment; briefly, they can *coordinate with the future* and not only with the present.

This capability depends on the double nature of goals, epistemic and motivational. Goals are not only a representation of the future: they have also a motivational, axiological, or deontic nature: saying us not only how the world is, was, will be, but how the world should be, *how the organism would like the world to be*. These two aspects of goals allow realizing the key capability of cognitive minds: conceiving, desiring and achieving what is not there. At the same time, this comes with risks: the degree of detachment from the

¹ Additional functions such as action planning, monitoring, and control, are also possible thanks to this functional organization of action. See also (Rosenblueth et al., 1943) for related ideas in behavior control in early cybernetics.

present state of affairs required for goal-directed action is also a presupposition for hallucinations, delirium, desires, and utopias.

Since anticipation is a key mechanisms for enabling goal-directedness and intentionality, we propose to consider it as an hallmark of cognition. The ability that characterizes and defines a true cognitive mind, as opposed to a merely adaptive system, is that of building representations of the non-existent, of what is not currently (yet) true or perceivable, and particularly goal representations. A real mental activity begins when the organism is able to endogenously (i.e. not as the consequence of current perceptual stimuli) produce internal representations of the world, like in perceptual simulation (Castelfranchi, 2005; Grush, 2004; Hesslow, 2002), and to use them to guide its conduct goal-directed, ‘from the future’.

Structure of the Paper In the rest of the paper we will provide an unitary account of anticipation as the key mechanism that allows representing and dealing with the future. Our theoretical analysis will be complemented by several examples of anticipation-based cognitive functions taken from empirical studies as well as computational models.

In the first part of the paper we discuss the crucial role of anticipation for several cognitive functions: we review empirical evidence of anticipatory behavior and representations in natural cognition and we describe how all that is modeled in artificial systems: cognitive robotics and machine learning.

In the second part of the paper, by noting that the investigated models use the concept of representation, we propose a critical review of this concept, introducing its role in cognitive science and describing the attacks it received. We thus explore a different view about anticipation, which is proposed in dynamical systems theory, that avoids representations. The point is understanding whether or not anticipatory representations are necessary for anticipatory behavior. We distinguish between implicit (behavioral) and explicit (representational) anticipation and we conclude that up to the moment representations are the best explanation we have.

2 Empirical Evidence about Anticipation

Evidence for a role of anticipatory, top-down processes is reported in the literature of visual attention (Mehta and Schaal, 2002) and motor control (Wolpert et al., 1995; Wolpert and Kawato, 1998). For example in the visual cortex a hierarchy of increasingly sophisticated representations has been claimed for visual processing (Riesenhuber and Poggio, 1999); this architecture can implement *predictive coding* (Rao and Ballard, 1999) in which anticipatory signals are spread top-down and error signal are spread bottom-up.

Similar control structures may exist for the preparation and execution of actions. According to Jeannerod (1994); Jeannerod and Decety (1995); Jeannerod (2001) there are striking neural resemblances between motor imagery and motor preparation. The frontal lobe, which typically is involved in motor preparation, is active even when the output of the primary motor cortex is suppressed; in this case it could act ‘in simulation’ of the action without actually activating the overt part of the movement. For these reasons perception and action can not be separated, since movement is part of the perceptual process (Jeannerod, 1979). Similarly, Edelman (1987) claims that for an exploring animal the motor behavior is essential to perception and that *our sensorimotor maps provide the basis for conceptualization and reasoning*.

Evidence of anticipatory aspects in motor activity and high level cognition is also reported in several psychological experiments (Hoffmann et al., 2004; Hommel et al., 2001; Kunde et al., 2004; Prinz, 2005). Kunde et al. (2004) report that during simple discrimination tasks humans are influenced by certain unrelated sensor stimuli, visual and auditory; for example, if those stimuli violate certain symmetries, humans respond slower. This very robust effect occurs both when those interfering stimuli are actually presented before the discrimination task or when they are expected to be presented after the completion of the task, highlighting the presence of active anticipatory representations (of the interfering stimuli) during the preparation of action. Consistent evidence about the role of explicit expectations comes from neuroimaging studies conducted with a variety of techniques, including fMRI (Carlsson et al., 2000). Lu and Ashe (2005) report that cells in the primary motor cortex of the monkey show anticipatory activity related to a specific memorized sequence of upcoming movements. Evidence for *prospective processing* and *prospective coding* has been reported in

animal perceptual systems (Colombo and Graziano, 1994; Rainer et al., 1999; Roitblat, 1980). In choice tasks many animal species including monkeys and pigeons are able to maintain a memory not only of past stimuli but also of expected stimuli, and this capability makes them able to direct attention toward expected critical events. Recently several researchers have investigated the biological mechanisms of anticipatory and goal-directed capabilities, and how they are learned (e.g., Balleine and Dickinson, 1998; Schultz et al., 1997).

Which is the neural substrate underlying prediction and selection for action? As discussed in (Fleischer, 2007), several brain structures are implied. For example, it has been supposed that the *cerebellum*, which is connected both with the cerebral cortex and with premotor and motor areas, is able to produce sensory predictions (Ghez and Thach, 2000; Kawato, 1990; Kosslyn and Sussman, 1994). On the same time the *basal ganglia*, that receive input from all parts of the cerebral cortex and have strong outputs to the premotor and prefrontal cortex, are supposed to be involved in action selection and movement initiation (Houk and Wise, 1995; Wickens, 1997) (computational models exist, too; e.g., Gurney et al., 2001). The cerebellum and the basal ganglia are thus supposed to create a loop (Middleton and Strick, 2000) permitting to simulate and select multiple alternative courses of actions, providing full neural support for anticipatory capabilities.

This evidence supports also the *ideomotor principle* (James, 1890), according to which an anticipatory representation of the goal state precedes, selects and regulates the act (Hoffmann, 1993; Hommel, 2003; James, 1890): *action planning takes place in terms of anticipated features of the intended goal*. A crucial role of anticipation in action selection has been claimed by Greenwald (1970), too: “a current response is selected on the basis of its own anticipated sensory feedback”. Prinz (1990) and Hommel et al. (2001) propose that a common coding exist in perception and action, suggesting that the motor system plays an important role in perception, cognition and the representation of goals. They suggest that action-effect relations can be learned and used to reverse the typical linear stage theory of human performance: instead of reacting from stimulus to response, as suggested by the sensorimotor view, it is possible to activate actions by selecting their effects. Neurobiological evidence for common mechanisms in perception and action are also reported in (Knoblich and Prinz, 2005).

Internal Models It has been proposed (Kawato, 1999) that *internal models*, which mimic the behavior of external processes, are used by the brain for motor control of action. *Inverse internal models* take actual stimuli and the goal state as inputs and provide the motor commands necessary to reach the desired state as output. *Forward internal models* take (an efferent copy of) the executed motor command as input and predicts the next sensori stimuli. Forward, predictive models have been claimed to be used for many purposes in motor control: compensating for delays in sensory feedback, cancel the self-produced part of the input from sensori stimuli, etc. (Wolpert et al., 1995; Wolpert, 1997; Wolpert and Kawato, 1998) and empirical evidence exists for their involvement in visuomotor tasks (Mehta and Schaal, 2002). However, their relevance extends to many if not all the spheres of cognition. Predictive models, in fact, have been widely claimed to be involved in the perception of action as well as in understanding of behavior (produced by self or by others) and prediction of dynamics in the environment (Blakemore and Decety, 2001; Jeannerod, 2001; Gallagher and Jeannerod, 2002; Gallese, 2001; Rizzolatti et al., 2001; Iacoboni, 2003; Schubotz, 2007; Wolpert et al., 1995; Wolpert and Flanagan, 2001; Wolpert et al., 2003).

Mirror Neurons The most striking evidence for the action-oriented nature of representations is the discovery of *mirror neurons* in the premotor area F5 of macaque monkeys (Rizzolatti et al., 1996), and successively in many other brain areas of monkeys and humans, indicating the existence of a *mirror system* (Rizzolatti and Craighero, 2004). In this respect, Gallese and Metzinger (2003) suggests that *the goal is represented as a goal-state, namely, as a successfully terminated action pattern*. Recent findings (Fogassi et al., 2005) also indicate that motor IPL neurons coding a specific act (e.g., grasping) show markedly different activations when this (observed) act is part of different courses of actions, leading to different distal goals (e.g., for eating or for placing). Since activation begins before the course of action starts, it has been proposed that those neurons do not only code the observed motor act but also indicate an anticipation (in an ideomotor coding) of the whole distal goal, i.e. understanding of the agents intentions.

Mirror neurons and the simulative processes that they enable have also been claimed to be involved in

a variety of social cognitive capabilities, such as understanding other’s intentions and imitation (Goldman, 2005; Hurley and Chater, 2004; Iacoboni, 2002; Meltzoff and Prinz, 2002), distinguish self from others (Decety and Chaminade, 2003), self identification and agency (Jeannerod and Pacherie, 2004; Ruby and Decety, 2001), mind reading (Frith and Frith, 2006; Gallese and Goldman, 1998), language production and understanding (Rizzolatti and Arbib, 1998). Pacherie and Dokic (2006) proposes a *motor understanding* which is different from a pure *visual understanding* and involves deeply the motor areas. The striking similarity of activated areas between performing, observing, simulating or imitating actions also suggests a shared neural space for private and social activity. Gallese (2001) thus proposes with the *shared manifold hypothesis* that a shared interpersonal space is preliminary to the sense of self, thus that ‘we’ precedes ‘I’. The shared space is acted and understood by means of the motor apparatus which can act really or be re-enacted in simulation. In a similar vein, Hurley (2005) proposes the *shared circuits hypothesis*, also suggesting a possible evolutionary pathway, and Sloman (2006) argues that the same mechanisms could be involved both in understanding own affordances and affordances for others. Several authors have also proposed unified accounts of individual and social action based on predictive capabilities (Frith, 2007; Jeannerod, 2001; Wolpert et al., 2003).

Simulative theories Especially after the discovery of mirror neurons (Rizzolatti et al., 1996), ‘simulative’ theories gained momentum in psychology and neurobiology: representations are conceived to be action-related and are thus grounded on the representations subserving the motor system. According to Hesslow’s *simulation hypothesis* (Hesslow, 2002) cognitive agents are able to engage in simulated interaction with the environment which serves to (prepare to) interact with it. According to Gallese (2000): “To observe objects is therefore equivalent to automatically evoking the most suitable motor program required to interact with them. Looking at objects means to unconsciously ‘simulate’ a potential action. In other words, the object-representation is transiently integrated with the action-simulation (the ongoing simulation of the potential action)”.

As we will see, computational models have been proposed of a ‘simulation and selection’ mechanism; the idea (Cotterill, 2001; Damasio, 1994; Hesslow, 2002) is executing and evaluating potential interaction by enacting the covert part of sensorimotor programs. Internal models such as forward models allow producing multiple candidate courses of actions, and the distal results of actions can be thus *evaluated* in advance, for example with a mechanism like Damasio’s *somatic markers* (Damasio, 1994). Moreover, there can be a *response preparation* to distal effects: agents can *react to the future* and not only to the present.

Barsalou (1999) and Grush (2004) report significant evidence for the activity of internal emulation of external reality; for example, hearing singing or speech involves the partial re-enactment of mechanisms for producing the heard sounds; and the same happens for the other senses. Evidence exists for similar timings in simulated reading and reading aloud (Landauer, 1962), and in actual and simulated rotation of objects (Kosslyn and Sussman, 1994). Several findings support the claim that imagined actions retain the same temporal characteristics as the corresponding real action when it comes to execution, and neuroimaging techniques have also revealed that they share a common neural substratum, suggesting that they are functionally the same (Decety et al., 1989; Decety, 1996; Decety et al., 1997).

These studies entail that at least some of the involved representations retain sensorimotor features and that a process of simulation retaining the timing of actual behavior is in play. However, simulative capabilities seem to go beyond behavioral capabilities, since we are able to understand and appreciate even music that we can not produce, and in general to recognize novelties, to conceive truly novel plans for fulfilling unexperienced tasks². This seems to imply that the process of *detachment* from sensorimotor cycle, which we will investigate in the details, do not extend only in time but also produces increasingly abstract levels of description, which are in turn functional to novel behavioral capabilities and competences. Whether all these levels retain part or all of the sensorimotor features, or if they include novel ontologies and the discovery of unperceived constraints such as causality as argued in (Sloman, 2006), is of course an empirical issue (for a discussion about how the process of abstraction is compatible with simulative theories, see (Barsalou, 2003)).

It is moreover questionable whether the whole manipulation machinery we humans use in order to act

² See also Schubotz (2007) for a discussion of how internal models developed for predicting and controlling one’s own motor system can be used for predicting and simulating external events

on our internal representations is shared with other animals. Many cognitive operations such as conceiving truly distal goals, counterfactuals, abstract representations or the non existent could be limited to our specie and there is thus the possibility that they depend crucially on *language* and other cultural artifacts (Tomasello et al., 1993) that can have open the possibility to act on our representations at a previously unmatched degree of abstraction. Or, there could be the possibility that things are the other way around: linguistic competence could have arisen thanks to externalization of pre-existing capabilities of running abstract competences (Sloman, 2006). However, all this goes beyond the scope of the present paper.

3 Anticipation in Artificial Systems

In cognitive robotics and machine learning several research programs are focused on how anticipatory representations allow realizing cognitive functions; some of them are explicitly inspired by (neuro)biology, others are not. Many principled approaches to the design of anticipatory systems have been proposed. Butz et al. (2003) reviews many anticipatory mechanisms and proposes a classification distinguishing four kinds of anticipation: *implicit*, *payoff*, *sensorial* and *state* anticipation. In implicit anticipation the agent do not produces expectations, but still exhibits anticipatory behavior which are for example selected by evolution to deal with the future; anticipation is functional but not represented. In payoff anticipation expectations about rewards for states of the environment are formed and used for planning. The two latter cases involve explicit prediction of the outcome of own actions. Sensorial anticipation involves predictions only influencing the sensory processing of the agent, not its behavior; examples are priming and visual imagery. State anticipation implies the construction of detailed predictive models of the agent itself or of the environment, directly influencing choices and behavior. We will come back to these distinctions, and in particular between implicit and explicit anticipation in sec. 6.3

Since prediction is a fundamental task in many scientific disciplines, several mechanisms for predicting e.g. time series and genetic series, or for weather forecasts have been developed, and these methodologies have been also used in artificial cognitive systems. In robotics, for example, a well known instrument developed in control theory has been often used, the *Kalman filter*. It realizes not only predictions but also many other related functionalities such as filtering and estimation, and for this reason is considered a prototypical emulator by Grush (2004). Various types of recurrent neural networks are widely used such as those proposed by Jordan and Rumelhart (1992) or the LSTM (Gers et al., 2003), which have been shown to predict at different time scales and address perceptual aliasing problems. Several comprehensive theoretical and computational frameworks have been proposed for studying and realizing anticipatory systems, too. The bayesian framework for prediction has been used in many studies (Friston, 2003; Kording and Wolpert, 2006). Witkowski (2003) investigates the anticipatory and predictive elements inherent to the most known forms of learning. Tsoukalas (1998) introduces a neurofuzzy methodology for incorporating predictions about the future in formal approaches to control. Schmidhuber (2002) discusses possible strategies to be adopted by artificial systems to learn to focus on the predictable components of spatio-temporal events. The capacity to predict and to discriminate between predictable and unpredictable elements in the environment is also fundamental for metalearning mechanisms such as *curiosity* (Oudeyer et al., 2005; Schmidhuber, 1991).

3.1 Modeling the functions of anticipation in artificial systems

Anticipation has been widely used for realizing a number of cognitive functions; here we review how many of them, such as action selection, filtering, action control, planning, etc. and provide some example of how they have been modeled in artificial systems. It is worth noting that some of these functionalities are based on more primitive ones which can be themselves anticipatory; here we do not try to order them according to criteria of complexity. Those interested in the evolution of anticipation can refer to (Gardenfors and Osvath, 2005; Hurley, 2005; Pezzulo and Castelfranchi, 2007) in which an evolutionary pathway is sketched. The idea is that the initial acquisition of anticipatory mechanisms is probably due to converging evolutionary pressures, related to the control of action: filtering noise, replacing missing or unreliable stimuli, compensating time delays, etc. (Hurley, 2005; Pezzulo and Castelfranchi, 2007). Later on, anticipation provides new evolutionary

possibilities and enables several complex cognitive functions, and definitely permits to act goal-directed.

Control and monitoring of action Anticipation has many related roles in the control of action. According to Hoffmann (1993): “action-induced anticipations are needed for stabilizing perception”. This is the *reafference principle* of von Holst and Mittelstaedt (1950) which is recently proposed by Grush (2004), too. In theory of control, several mechanisms realizing those functionalities such as *Kalman Filters* are studied and Grush (2004) provides evidence for the fact that similar mechanisms could be implied in natural cognition. Anticipatory mechanisms can also be used for stabilizing feedback. Like in *Smith Predictors*, a forward model is used both for mimicking the plant and for canceling the predictable part of the feedback; in this way only the unpredictable part of the feedback is used for correcting errors within the feedback loop (Miall and Wolpert, 1996). Another function related to the control of action is erasing stimuli produced e.g. by the body of the agent (Desmurget and Grafton, 2000). For example, an agent which is able to track other agents can avoid tracking its own arms moving in its visual field by erasing the portion of incoming stimuli which is self-generated. Monitoring actions during their courses is another function facilitated by anticipation. By comparing actual with anticipated stimuli, an agent can regulate its behavior. Lastly, according to Adams (1971), anticipations can be used as a reference signal for the control of voluntary acts: “Beginning the movement brings an anticipatory arousal of the [perceptual] trace, and the feedback from the ongoing movement is compared with it”. In robotics there are several examples of anticipation for the control of action. For example, the robot Murphy (Mel, 1990) can exploit efference copies of motor commands for generating simulated perceptual inputs and is thus able to maneuver its arm robustly even in partial absence of sensory stimuli. Gross et al. (1999) use internal simulation of the sensory consequences of multiple possible motor actions to perform robust planning in presence of noise.

Learning Jordan and Rumelhart (1992) exploits expectations produced by forward models for ‘vicarious learning’, similarly to what proposed by Tolman (1932). Reinforcement learning can be used to learn a predictive model of the worlds dynamics and e.g. calculate expected reward. In his seminal paper Sutton and Barto (1981) introduced a framework for *temporal difference learning* which was successively developed (Sutton, 1988; Doya, 1996). Recently the role of anticipation in machine learning has been revitalized thanks to the concept of *predictive state representation* (Littman et al., 2001; Wolfe and Singh, 2006), which has been shown to be effective for example in learning (Wolfe et al., 2005) and planning (James et al., 2004). Butz (2002) proposes an *anticipatory classifier system* which autonomously learns action-effect rules and chains them for planning and action control, improving the seminal work of Drescher (1991).

Anticipation can also be used for learning ‘hidden structures’ in data series. Elman (1990) firstly showed that a connectionist network can learn the formal structure of language, its syntax and semantics, by learning to predict the next word, without any innate knowledge. A similar methodology has been applied to series of stimuli sensed by robots, which have a rich structure, too. As shown by Pierce and Kuipers (1997), by learning to anticipate the effects of its actions, a robot can learn a rich hierarchical model of the robot’s sensory and motor apparatus, such as position and type of its sensors and the degrees of freedom of its effectors. This knowledge can bootstrap the representation of own body schema.

Self-modeling and body schema Anticipation can be used for distinguishing self-produced motion from sensory stimuli which are caused by interaction with objects in the environment. As shown in the developmental literature, children at early stage attribute to themselves the cause of everything happens; or, better, of everything they can predict. They learn only successively to discriminate among those effects which are under their direct control and those depending on dynamics and causality in the environment. But for this activity a full fledged *body schema* is necessary: understanding the boundaries of own prediction and control permit to discriminate self from others and from the environment (Maravita et al., 2003; Piaget, 1954). Bongard et al. (2006a,b) describes how robots might evolve body schemas by means of interactions and their anticipation. Mohan and Morasso (2006) also investigated how carrying one stick leads to extensions of the body schema which can however be adapted to the new situation by means of adapting its predictions.

Selection of action adequate to the context A combination of forward and inverse models has been used for action control in robotics both in distributed approaches (Tani, 2003; Tani et al., 2004) and in localist ones (Demiris and Khadhour, 2005; Mohan and Morasso, 2006; Pezzulo and Calvi, 2006a; Tani and Nolfi, 1999; Wolpert and Kawato, 1998), the latter based on the mixture of experts architecture (Jacobs et al., 1991). By introducing many competing combined forward-inverse models, many competing motor plans can be generated and maintained for the same or for different targets, and choice depends on predictive accuracy. Models predicting better are selected; the rationale is that they are well attuned with the current context. As an example, Wolpert et al. (1995); Wolpert and Kawato (1998) use competing models as alternative motor plans entailing alternative hypotheses about the context (*full glass* and *empty glass*); prediction of the right sensorimotor flux is thus used for an evaluation of the hypotheses and the selection of the best modules. Since modules' outputs can be combined linearly to realize mixed behavior, these models provide generalization, too. Algorithms for learning and combining context have been proposed (Vijayakumar et al., 2005) which extend to non linear dynamics, too.

Planning and selecting via simulation The anticipatory representations provided by the forward models can not only be used on-line for the selection of action, but also for more complex functionalities such as offline planning: possible outcomes of events can be simulated and compared offline by exploiting the same machinery involved in online visual and motor planning, but without sending commands to the effectors. During this operation the expected stimuli replace actual ones and serve as inputs for chaining the schemas. This offers two more advantages: detaching the representations from the sensorimotor loop by setting up hierarchies of schemas representing abstract concepts; and using goal states, and not current stimuli, for the selection of action. The *simulation hypothesis* (Hesslow, 2002) has been recently tested in simulated and robotic settings (Stephan and Gross, 2003; Toussaint, 2006; Ziemke et al., 2005). In these studies multiple candidate long-term plans are generated, for example by chaining forward models (Hoffmann and Moller, 2004), and then compared for the sake of selecting the best one. Another possibility is to generate long-term predictions related to current active courses of action in order to receive 'feedback from the future', e.g. with a mechanism similar to the *somatic markers* (Damasio, 1994), as done by Shanahan (2005). A methodology for using simulate planning in order to select the action to perform in the real world is described in (Vaughan and Zuluaga, 2006), too.

Active sensing and action control Balkenius and Hulth (1999) implemented several systems in which attention is conceived as a mechanism of selection for action; in this sense, attention is oriented proactively for the sake of gathering information useful for action. Pezzulo and Calvi (2006a) introduced a schema-based architecture in which predictions by the forward models are also used for orienting the motor apparatus (fovea and motors) in order to gather relevant information and to determine part of the next stimuli, like in *active sensing*. In the case of perceptual schemas, this means orienting the fovea toward relevant inputs (e.g., relevant colors and trajectories); in the case of the motor schemas, this means selecting the most appropriate motor action (e.g., specialized for following or escaping from quick or slow, big or small entities). Moreover, some perceptual and motor schemas are coupled: active perceptual schemas specialized in tracking some trajectories or colors prime motor schemas for dealing with related entities and vice versa. A significant advantage of anticipatory versus reactive strategies for action control is reported in this study for simulated agents dwelling in a complex and dynamical environment offering multiple possibilities for action. Three related architectures exploiting the principles of the ideomotor principle and the TOTE are described in (Pezzulo et al., 2007) which demonstrated the role of anticipation for the control of action in visual search and in reaching a target. In one of the models described there, schemas are arranged hierarchically; expectations produced at the high level canalize in a top-down way search at the lower level, while bottom-up signals serve mainly to confirm or disconfirm concurrent running hypotheses, like in predictive coding (Rao and Ballard, 1999).

Categorization In ART (adaptive resonance theory) (Carpenter and Grossberg, 1988) bottom-up signals and top-down expectations are used for self-organizing categories and resolving the stability-plasticity

dilemma. Pezzulo and Calvi (2006b) showed how perceptual and abstract categories can be evolved in a situated agent by means of interaction with an environment; the results are also interpreted according to the theory of perceptual symbols (Barsalou, 1999). Schenck and Moller (2006) used visuomotor anticipation for navigating visual scenes; objects and shapes are recognized by means of their compliance to certain sensorimotor transformation produced by the movements of the camera.

Grounding As we will discuss, *symbol grounding* (Harnad, 1990) is a key problem in artificial cognition; and we will argue that anticipation is a suitable solution. Roy (2005) developed a concept of grounding which depends on two mechanisms relating agent and environment: causation (from environment to agent) and anticipation (from agent to environment). According to this idea, concepts for objects which are e.g. reachable or graspable are grounded by schemas which regulate actual behavior and on the same time encode predictions the consequences an expected interaction. Hoffmann (2007) used internal simulation of possible trajectories for grounding concepts related to navigation; for example, distance from obstacles is grounded and estimated by running simulations until they encounter the obstacle. Dead-ends are recognized through simulated obstacle avoidance, while passages are grounded in successfully terminated simulations of navigation.

Imitation, Perspective taking, Joint Attention As discussed above, some social capabilities, such as imitation, perspective taking and joint attention are crucial in the development of both social and individual capabilities and are supposed to rely on a neural substrate which is common to other future-oriented activities. Many computational studies exist on these themes (Breazeal and Scassellati, 1999; Dautenhahn and Nehaniv, 2002; Demiris and Hayes, 1996; Kaplan and Hafner, 2006; Scassellati, 1999; Wohlschlger et al., 2003) and in many cases common anticipatory mechanisms have been developed for the control of action, imitation, and attention. For example, Johnson and Demiris (2005a) have developed Hammer, a system which uses several coupled inverse and forward models for generating actions and for imitating actions performed by others. In this case several forward models compete for predicting the perceptual input; the most accurate one activates its related inverse model for the sake of generating a comparable behavior (i.e. a behavior which would have produced the same perceptual input). Johnson and Demiris (2005b) demonstrated how to implement perspective taking via simulating the other’s perspective. The ingredients are similar to those described in the previous system, but an allocentric map of the scene is added on. By simulating the other point of view the system is able to take its perspective and know e.g. that while from its own perspective two objects are one in front of the other, from the other point of view they are side by side. The robot Ripley (Roy et al., 2006) is able to understand how object appear from its own as well as from another’s perspective and to fulfill requests such as *hand me the cup at my left* or *hand me the cup at your left* by running a real time, internal simulation of the scene. The running internal model of Ripley is also used for maintaining in memory the state of known objects in the environment, so it is very close to the typical small-scale models in cognitive science apart from the fact that objects are understood by the robot through anticipatory schemas describing (expected) sensorimotor consequences of possible interactions, and are not encoded as a collection of attributes or properties.

Gathering and selecting affordances The theory of affordances by Gibson (1979) proposes direct perception to the possibilities of actions and the constraints of actions; even the perception of a static scene may imply understanding the possibilities for action. This implies an implicit prediction and evaluation of the consequences that might occur as a consequence of actions. However, some affordances are more available, for example as a consequence of running intentions; this is what Wittgenstein (1953) calls *to-see-as*: a chair can be seen as a chair if the agent wants to seat, as an obstacle if it is navigating a room and as wood if it needs fire. Understanding affordances actually involves forming multiple (competing) hypotheses, taking different perspectives on the phenomena (related to different possible goals), and possibly with different levels of abstraction. This is also the point of the *multiple drafts model* proposed by Dennett (1991) which describes sensing as producing multiple interpretations: “When we hear other people talk we (mostly unconsciously) analyze and interpret the sounds they make, building interpretations of different sorts concurrently”. In a

similar way Sloman (2006) argues that *seeing involves multi-level process-simulations in partial registration using different ontologies, with rich (but changing) structural relations between levels*. In this vein Pezzulo and Calvi (2006a) describe a schema-based architecture in which multiple behaviors are concurrently maintained and evaluated against current drives and affordances, and their level of involvement in the control of action depends mainly on their efficacy in predicting. The system performs selection of affordances: for example, it *sees* a pray as a pray or as an obstacle depending on its current needs. Gorniak and Roy (2007) describe instead a system in which *language* is used for the selection of affordances.

Building up concepts and ontologies Anticipation plays a crucial role in schema-based and related architectures (Arbib, 1992; Drescher, 1991; Pezzulo, 2006; Roy, 2005), which in some cases can bootstrap their ontologies by interacting with the environment, like in constructivist theories (Bickhard, 1998; Piaget, 1954). Drescher (1991) investigated how to build sensorimotor schemas through interaction with a simple environment. Schemas have the form of action-effect rules; moreover, the agent interactively enlarges its ontology by learning new *synthetic items* which are conceived as the common cause of a set of related interactions. This approach also permits to model why learning and abstraction is discontinuous. Learning new synthetic items, in facts, opens the possibility to learn more; for example a children can learn new competences or about new kinds of objects or processes, and those representations can scaffold new learning. Roy et al. (2006) have built a robot which is able to build up representations of objects based on their sensorimotor structures; and to exploit (e.g. associate) those representations in order to manipulate (e.g. pick up) the objects, on the basis of verbal instructions which are understood in sensorimotor terms, too. Kuipers et al. (2006) proposed a framework for bootstrapping perceptual representations whose main criterion is predictive power.

After reviewing many cognitive functions based on anticipation, both in natural and artificial systems, it has to be noted that many of them make use of the notions of *internal models* and *representations*, which are typically supposed to be needed for future-oriented behavior. However, the concept of representation as used in cognitive systems and AI is quite problematic; here we discuss its main features for the sake of understanding if it can be maintained as an explanatory construct or not.

4 Features and Problems of Representations

Representations are often conceived as ‘vicarious’ stimuli, i.e. internally stored states which serve to activate and regulate behavior. However, this simple definition fails to discriminate between *information states* and true *representations*; or, according to Clark and Grush (1999), between *presentations* and *representations* (see also Grush, 1997).

Information states *guide our activities by creating closed loop interactions*; an example is the temperature as used by a thermostat. They have two of the characteristic features attributed to representations:

- *Internal storage*. They are often (but not always) related to some internally stored variable (or set-point of the system), although they retain a close coupling with the environment. The mechanism can not work properly if the coupling is broken (in some rare cases, if the external stimulus temporarily lacks there is the possibility to recover).
- *Causal power*. They activate and/or regulate behavior. Properly speaking, it is not the temperature of a room which activates the thermostat, but the mismatch between the sensed and the desired temperature.

However, information states lack two other relevant features:

- *Intentionality*. They *are about, stand-in* for something else. According to Brentano (1985): “As every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here as meaning

a thing), or immanent objectivity. Every mental phenomenon includes something as object within itself [...]”. Internal storage per se is not sufficient to guarantee intentionality. Representations are not required to be related to the *content* of the represented entity (it is not required that a representation of fire is hot), and in fact they are often supposed to be arbitrary and amodal symbols. However, they must have a specific *functional relationship* with the represented entity (such as to covary in a certain way). If all functional relationship is broken, there is the risk of losing aboutness and grounding, as we will discuss.

- *Detachment*. They can be used by the agent even *in absence* of their reference. According to Taylor (1971): “To be able to talk about things is to be potentially aware of them outside of any particular transaction with them; it is to be potentially aware of them not just in their behavioral relevance to some activity we are now engaged in, but also in a *disengaged* way”. Disengagement has two facets. Firstly, it permits to break the sensorimotor coupling to an extent which is impossible for information states: representations are conceived for dealing with the absence of their referent, thus mechanisms based on them are robust with respect to lacking stimuli. Secondly, while information states serve only to act on the environment (in the case of a thermostat, raise its temperature), an agent can act directly on his representations *instead of acting on their reference*. For example, an agent can build up, evaluate, select or discard one of his plans before acting in the world; Popper (1996) says that it can *let its hypotheses die in its stead*.

Adapted vs. Cognitive Systems The crucial difference between an *adapted* and a *cognitive* system thus is the presence of representations and the capability to internally manipulate them. An adaptive system can have and exploit information states, while a cognitive system also has (and manipulates) detached representations. The former entails a closed-loop with the environment and is therefore only capable of on-line processing, while the latter develops decoupled, stand-in representations that can be manipulated even in the absence of stimuli and is henceforth capable of off-line reasoning and deliberation³. Flexible manipulation of detached representations is particularly relevant for the realization of goal-directed action, that involve evaluating possible future outcomes and coordinating one’s own actions in the present and the future for realizing (one of) them. If we assume, for instance, an ideomotor functional structure of action (with action-effect relations), representing the future is necessary in order to trigger and control present action(s).

4.1 Representations under attack

The traditional notion of representations as used in cognitive science and in AI has been widely attacked from many perspectives; for useful reviews, see (Clark, 1998; Markman and Dietrich, 2000). Here we review two related problems which, as we will see, can lead either to rethink or to abandon the concept of representation.

4.1.1 Representations, as currently conceived, are not grounded

There seems to be a trade-off between intentionality and detachment as previously described, since the *aboutness* implied in the former requirement presupposes a specific functional relationship which at least suggests that the representation should retain some of the characteristics of the represented entity. On the contrary, the second requirement seems to indicate a high degree of arbitrariness which, however, has a price. There is the risk for representations to be so detached to lose contact with reality: this is the *symbol grounding problem*, formulated by Harnad (1990). Amodal, arbitrary symbol systems, especially those encoded directly by the designers, are exposed to the grounding problem since they are not constrained by physical interaction with the environment. In this case, it is argued, detachment from the sensorimotor cycle produces arbitrary symbols which lack intentionality and aboutness. The trade-off can only be resolved by indicating a suitable functional relationship between representation and environment permitting intentionality and, on the same

³ Gardenfors (2004) uses a similar notion of detachment for distinguishing agents which are able to refer to situations which are not motivated by actual or recent stimuli.

time is not too strict (coming back to information states) to deny detachment. In this sense the proposal of Harnad (1990) to ground all representations on sensorimotor features seems to be too restrictive. For example, Chalmers (1996) argues that other forms of grounding are possible. Since the structure of some mechanisms for manipulating concepts (such as connectionist ones) carries on semantic content, there is the possibility for some concepts to be grounded not via interaction with external reality but to have ‘internal grounding’. As we will see, alternative proposals exist which imply ‘extending the notion of mind’.

4.1.2 For many tasks we don’t need representations because agent-environment coupling is sufficient

According to the European phenomenologists Merleau-Ponty (1945), Husserl (1982) and Heidegger (1977), prior to any form of mental life there is an embodied interaction with the world. Maturana and Varela (1980) stress the fact that a living being and its environment are in fact not two systems but a unique system which has to be understood in a unitary way; as von Uexküll (1985) puts it, the animal and its *umwelt* are not separable. All the field of situated, embodied cognitive science (Brooks, 1991; Chiel and Beer, 1997; Nolfi, 2005; Pfeifer and Scheier, 1999; Varela et al., 1992), stresses three fundamental elements which were missing in cognitive science and AI: (1) *Situatedness*. the agent is situated in an environment and interacts continuously with it. (2) *Embodiment*. the body produces sensorimotor loops consisting in generating motor actions which produce and guide toward processing of stimuli. (3) *Emergence*. behavior is the emerging result of fine-grained and dynamical interactions with the environment.

This view implies that the brain, the body and the environment can not be studied in isolation, since they form a unique system, and their *coordination* is considered of great importance. While classical AI took as paradigmatic of intelligence the process searching in the space of internalized representations, several researchers in the field (Brooks, 1991; Beer, 1995; Port and van Gelder, 1995) show that many tasks which were conceived as representational can be executed in a more plausible and robust way without internally representing the environment but only by being coupled, engaged with it. Representations are in fact not only redundant, since coupling is sufficient for many tasks; even worse, they are considered bad: processing internal representations is too slow to permit coordination to arise, and this is why model-based AI experiences so many problems. For the philosophically inclined researches, this fact depends on an even deeper problem: since representations often consist of arbitrary symbols which are not grounded (Harnad, 1990), processing them produces results that are not useful for realizing behavior. This is why systems based on internal representations are brittle and not adaptive.

Empirical evidence supports the claim that for many tasks we don’t exploit a rich internal representation of the external world, or at least our internalization is much less widespread or relevant than assumed before; as a paradigmatic example, *change blindness* (Simons and Levin, 1997) is often cited. The proponents of the ecological approach to *direct perception* such as Gibson (1979) argue that the environment makes its *affordances* (i.e., interaction potentialities) directly available to us. This means that there is no need of having two processes, the former for representing (or differentiating) a possible interaction, and the second for performing the actual interaction; but there is only one process which is the (inter)action.

In a similar way, Bickhard (1998) offers an interactivist account of objects: the discovery and exploitation of possibilities of interaction rather than an internal storage of representations to be acted on. One important feature of this view is that by exploiting some opportunities for interaction, some other become available; for example, by taking an apple from a bag, some affordances such as *eatability* become available which were formerly unavailable.

The close relation between manipulation and the perceptual discovery of the ‘hidden’ structure of the environment was, again, underlined by Gibson (1979): perception is not the response to a stimulus but the act of picking up information. For this reason, it is incorrect to conceive a sensorimotor system as lacking memory; the system and its environment are a unique system and memory is not internalized but provided by the environment in the engagement and manipulation process. By acting in the right way, in fact, the system can access all the information it needs at the right moment: this is why Brooks (1991) says that those systems can *use the world as its own best representation*. The point raised by these theories is that even part of anticipatory behavior can be explained without the costly activity of producing expectations, for example by partially selecting the next stimuli in active vision (Brooks, 1991; Nolfi, 2002). Recently O’Regan

and Noe (2001) propose similarly that perception is a way of exploring the environment by exploiting the sensorimotor contingencies, i.e. the structure of the rules governing the changes in sensorimotor patterns which are caused by motor actions. Thus, actions in the environment assume both a *pragmatic* function, to fulfill a task, and an *epistemic* function, to discover or make it available some information which is useful for that task.

In the dynamical system literature (Beer, 1995; Port and van Gelder, 1995; Thelen and Smith, 1994) the most famous example of a system based on coupling and not representations is the Watt Governor. In a thought experiment van Gelder (1995) argues that it performs a ‘cognitive’ task by using the same kind of physics which underlies dynamical systems, while a sophisticated system would be required in order to implement the same function computationally. The computational approach is thus inessential for these tasks; whether they can be considered really cognitive is however questionable, see (Clark and Grush, 1999).

Taking at the extreme consequences this view, in philosophy eliminativists such as Churchland (1992) adopt a reductionist perspective and propose to eliminate the entire mentalistic lexicon of folk psychology, including not only representations but also beliefs, goals, etc. in favor of a (hopefully forthcoming) full fledged explanation of mental phenomena provided by neurosciences of perhaps physics. This kind of attack to representations is thus ontological, denying autonomy of a mentalistic account of mind⁴.

How can the concept of representation deal with these two criticisms? As previously stated, there are two possible ways currently explored, abandoning or rethinking them. These opportunities are now discussed, especially with respect to the problem of how they account for anticipation.

5 Abandoning representations: Dynamical Systems

The dynamical system approach (Beer, 1995; van Gelder, 1995) is focused on how continuous reciprocal causation between brain, body and environment gives rise to complex dynamics which make the behavior of an agent robust and adaptive. Representations or other kinds of stored states are often not an issue at all in dynamical systems literature. Of course, if it is well known that in order to describe a system as dynamical we must firstly extract which variables to track; but what is considered relevant is what changes and the characteristic patterns of transformation which give rise to some regularities such as attractors. For example (Schöner and Kelso, 1988; Schöner, 2002) propose to use dynamic fields as the meaningful variables; however the idea here is very far from internal states as representations are often conceived. There are two reasons to investigate this literature. Firstly, dynamical systems are being increasingly successful in explaining a variety of natural phenomena with an elegant and parsimonious theoretical apparatus. Secondly, they provide the most comprehensive account of how representations can be avoided; in a sense, they have shown how to really implement many ideas which were already under development but deserved a precise foundation.

5.1 The alternative to representations to realize anticipation

Anticipation is a rare issue in the dynamical systems literature, with some exceptions, most notably the theory of Kelso (1991) about *anticipatory dynamical systems*. Even if the mathematical apparatus permits to model selection among multiple behavioral possibilities, in fact, typically the temporal horizon of choice is limited to the present; a tight coupling with the present is in fact very relevant for achieving coordination. Dynamical systems theories (e.g., Port and van Gelder, 1995) typically focus on examples which Clark and Grush (1999) call *not representation-hungry*. The question whether or not dynamical systems can scale up to representation-hungry problems is still waiting an answer; dealing with the future is a good example of this kind of problem, and in fact future-oriented behavior is typically the province of representational systems. Thus, whether or not dynamical systems can account for anticipatory behavior in a satisfactory way is a crucial issue for deciding whether or not the notion of representation has to be abandoned.

⁴ It is worth noting that this perspective is at odds with the current view about emergence and self-organizing systems (Haken, 1988; Kelso, 1995) in which multiple levels of reality and of explanation are recognized to be autonomous (without denying interdependences), against any reductionism. This does not mean that in principle any level of explanation is correct, but that those who prove to be good according to the usual scientific criteria, are guaranteed a role in science: and mentalistic theories have of course to pass this test.

Recently Keijzer (2001) accepts this challenge and asks: “can dynamical systems deal with distal, future events?” He thus proposes the most comprehensive theory by now about how to extend the notion of behavior used by dynamical systems to be compatible with the realization of *distal* events; the term *anticipation* is used in this sense. More precisely, his question is: *how can we be guaranteed that a distal goal is realized by a series of proximal actions?* For example, how can I say that my hand moving toward a glass is part of or is involved on, or is needed for, taking the glass? We think that the problem is well posed, since the point in anticipatory behavior, as we will claim, is that we act in some way to coordinate with a future effect which we want to achieve, and we do so via proximal acts which are however intertwined with the distal effect (thanks to the intrinsic teleonomy of action which guarantees the instrumentality of the proximal acts with respect to the distal effect). Typically in a representationalist framework a crucial role is assigned to anticipatory representations of the distal effect that guide action; is there an alternative, not representational way?

Elaborating on theories of self-organizing systems and morphogenesis (Kauffman, 1993), Keijzer (2001) proposes that in analyzing natural systems it is fundamental to focus on their evolved, self-organized *structure*. The point is that in nature effects at small time scales, such as patterns of motion, auto-organize and realize higher-level (and longer time-scale) behavioral effects such as reaching an object (in the self-organization literature these would be called *order parameters* or *collective variables*) which on the turn *enslave* them in a form of circular causality (Haken, 1988; Kelso, 1995). Looking at behavior as an emergent phenomenon is also crucial for building artificial systems. The traditional encodingism in AI has focused directly on the macro, or behavioral, level. But, if behavior is an emergent property, it can only be effectively realized by reproducing the significant interactions between brain, body and environment occurring at the micro-levels of organization. Promising directions are for example evolutionary techniques, in which multiple levels of behavioral organization emerge (Chiel and Beer, 1997; Nolfi, 2005).

Modeling all the levels of organization has also an important functional role with respect to realizing distal goals. From the evolutionary and self-organizing processes a structure emerges which ensures stable and functioning behavior under a wide set of starting conditions. Proximal actions can be concatenated and realize a distal effect because evolution has shaped the right structure. Paralleling morphogenesis, behavior realizing a distal effect emerges as the result of self-organization occurring at different time scales: *a regular trajectory (behavior) spanning several scales of organization, emerges when the dynamics at multiple scales became coupled to one another* (Keijzer, 2001).

In this framework there is no need of a *representation* of the goal state; on the contrary, like in morphogenesis, the behavior can be regulated by *internal control parameters (ICPs)* situated at the lower level and the guarantee for the realization of the goal is in a sense encoded in the (evolved) structure. The term *control* is misleading, since ICPs play the role of behavior *modulators* by causing bifurcations in the dynamical systems, and not controlling it. Typically ICPs are non-specific, in the sense that they do not encode directly the result of their action, but nevertheless produce qualitative changes in behavior. An example in morphogenesis are enzymes which have a modulatory role in the growth processes; another example is walking speed, which causes qualitative changes in the gait, e.g. walk, trot or gallop in horses.

The sophisticated notion of behavior proposed by Keijzer (2001) thus avoids representations while successfully describing how an evolved dynamical system can show anticipatory behavior, at least in the sense of realizing distal goals. He concludes that, up to the moment, humans behave and robots simply move exactly because of structural differences between their motor apparatus. However, a suitable research program (such as (Beer, 1997; Nolfi, 2005)) is in principle able to account also for distal goals in artificial systems.

A similar account of suitable ways of dealing with the future without any anticipatory representation is provided by O'Regan and Noe (2001). According to their sensorimotor view, perception is coordinating with the dynamical structure of sensori stimuli, which is acquired by learning structure in stimulation from the environment. This activity is clearly anticipatory in the sense that the sensory apparatus learns to attend to relevant stimuli by anticipating e.g. the movement of an object or the patterns of transformation of sensory stimuli under a certain motor operation. However, there is no explicit representation of stimuli. For example, in tracking a target the agent learns to foveate beyond the target in anticipation of its movement, but it has no anticipatory representation of the next visual stimulus. According to this *active vision* (Ballard, 1991; Gibson, 1966) account of perception, anticipation only relies upon a learned pattern of sensorimotor

transformations, and the environment is used as *the best representation of itself*. We will call this form of anticipation *behavioral* or *implicit*.

5.2 Anticipation in a dynamical systems approach to robotics

The mathematical apparatus of dynamical systems has been widely used in robotics and research programs have been established in this sense (Beer, 1997). Anticipation has rarely been investigated in this field, even if the notion of *central pattern generators* (Delcomyn, 1980) which is widely used entails a ‘productive’ view of behavior and control. However, in recent robotic studies top-down, anticipatory signals have been used to select the appropriate internal control parameters which can cause an adequate bifurcation in the dynamical system, by activating opportune ICPs (called *parametric bias* in (Tani, 2003; Tani et al., 2004) or *neuromodulatory bias* in (Kondo and Ito, 2006)).

For example, Tani (2003) describes a system in which two recurrent neural networks operating at different hierarchical levels are coupled. The former (higher level) network predicts the sensorimotor flow and influences the latter by setting its internal control parameters, called *parametric biases*; depending on the values of the biases the network manifests one of its behaviors encoded in the multiple attractor dynamics. This architecture can also serve for recognizing the actual sensorimotor flow: in this case, the lower level network sends back to the higher level the values of the biases (modulated by an error signal, i.e. the discrepancy between predicted and actual sensory input) which are thus used for classification.

In a related work, Paine and Tani (2005) show how a similar hierarchical structure can self-organize in a continuous-time recurrent neural network by simply introducing topological constraints (a bottleneck dividing the network in two parts, whose neurons roughly play the role of parametric biases). In that architecture the higher level part of the network adapts to operate at slow time scales and to represent macro-actions, also modulating the dynamics of the lower level part of the network, which in turns operates at fast time scales and encodes primitive behaviors.

These studies show that anticipatory mechanisms can play a fundamental role even if they are not conceived as producers of representations but as modulatory mechanisms. In this case the ICPs are as usual *non-specific*, since they do not represent the distal goal to achieve. However, it is worth noting that ICPs are *selected by evolution for* (or, in these studies, learned for) achieving such a specific goal, and they do it by selecting appropriate behavior. This is why *functionally* they have the same role of anticipatory representations: guiding toward a distal goal (and, in the meantime, being sensitive to anticipated phenomena). Their functional role is provided by evolution and not by design.

This is in brief the alternative offered by dynamical systems to the notion of representations for realizing anticipatory behavior: non-specific ICPs modulating a behavioral structure (encompassing many levels and time scales) which is evolved for realizing specific distal goals. And we have shown that this option is able to account for at least some facets of anticipatory behavior.

6 Rethinking representations

As above discussed, the notion of representation as traditionally conceived in cognitive science (i.e., internal reconstruction of external reality) received several attacks. In order to maintain its role as an explanatory concept, two challenges have to be addressed:

- there is a notion of representation which maintains its peculiarities such as detachment without occurring in lacks of intentionality and grounding;
- there are cognitive functions which are only possible by using representations. As already discussed, the focus on the power of reactive behavior for producing phenomena which were conceived as representational has produced very good results but has also lead to neglect other crucial kinds of anticipatory behavior in natural cognition.

Recently many significant theoretical attempts to re-introduce representations in cognitive robotics have been formulated. For example, Steels (2003) argues, against the representation-less approach in behavior-based robotics (Brooks, 1991), that representation-making as a crucial bootstrapping device for higher mental function; for example, an internalization of external representation-making can be the basis of an ‘inner life’ for an agent, including inner dialogs and mental imaging. Clark and Grush (1999) propose that anticipatory representations, conceived in the framework of the above discussed simulative and emulative theories of cognition, will permit to cognitive robotics to scale-up from sensorimotor coordination to high level cognitive capabilities. In the spirit of these proposals, here we will defend two claims: that anticipatory representations are detached without losing intentionality and grounding; and that anticipatory representations (and, as far as we know now, only them) permit to *coordinate with the future* and to act goal-directed. Note also that our emphasis on goals and goal-directedness indicates that ‘representation’ has not (always) to be intended as a synonym of ‘knowledge’ but it is deeply related to (potential or actual) action.

6.1 Anticipatory representations are detached without losing intentionality and grounding

Recently in cognitive science the notion of representations is being revised, basically accepting many of the critics about grounding and situatedness. As above discussed, the critical point is that representations can lose contact with reality; as we have argued, there is a risk in the process of detachment to lose intentionality (and thus causal power, too). In order to resolve this problem, amodal and abstract accounts of representations such as the *language of thought* (Fodor, 1975) and the *physical symbol system* (Newell, 1990) should be abandoned in favor of different alternatives. The problem with these theories is that detachment entails acquiring arbitrariness, and this entails losing intentionality and grounding, too. The challenge now is finding other ways to resolve the trade-off between detachment and intentionality. Many converging perspectives have been proposed.

6.1.1 The situated perspective

Barsalou (1999) proposes that in order to account for *situatedness*, we have to abandon the idea that representations are amodal and arbitrary; see also (Barsalou et al., 2003). In opposition to the *physical symbol system*, he proposes a *perceptual symbol system* that is composed of modality-specific representations. It can also form multi-modal arrays, the *simulators*, which are productive structures able to generate virtually unlimited simulations of the phenomena they refer to. Barsalou thus stresses a situated and simulative vision of cognition and discusses the use of anticipatory representations for action and the influence of situated action in higher-level cognition. A similar situated approach to grounding high-level cognition is proposed by cognitive linguists such as Lakoff (1987); Lakoff and Johnson (1999). Schematic structures such as *part-whole* and *origin-trajectory-destination*, the ‘image schemas’, acquire grounded structure via sensorimotor interaction and transfer it from concrete to abstract domains in a constructive process of building metaphors. Gallese and Lakoff (2005) propose an account of how this process could be implemented in the brain by exploiting mirror neurons.

6.1.2 The externalist perspective

Even Clark (1998) tries to formulate a new, proper concept of representation which is not at odds with the traditional roles attributed to it, such as the possibility to manipulate them outside the sensorimotor loop, but nevertheless is able to relate them to situated action in order to avoid the symbol grounding problem. In order to do so, the mind has to negotiate the meanings it carries on with the external reality: this is where *externalism* comes in play.

Fodor (1980) provided an internalist theory of mind: the content of an agent’s intentional states depend only on the formal properties of the representations they are related with and not on any reference or relationship with the external reality (as the title says, it is a form of solipsism). This representational view

is coupled with a computational theory of how representations are manipulated: mental processes manipulate symbol-like representations only via their syntactic properties and have no access to their semantic content.

On the contrary, for externalists such as Ryle (1949) the representational content of an intentional state is a relational property: it depends on an object or event which is external to the subject holding the representation (i.e. to his brain and/or body). This is not the case for internalists, for which the external reality is obviously relevant as a provider of inputs, but irrelevant for determining the content of intentional states, for which only internal, non-relational criteria are sufficient.

In an externalist perspective representations can be conceived as a way of coordinating with the environment which is selected by evolution. Externalism shifts the focus from reasoning to acting and from the internal, solipsist activity *to reason about the world* to the external perspective of *reasoning through the world*. The mind is no more only an ‘inner’ activity, but a matter of what we do and how we behave. However, the role of representations in cognitive processes is not denied; what is needed is *extending the mind*.

There are two main possibilities to do that in an externalist perspective. The former consists in *going ‘outside the skull’*. According to the *extended mind* proposal by Clark and Chalmers (1998) many external phenomena (most notably artifacts) are proper part of the cognitive mind since we operate many of our cognitive processes through them. It has been also shown that reasoning with graphs and diagrams (Gattis, 2001) facilitates cognitive processes. The latter consists in *extending cognition including social dynamics*. the social dimension is crucial, since cognition grows and develops in a social environment (Vygotsky, 1978; Wittgenstein, 1953); some mechanisms crucially involved in development such as joint attention and imitation are social (Tomasello et al., 1993); some actions we perform are intrinsically social (Conte and Castelfranchi, 1995); and a collective negotiation for the selection of shared symbols (words) and their grounded meanings has been suggested (Cangelosi, 2006). Cultures also provide very powerful *artifacts*, too, such as language, institutions and practices (Tomasello et al., 1993) which, according to Vygotsky (1978), can be partially *internalized*. Not only we typically build external artifacts which are indeed affordable by us, but we can build up internal reconstructions of external reality; and many goals and conducts can have a social origin. Internal and external manipulation and simulations can overlap, too.

6.1.3 The interactivist perspective

In a way which is similar to externalism, the relationship between representations and reality is not conceived as mirroring but as a functional adaptation by constructivists such as Piaget (1985), interactivists such as Bickhard (1998) and enactivists such as Maturana and Varela (1980). This entails rejecting encodingism, which is typical of a large part of AI, in favor of a concept of representation which is *for* and *usable by* the embodied agent (Bickhard and Terveen, 1995). The external reality is not internally represented but acted upon via appropriate interaction-enabling neural structures. For example, neural maps as traditionally conceived by neurobiologists have not to be conceived as representations in the sense of corresponding to an external reality. On the contrary, their primary function is to permit an organism to coordinate and make sense of its environment and they can only be understood in the framework of this dynamic interaction. Only on the basis of these forms of interactions higher level cognitive capabilities can grow up.

6.1.4 The developmental perspective

While cognitive science has often focused on the mind as a ‘final’ product of evolution, the diacronic perspective can be extremely useful. In this sense, the proposal is *understanding mind in a larger temporal window which includes learning and development*. Mind and representations are not monolithic but they grow through ontogenetic and phylogenetic processes and can thus only be understood in a larger temporal window (Piaget, 1985), and behavior only constituted of parts which are constructed by an evolutionary process (this is akin to what described earlier about dynamical systems, but in a representational perspective). The role of developmental and epigenetic processes has also been recently recognized in robotics (Lungarella et al., 2004; Weng and Zhang, 2002; Zlatev and Balkenius, 2001). In order to understand developmental processes, a deep investigation of emerging and self-organization phenomena (Haken, 1988; Kelso,

1995) is needed too, and in particular how it is realized in biological systems (Camazine et al., 2001).

6.1.5 The emotional perspective

A different approach is to re-relate representations and emotions, which have been often conceived in isolation in traditional cognitive science. Often representations, including very simple ones, carry on a *value* for the individual, for example with respect to the satisfaction of its drives. Organisms not only perceive the environment, but *feel* part of it, and in particular its bodily states; it can become able to relate its actions to actual or anticipated feelings, as in the case of the *somatic markers* (Damasio, 1994). The inner life of an organism does not only include a coupling between perception and action, but also the internal, sensed states have a crucial part in this process. Representations thus can be fully grounded only by referring to the value they have for the organism, and not only to the power they have to produce a given behavior, since each conduct is ultimately driven by its value.

6.1.6 The anticipatory perspective

As argued in the literature about dynamical systems, mind, body and environment form an unique, coupled system. However, as we will argue, a cognitive mind is not limited to be coupled with the present, but can also *coordinate with the future* by means of anticipatory representations. In this sense, an essential part of the concept of representation is that it permits to *extend the temporal framework of coupling*, and to conceive equally the present and the future. Clark and Grush (1999) propose to view representations as oriented toward the future. In *emulation* and *simulation* theories, a mechanism permitting to re-enact sensorimotor structure used for actual interaction and exploit them ‘in simulation’ has been widely proposed (Gallese and Goldman, 1998; Grush, 2004; Hesslow, 2002). The future orientedness of representation has also been recognized in different philosophical traditions such as husserlian phenomenology, where the perception of the *present moment* has an anticipatory aspect: *being aware of what I have just seen, heard or thought retention as well as anticipating somehow a continuation of what I am now seeing, hearing or thinking protention* (Stern, 2004). Pezzulo and Castelfranchi (2007) investigated the process permitting representations to *detach* from the sensorimotor cycle without losing intentionality and grounding and propose that the capability to formulate expectations and to run internal simulations give rise to ‘stages’ in which increasingly sophisticated cognitive functions can be realized by the organism. Conceptual structures of a cognitive agent can also include ‘nested’ simulations (Grush, 2004) to conceive phenomena which progressively achieve higher levels of abstraction and detachment.

Anticipatory representations are grounded All the proposals sketched here for rethinking representations, in brief, try to resolve the above described trade-off between detachment and groundedness and to explain why detachment provided by representations is nevertheless guaranteed to preserve intentionality. The mental dimension is extended to external reality, or to social reality, or to developmental processes, or to a larger temporal dimension including the future. All these dimensions also overlap to some extent; for example, it is suggested that social mechanisms of imitation and mimesis, which depend on anticipatory capabilities, are crucial for development (Zlatev et al., 2005); and that mirror neurons are a common precursor of many anticipatory and social capabilities (Rizzolatti and Arbib, 1998). Gardenfors and Osvath (2005) suggested that anticipation facilitated the emergence of communication, since communication about future goals became beneficial by enabling long-term planning. Moreover, a common computational framework (Wolpert et al., 2003) has been proposed for explaining motor control, action observation, imitation, social interaction and theory of mind.

In the rest of the paper, although all these proposals are interesting and interrelated, we will only focus on the last one, which is concerned with detachment and anticipation. We argue that there is a kind of internal representation which permit detached manipulation although retaining intentionality and grounding: *anticipatory representations*. As claimed by many authors (Barsalou, 1999; Grush, 2004; Hurley, 2005; Pezzulo and Castelfranchi, 2007), they crucially depend on mechanisms permitting to *emulate* the external reality. Grush (2004) suggested that representing consists for the agent in emulating its environment

(or at least some of its aspects); this can happen both if the emulated entities are indeed present or absent. This capability permits not only to stay attuned with the environment -this could be possible even with the tight sensorimotor coupling which is typical of dynamical systems. Representations are instead needed in order to operate *in absence* of their referents, or in order to emulate and predict their behavior outside the sensorimotor cycle, outside the spatial and temporal dimensions of the interaction. As Grush (2004) puts it: “These models can also be run off-line in order to produce imagery, estimate outcomes of different actions, and evaluate and develop motor plans”.

Pezzulo and Castelfranchi (2007) described how anticipatory representations and mechanisms for manipulating them satisfy the criteria formulated earlier. Briefly, the guarantee for intentionality and grounding is the fact that those mechanisms are learned from previous experience about actions and their consequences; their original role is to control motor programs and thus retain their *aboutness* with respect to situated action. The guarantee for detachment is that the activity to emulate has successively been exploited per se, extending the time scale and the level of abstraction of the predictions, thus progressively disengaging from the sensorimotor loop. Thus, since representations evolve from information states, they are however attuned to the entity they represent. This is why cognitive systems, which evolve from adapted systems, even having decoupled representations remain attuned and grounded to their environment; and this makes them able to survive.

In the same study a possible evolutionary pathway is proposed. Briefly, first anticipatory systems have evolved their anticipatory capabilities, and in particular the ability to internally emulate a portion of external reality, for the sake of coordinating better with it. Successively a new evolutionary possibility has been offered to them: to exploit their forward models detaching from the sensorimotor cycle, for example by predicting long-term effects of possible actions, to conceive, evaluate and compare distal goals.

6.2 (Only) Anticipatory representations permit a real detachment

The second problem to address is whether representations are a good explanatory concept or if they can be safely replaced by other concepts which explain the whole range of anticipatory behavior, such as those proposed in dynamical systems. The point is: *what can representations do that other forms of coordination don't?*

As already discussed, Clark and Grush (1999) conclude that dynamical systems are currently unable to deal with *representation-hungry problems*. According to our analysis, this depends on the fact that many cognitive phenomena require detachment, and up to the moment representations are the only available option for modeling it.

Detachment in representations is highly related to *disengagement* from situated action, i.e. conceiving situations which are not here-and-now. All that seems to be impossible without building up and manipulating representations in the sense of (semi) internalized models of states of affair which are external to the agent, either temporally (e.g. in the future) or behaviorally (beyond its immediate range of action, i.e. other agents or the environment). As we have discussed, Keijzer (2001) provides examples of how that can be done, but we argue that his solution, although valid, only covers a part of the problem.

Detachment, in fact, implies manipulating internal models having dynamics which differ from the current sensorimotor flux (the term disengagement is used in this sense), although the simulative hypothesis maintains that these dynamics are evolved on the basis of the emulation of actual sensorimotor fluxes, and representing may well consist in re-enacting the same schemas used for simulated action (Hesslow, 2002; Grush, 2004).

One crucial point is that these models serve to regulate and to evaluate the action, and thus have to have a certain independence from the action to be produced: detachment is necessary for example in goals, in which a desired state of affairs is evaluated against the current one. The second point, related to the former one, is that often internal models have to be maintained concurrently with the current sensorimotor cycle; for example, having two models in order to compare them seems to be crucial both in individual cognition (e.g. comparing multiple alternative potential actions in planning or in the grounding of distance as described in Hoffmann and Moller, 2004; Hoffmann, 2007) and in social cognition (e.g. in perspective taking, see Johnson and Demiris, 2005b).

Here we provide some examples of tasks which depend on the two above mentioned conditions; and, up to the moment, have only been successfully modeled by means of anticipatory representations.

6.2.1 Conceiving the non-existent: the case of Goals

A cue for understanding *what representations are for?* is the notion of *goal*. A cognitive agent differs from a simply adapted one because it can endogenously build up representations which are adequate for their purposes and change the world according to them. In fact, there is a type of representations which do not only *describe* the world but *prescribe potential actions*: goals. Even Glenberg (1997), in asking *what memory is for?*, concludes that it is not simply a storage of past events but a way to constrain and to regulate the conduct in the future, we would say by endogenously producing goals to be achieved.

Goals are a difficulty for dynamical systems because they do not exist here and now (and in our terms imply detachment). Keijzer (2001) showed that the structure of behavior can implicitly deal with distal results, but still some crucial aspects of goals are missing. For example, humans and other living organisms are able to understand whether or not they have reached their goal, and are surprised if something unexpected happens. All that is naturally modeled in terms of explicit representations of the goal state which regulate the conduct and evaluates the current state with respect to the goal state, and much less in terms of implicit dynamics.

Expectations for regulation and evaluation of action In accordance with Castelfranchi (2005), we have defined a goal as an anticipatory representation which has the potentiality and function to *constrain and govern behavior toward its realization* (of some specific state of affairs). A goal also *implies an evaluation function*: it includes expectations permitting to evaluate states of the world with respect to the compliance with the goal. In goal-directed action then expectations are not only used for monitoring and controlling action; they also serve for checking whether or not it has to trigger actions (for example, this is not the case if the goal state is already achieved), whether or not it has to stop (for example, if the goal state is achieved, or because an independent course of actions will realize the goal by itself).

The notion of goals as defined here is different with respect to homeostatic mechanisms as conceived in cybernetics, or attractors in dynamical systems. In facts, these are models of *drives* and not *goals*. Goal-directed conducts are unique in that they consist in conceiving the non existent and operating in the world in order to realize it. Goal conditions, unlike drives, are not achieved periodically but rather as a result of a selection or a choice.

Do goals involve an anticipatory representation of the desired result? We argue that an anticipatory representation is needed for *guiding the behavior* and *deciding whether or not the goal is achieved*. It is crucial for example in the realization of a never experienced state of affairs, or in the case of goals which are not related to immediate pleasure (for which an edonistic, drive-related evaluation function could be in play), or in the case of truly distal goals which are out of reach of immediate behavior, whose regulatory function has to be maintained for long time. All that cases are typical of natural cognition. Moreover, typically goals are (partially) selected by comparing the expectations they embed, and thus representations are also involved in *choice* which is preliminary to triggering actions. All these theoretical requirements are closely related to the notions of goals as indicated in the empirical literature (Gallese and Metzinger, 2003; James, 1890; Jeannerod, 1997).

Planning vs. exploiting affordances Being able to formulate alternative courses of actions for realizing goals is another distinctive feature of detachment. Often, in fact, the environment offers affordances, but cognitive agents can adopt courses of actions which are not (immediately) related to them. This is for example the case of the Washoe monkey which has to ‘discover’ how to use a stick for catching food. It is immediately attracted by the goal (food) but it has to formulate a plan which is a detour from the most immediately perceptually relevant aspects of the task and on the same time has to use objects in the environment in a way which goes against their most immediate affordances, losing its *functional fixity* (Dunker, 1935). Another way to put it is that cognitive agents can act in order to force the world to show hidden affordances; and

typically they do that firstly acting with their representations (e.g. in simulation) and not directly in the world.

Planning for future needs Another crucial problem with goals arises when one considers distal effects which are loosely related to proximal actions. While the example of Keijzer (2001) are about realizing, or planning to realize, present needs, a different case is realizing future needs, which we humans often do. The main problem is that in this activity it is required to maintain both the situated activity (in order to maintain coordination with reality), and a model of the future (in order to guide the conduct toward the distal goal), and sometimes to compare and cross-evaluate them. It is really unclear how to maintain and exploit two *running processes* without using internal models; again, detachment is in play. Another relevant issue is that long-term plans are quite flexible structures and coordination is required not only with the distal, final goal, but even with intermediate steps which have to be evaluated e.g. in order to abandoning or modifying a plan on the fly. Representations of both intermediate (instrumental) and distal (final) goals are required to be evaluated and they can have different success criteria. Moreover, typically plans are firstly deliberated at a very abstract level, and only when the situation occurs specified in the details; in that case, multiple models having different degrees of abstraction and different time scales can require diverse explicit expectations to be evaluated differently.

A related problem: abstraction In a sense, all goals are formulated in a quite abstract way. I don't have the goal to eat exactly that slice of pizza at that exact time, but many occurrences of 'eating a pizza' satisfies the goal. There is thus the problem of how the success of such a goal can be stated. The notion of an explicitly represented goal does not imply a specific state of affairs to be anticipated or matched against an actual situation in a common sensorimotor format, which is the case in the so-called *perceptual anticipation* (Butz et al., 2003). A real matching occurs between two representations which are expressed in the same format, for example in Kalman filters, but this could not be the case for more abstract or sophisticated forms of expectations.

In the case indicated a real perceptual matching could be impossible. But an expectation can even be conceived simply as a set of prescriptions, or check-points, instead of a matching; the point is that I have to be able to discriminate whether the achieved state of affair satisfies or do not satisfies my expectation (and at the end my goal), and I know the criteria in advance. In schemas as used by (Roy, 2005; Pezzulo and Calvi, 2006a) the criterion is different: it is the *success* of the schema, i.e. the compliance of its operations in a given span of time which is functionally equivalent to achieving the goal. In an hypothetical *eat-a-pizza* schema, there is no match with a perceptual representation of the pizza to be eat, but the criterion is the compliance of a set of operations which involve a chain of expectations, finally leading to the success of the whole schema. Bickhard's interactivism (Bickhard, 1998) suggests a similar perspective: if an active interaction fails, than the 'indication for action' (and the content of the representation) is false. This model has the important property that the system can learn to exploit these failures, e.g. for assigning different responsibilities to active schemas (Pezzulo, 2006; Wolpert et al., 1995; Wolpert and Kawato, 1998). It can also be noted that, while very often sensorimotor information are involved, expectations can also be satisfied very indirectly by *signs*; for example, by the fact that 'pizza' is written on the menu we choose. This fact is widely recognized in semiotics; for example, Peirce (1940) intends signs as a mechanism for grounding ideas.

In what sense the above process implies representations, if only processes are in play e.g. in schemas? Inside schemas many perceptual matchings can be operated, but what about the success criteria of the whole goal? Is it representational or not? The idea is that in those operations the representational content of an expectation can be another running process, i.e. another emulation. The process of assigning success of a goal can be described as schemas that learn to formulate expectations about the compliance of other active schemas, or about typical patterns in their activation levels. According to Grush (2004) simulations can nest to produce increasingly abstract levels of description, in which the criteria for 'matching' are increasingly distant from perceptual matching, although they remain grounded on sensorimotor interaction, or emulation of interaction. Very abstract goals such as *become famous* can be conceived in that way, too. They do not prescribe any precise conduct but they serve as a *regulatory ideal* (Kant, 1998) for evaluating the compliance

of all the conducts of the agent. We have approached abstraction talking about goals, however abstraction exists in concepts, too; but, again, the idea is to rely on emulations in which the criteria for success and the expectations are nested. Of course, in order to become more precise the idea of grounding abstraction on multiple, nested simulations needs empirical investigation, and cognitive robotics can provide a suitable ‘simulated laboratory’. However, again here it is shown how detachment provided by emulation can indicate a suitable research direction. As an example in this direction is the tentative to realize in a schema based architecture the simulation theory of Barsalou (1999); in that study (Pezzulo and Calvi, 2006b) perceptual and abstract categories emerge during situated interaction by means of the dynamics of schemas which synchronize thanks to coherent patterns of predictions. Other examples of evolving increasingly complex concepts via interaction and simulation have been described when discussing about grounding in artificial systems (see e.g., Hoffmann and Moller, 2004; Hoffmann, 2007).

6.2.2 Perspective Taking

Even if the case of goals we have described is paradigmatic, there are many other cognitive functions which seem to require representations and detachment. For example, many cognitive capabilities involve conceiving, assuming and comparing different perspectives. We individually have the ability to formulate and compare in advance alternative plans, which are about the future and are thus detached from sensorimotor reality. Also many social attitudes such as perspective taking or mind reading seem to imply maintaining models of external reality which run concurrently with the sensorimotor cycle.

The same problem which exist in maintaining two running processes (for now and the future) in distal goals and in many social situations, such as maintaining two perspectives (for me and for you) and sometimes comparing them, as happens in conceiving other’s points of view, in taking multiple perspectives and in distinguishing first person view from other perspectives.

We think that these similarities between individual and social cognition are not casual but, again, depend on the dynamics of detachment of representations. As we have seen, in fact, many social phenomena depend on the ability to detach, such as understanding other’s points of view. In this case the aim is to realize not temporally distal state of affairs (as in goals) but *behaviorally distal* ones, in the sense that they are out of range of immediate action by the agent. Learning to coordinate with others and to conceive situations which do not depend on us are not innate, but developed abilities, as already pointed out when discussing about body schemas. According to Piaget (1954) children learn slowly to distinguish two kinds of causalities: those occurring depending on my actions, and those occurring independently of me, i.e. depending on the environment by itself. These latter are behaviorally distal phenomena, and we argue that the same mechanisms underlying detachment from current sensorimotor cycle (to conceive the future) are also involved in detaching from behaviorally in-the-range-of-my-action phenomena (to conceive others and their perspectives, or the intersubjective/objective reality). Mirror neurons have been suggested to be the common neural substrate for simulative aspects of individual and social cognition (Gallese, 2001; Hurley, 2005; Pezzulo and Castelfranchi, 2007; Rizzolatti and Craighero, 2004), and suitable mechanisms have been proposed that can realize computationally several social abilities, including perspective taking, understanding other’s actions, and imitation (Demiris, 2007; Wolpert et al., 2003).

There are many possible way to realize that detachment process. For example, the ‘Piagetian’ way (Piaget, 1954) is to realize first and foremost all that at the individual level, by developing schemas for anticipatory coordination with the environment; and than to transfer all that at the social level for coordinating with others. The other way around, from social to individual cognition, is the ‘Vygotskian’ way (Vygotsky, 1978). Gallese (2001) proposed that a shared ‘we’ comes first than ‘me’ and ‘other’. All these possibilities are interesting enough to investigate. Our guess (see Pezzulo and Castelfranchi, 2007) is that since anticipatory mechanisms could have been evolved for the need of controlling action and only after served to more complex functionalities, the first, piagetian way is at least at the beginning very relevant. This is why we stress the aspect of anticipation which is coordination with the future as more fundamental to other implications of detachment, and seems to be implied in all of them in one form or another. However, all that goes beyond the scope of the current paper. What is important here is to remark that there are many anticipatory phenomena which imply detachment and in many cases engaging in many relations with the environment, perhaps one

related to the sensorimotor engagement, another (or many concurrent other) to a future goal, and other to other agents' perspectives. All these cognitive functions require a flexible medium to manipulate for different purposes, which we argue can be anticipatory representations as conceived in simulative approaches. The examples provided in cognitive robotics indicate that the solution is suitable and worth a deep investigation in the future.

6.3 Implicit (behavioral) vs. explicit (representational) anticipation

Lastly, in order to put in a coherent frame the distinctions we have described about different kinds of anticipatory conducts, here we propose a distinction between *implicit* and *explicit* anticipatory behavior. The distinction is related to that proposed in (Butz et al., 2003), but here we use it as a demarcation between representational and not representational accounts of anticipatory behavior; in fact, only in the last case the goal state is explicitly encoded.

6.3.1 Implicit anticipatory behavior

It is not true that only an anticipatory system can show anticipatory behavior; some animal conducts (e.g. reacting to a noise by escaping) can be evolved without any mechanism for anticipating future stimuli (e.g. the future danger). Pezzulo and Castelfranchi (2007) argued that those are still *implicit anticipatory behaviors* (i.e. anticipatory conducts without an anticipatory representation), since their origin is functionally anticipatory. The animal, in fact, apparently reacts to the current stimulus (the noise) but this mechanism was evolved because of its function of dealing with a future event (the danger); thus, functionally the animal acts in view of the future even if a representation of the future is not implied.

Even a dynamical systems approach focused on evolution (such as the one proposed in Keijzer, 2001) can account for this kind of phenomena. As discussed above, an evolved structure is implicitly anticipatory, in the sense that it manages the coordination of proximal events to realize distal effects. It is thus the evolved structure that implicitly encodes anticipation and permits to be ready to coordinate with effects occurring in the future, without a need for representing them. However, we argue that not all anticipatory behavior works in that way: the successive functional steps involve explicit representations of anticipated future states (see also Pezzulo and Castelfranchi, 2007). Implicit anticipatory behavior can be the case of habits for animals, or behaviors related to homeostatic variables (even if represented as set points or attractors in dynamical systems). But if the right structure is not already in place it is impossible to use it for coordinating with a distal effect. This is the case of an agent realizing a novel goal, i.e. a situation which is never experienced. Even if structures permitting implicit anticipatory behavior may achieve a good generalization over novel situations, the range of anticipatory behavioral capabilities they can express is not comparable with that realized in natural cognition.

6.3.2 Explicit anticipatory behavior

Butz et al. (2003) proposed to distinguish two forms of anticipation implying explicit representations, *perceptual* and *state-based* anticipation. Many cognitive functions are not possible without a state-based anticipation such as planning via mental simulations, and conceiving novel goals. In the case of goals, when a novel state of affairs is desired and has to be achieved by an agent, the right structure to deal with future stimuli has to be shaped on-the-fly. In order to do so, the agent has to anticipate (at least to a certain degree of abstraction) the state of affairs to achieve in order to prepare its resources; once this is done, the goal state acts as a regulatory mechanism for the system.

The distinction between implicit and explicit anticipatory behavior serves to draw a distinction between different kinds of anticipatory phenomena all existing in nature, and to analyze how they are modeled in artificial cognition. In brief, the whole spectrum of anticipatory capabilities occurring in nature is, up to the moment, only possible by means of building up and maintaining (partially) internal models of external reality and manipulating them in ways which are *disengaged* from situated action: this is the case of explicit anticipatory behavior.

It is important to remark that, in claiming the relevance of anticipatory representations, we do not want to deny the fundamental role of dynamical aspects in individual and social cognition. On the contrary, this analysis might lead to investigating intermediate solutions between representational and dynamical approaches. This is also suggested by Clark (1998), who proposes the *minimal representationalism* implying that cognition is both a dynamical and representational phenomenon: *minds may be essentially embodied and embedded and still depend crucially on brains which compute and represent*.

One possible solution about goal oriented behavior is sketched here: the mind is capable of endogenously generating representations in order to influence its conduct, to prescribe actions, etc. Those anticipatory representations do not have a mysterious origin, but they are simply produced by emulation mechanisms which run out of the sensorimotor cycle. The hypothesis is that if novel states which are endogenously produced by running those mechanisms are evaluated as good (for example by a *somatic marker* mechanism), they can be stored and endogenously re-enacted for future uses. Anticipatory representations embedded in goals can be conceived as *specific ICPs* which modulate and trigger dynamical processes, which lead to the desired state of affair. There are resemblances with the approach of Tani (2003); Tani et al. (2004), the difference being that ICPs are specific and involve a representation of the goal state which triggers, regulates and stops the dynamical process (like in e.g., Pezzulo et al., 2007). An anticipatory representation of the to-be-achieved state permits not only to regulate the conduct ‘from the future’, but also to understand whether or not the goal is achieved by matching expectations with reality, being surprised if this is not the case.

7 Conclusions

Anticipation is essential in several cognitive functions. Humans and other animals can generate internal, perceptual representations of their environments for matching them against perceptual inputs while actively searching for a given object or stimulus. Similarly, several animals do not passively attend stimuli, but predict those that will probably arrive; when predictions are then matched against actual stimuli, they can be confirmed, disconfirmed, or generate surprise. Some animals can also form mental representations of the current world to work on them, virtually exploring and evaluating possible actions, events, results: *what will happen if ... ?*; or maintain concurrent representations, such as motor plans, and select among them; or maintain multiple perspectives on the same situation in order to understand the other’s point of view. Humans exploit the potential offered by anticipation with unprecedented flexibility to conceive not only the immediately perceptible world, but also other possible, future or imaginary worlds, in order to build up increasingly complex representations and act on them. For example, we can plan how to build up or restructure our house, try different solutions and compare them by acting on *models* of it, before doing it in practice; and we can also decide never to do that in practice.

Such anticipatory capabilities are considered by cognitive scientists a presupposition for autonomous mental life, since they permit to cognitive agents to *disengage* from their sensorimotor loops and to conceive and pursue their goals. Anticipatory capabilities permit to imagine, to reason about the possible and the non existent, to build up plans, to act goal-directed, and on the same time to build up deceptive and illusory worlds, to dream and to hallucinate.

Typically in analyzing natural systems and in designing artificial systems anticipatory capabilities are related to internal models and to representations. As we have discussed, however, representations have received many criticisms. One possible option is to abandon them, as proposed in the dynamical systems literature. The other one is rethinking them, trying to address their problems. We have then analyzed the two main challenges to a representationalist view.

Detachment and intentionality The first challenge is how to develop a notion of representation which permits detachment without losing intentionality and grounding. We have argued that anticipatory representations can do that: even if they are used for detaching from reality (as in visual imagery, or in planning), anticipatory representations are fully grounded, since they are originally acquired to be used in the sensorimotor cycle and subsequently, via mechanisms permitting to emulate reality, have been made available for

other, more complex uses.

What are representations for The second challenge, posed by many researchers and in particular by dynamical systems theory is: what are representations for, if tight agent-environment coupling is sufficient for accounting many, if not all, cognitive phenomena, including (at least some) anticipatory behavior? The answer is: while coupling is a fundamental mechanism for coordinating with the present (here and now), it is questionable how it can account for other capabilities of cognitive systems, for example those consisting in conceiving and acting in view of the future, i.e. to realize distal goals, or *coordinating with the future*. We have analyzed the alternative notion of anticipatory behavior (realizing distal goals) which is being developed in dynamical systems (Keijzer, 2001) and avoids representations, and we have concluded that some crucial elements of goals are still missing. We have also discussed other cases in which it is necessary to conceive (and compare) multiple perspectives at once, such as understanding other's points of view as distinct from one's own and perspective taking. So we conclude that, as far as we know by now, only a representational framework can do that. The alternative, *implicit* notion of anticipatory behavior provided by dynamical systems theory seems to be unable to explain the whole range of anticipatory capabilities existing in nature. Of course, dynamical systems are quite a novel discipline and further development is needed in order to understand if that is a temporary or intrinsic limitation of the approach.

7.1 Coordinating with the future

We have thus argued that conceiving representations as anticipatory permit to address both the challenges, and to reframe many of the classical problems in cognitive science and AI in a perspective which is compatible with the advancements in theories and techniques of cognition, both natural and artificial. One of the central constructs in dynamical system is in fact the concept of *coordination*. In this sense, the more central aspect of anticipation can be described as *coordinating with the future*: realize some actions here and now in order to coordinate with a distal effect (the goal). As above discussed, we can exploit sensorimotor loops (without representations) for coordinating with present situations; for example, active vision permits both to act and to discover information in the environment. However, descriptions of cognition which only focus on sensorimotor coupling fail to account for anticipatory phenomena which are typical of natural cognition, such as future-oriented, *distal* effects and goals. *The mind as a device whose main function is to coordinate with the future, and anticipatory representations (produced by forward models) are the means for doing that.*

As we have discussed, one strategy of the representational solution is to extend the temporal framework of cognition, including the future. It is a fact that we have evolved mechanisms for coordinating with the present. In a similar way, we have evolved mechanisms for coordinating with the future; coordinating either with future events which are predictable; or with a goal which can be endogenously generated by the agent (i.e. in absence of any stimulus). These mechanisms, which Grush (2004) describes as *emulating* the external reality, are developed for the sake of action control, but once established permit to act vicariously in a simulated sensorimotor cycle and progressively to detach from it, enabling the immense range of anticipatory behavior existing in nature. The same framework can be extended to conceive anticipation as used in social cognition. Here mechanisms for detachment from the self are involved which are closely related, if not the same, to mechanism for detaching from the present sensorimotor engagement. And, since the conceptual problem is the same, dynamical systems approaches are currently unable to deal with it, too.

Anticipatory behavior poses a big challenge for the future: it should be explained according to the four types of explanation proposed by Tinbergen (1951): adaptation (related to function, i.e. the 'why'); mechanism (related to causation, i.e. the 'how'); ontogeny (related to individual development); phylogeny (related to evolutionary development of the species). Here we have offered a tentative functional view: thanks to anticipation, actions can be selected because they are expected to be successful, and this permits to coordinate with the future. We also described some of the mechanisms for producing expectations and selecting actions depending on prediction, providing examples from natural and artificial cognition. Ontogenetic and phylogenetic explanations are still open issues (see Hurley, 2005; Pezzulo and Castelfranchi, 2007).

Investigating these four kinds of explanations is of course very important for cognitive science, since

anticipatory behavior is crucial for several cognitive functions, ranging from short-ranged control of action to realizing distal events and engaging in social dynamics. On the same time, it is highly beneficial for cognitive robotics, too, since anticipation has the potential to become a key issue in designing and developing the artificial cognitive systems of the future. Several researches in the ‘new AI’ focused on reactive mechanisms, implicit anticipatory behavior and agent-environment engagement. This has produced several results, most notably a clarification of the phenomenon of cognition and of its roots in sensorimotor interactions. Now that this approach is widely accepted, it is time to think again about how to deal with the huge variety of anticipatory and goal-directed conducts in natural systems which seem to be representational in nature, and to understand how they can emerge from more primitive forms of coordination and interaction. Our guess here is that a theoretical, empirical and computational investigation of simulative and emulative theories will permit to understand and to model the evolutionary leap from reactive to anticipatory cognitive embodied systems, and ultimately to goal-directedness in action. Anticipation will then be a crucial challenge for the future of artificial cognition.

Acknowledgments

This work is supported by the EU-funded projects **MindRACES: from Reactive to Anticipatory Cognitive Embodied Systems** (FP6-511931) and **euCognition: The European Network for the Advancement of Artificial Cognitive Systems** (FP6-26408). The author wants to thank Prof. Cristiano Castelfranchi for countless discussions and insightful comments.

References

- Adams, J. A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behavior*, 3:111–149.
- Arbib, M. (1992). Schema theory. In Shapiro, S., editor, *Encyclopedia of Artificial Intelligence, 2nd Edition*, volume 2, pages 1427–1443. Wiley.
- Balkenius, C. and Hulth, N. (1999). Attention as selection-for-action: a scheme for active perception. In *Proceedings of EuRobot-1999*, Zurich.
- Ballard, D. (1991). Animate vision. *Artificial Intelligence*, 48(1):1–27.
- Balleine, B. W. and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, 37(4-5):407–419.
- Barsalou, L., Simmons, W. K., Barbey, A., and Wilson, C. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends Cogn Sci*, 7(2):84–91.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22:577–600.
- Barsalou, L. W. (2003). Abstraction in perceptual symbol systems. *Philos Trans R Soc Lond B Biol Sci*, 358(1435):1177–1187.
- Beer, R. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, 72:173–215.
- Beer, R. (1997). The dynamics of adaptive behavior: A research program. *Robotics and Autonomous Systems*, 20:257–289.
- Bickhard, M. H. (1998). Levels of representationality. *Journal of Experimental and Theoretical Artificial Intelligence*, 10(2):179–215.
- Bickhard, M. H. (2001). Function, anticipation and representation. In Dubois, D. M., editor, *Computing Anticipatory Systems. CASYS 2000 - Fourth International Conference*, pages 459–469, Melville, NY. American Institute of Physics.

- Bickhard, M. H. and Terveen, L. (1995). *Foundational Issues in Artificial Intelligence and Cognitive Science: Impasse and Solution*. Elsevier Scientific, Amsterdam.
- Blakemore, S.-J. and Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience*, 2:561–7.
- Bongard, J., Zykov, V., and Lipson, H. (2006a). Resilient machines through continuous self-modeling. *Science*, 314(5802):1118–1121.
- Bongard, J. C., Zykov, V., and Lipson, H. (2006b). Automated synthesis of body schema using multiple sensor modalities. In *Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems (ALIFEX)*, pages 220–226.
- Breazeal, C. and Scassellati, B. (1999). A context-dependent attention system for a social robot. In *IJCAI '99: Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, pages 1146–1153. Morgan Kaufmann Publishers Inc.
- Brentano, F. (1985). *Psychology from an Empirical Standpoint*. Routledge, London. Translated by A. C. Rancurello, D. B. Terrell and L. L. McAlister.
- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47(47):139–159.
- Butz, M. V. (2002). *Anticipatory learning classifier systems*. Kluwer Academic Publishers, Boston, MA.
- Butz, M. V., Sigaud, O., and Gérard, P. (2003). Internal models and anticipations in adaptive learning systems. In Butz, M. V., Sigaud, O., and Gérard, P., editors, *Anticipatory Behavior in Adaptive Learning Systems: Foundations, Theories, and Systems*, pages 86–109. Springer-Verlag, Berlin Heidelberg.
- Camazine, S., Franks, N. R., Sneyd, J., Bonabeau, E., Deneubourg, J.-L., and Theraula, G. (2001). *Self-Organization in Biological Systems*. Princeton University Press, Princeton, NJ, USA.
- Cangelosi, A. (2006). The grounding and sharing of symbols. *Pragmatics and Cognition*, 14:275–285.
- Carlsson, K., Petrovic, P., Skare, S., Petersson, K. M., and Ingvar, M. (2000). Tickling expectations: Neural processing in anticipation of a sensory stimulus. *J. Cognitive Neuroscience*, 12(4):691–703.
- Carpenter, G. A. and Grossberg, S. (1988). The art of adaptive pattern recognition by a self-organizing neural network. *Computer*, 21(3):77–88.
- Castelfranchi, C. (2005). Mind as an anticipatory device: For a theory of expectations. In *BVAI 2005*, pages 258–276.
- Chalmers, D. (1996). *The conscious mind*. Oxford University Press, New York. Tr. It., La mente cosciente, McGraw- Hill, 1999.
- Chiel, H. J. and Beer, R. D. (1997). The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences*, 20:553–557.
- Churchland, P. (1992). *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. MIT Press, Cambridge, MA.
- Cisek, P. and Kalaska, J. (2001). Common codes for situated interaction. *Behavioral and Brain Sciences*, 24(5):883–884.
- Clark, A. (1998). *Being There. Putting Brain, Body, and World Together*. MIT Press.
- Clark, A. and Chalmers, D. J. (1998). The extended mind. *Analysis*, 58:10–23.
- Clark, A. and Grush, R. (1999). Towards a cognitive robotics. *Adaptive Behavior*, 7(1):5–16.

- Colombo, M. and Graziano, M. (1994). Effects of auditory and visual interference on auditory-visual delayed matching to sample in monkeys (maca fascicularis). *Behav Neurosci*, 108:636–639.
- Conte, R. and Castelfranchi, C. (1995). *Cognitive and Social Action*. University College London, London, UK.
- Cotterill, R. (1998). *Enchanted Looms: Conscious Networks in Brains and Computers*. Cambridge University Press.
- Cotterill, R. (2001). Cooperation of the basal ganglia, cerebellum, sensory cerebrum and hippocampus: Possible implications for cognition, consciousness, intelligence and creativity. *Progress in Neurobiology*, 64:1–33.
- Craik, K. (1943). *The Nature of Explanation*. Cambridge University Press, Cambridge.
- Damasio, A. R. (1994). *Descartes’ Error: Emotion, Reason and the Human Brain*. Grosset/Putnam, New York.
- Dautenhahn, K. and Nehaniv, C. (2002). *Imitation in animals and artifacts*. MIT Press.
- Decety, J. (1996). Do imagined and executed actions share the same neural substrate? *Brain Res Cogn*, 3:87–93.
- Decety, J. and Chaminade, T. (2003). When the self represents the other: A new cognitive neuroscience view on psychological identification. *Consciousness and Cognition*, 12(20):577–596.
- Decety, J., Grzes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., and Fazio, F. (1997). Brain activity during observation of actions. influence of action content and subject’s strategy. *Brain*, 120 (Pt 10):1763–1777.
- Decety, J., Jeannerod, M., and Prablanc, C. (1989). The timing of mentally represented actions. *Behav Brain Res*, 34(1-2):35–42.
- Delcomyn, F. (1980). Neural basis for rhythmic behaviour in animals. *Science*, 210:492–498.
- Demiris, J. and Hayes, G. (1996). Imitative learning mechanisms in robots and humans. In *Proceedings of the 5th European Workshop on Learning Robot*, pages 9–16.
- Demiris, Y. (2007). Prediction of intent in robotics and multi-agent systems. *Cognitive Processing*, 8(3):151–158.
- Demiris, Y. and Khadhour, B. (2005). Hierarchical attentive multiple models for execution and recognition (hammer). *Robotics and Autonomous Systems Journal*, 54:361–369.
- Dennett, D. C. (1991). *Consciousness Explained*. Little, Brown & Co.
- Desmurget, M. and Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends Cogn. Sci.*, 4:423–431.
- Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review*, 3:357–370.
- Doya, K. (1996). Temporal difference learning in continuous time and space. In Touretzky, D. S., Mozer, M. C., and Hasselmo, M. E., editors, *Advances in Neural Information Processing Systems*, volume 8, pages 1073–1079. The MIT Press.
- Drescher, G. L. (1991). *Made-Up Minds: A Constructivist Approach to Artificial Intelligence*. MIT Press, Cambridge, MA.
- Dunker, K. (1935). *Psychology des produktiven Denkens*. Springer, Berlin.

- Edelman, G. M. (1987). *Neural Darwinism: The Theory of Neuronal Group Selection*. Basic Books, New York.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2):179–211.
- Fleischer, J. G. (2007). Neural correlates of anticipation in cerebellum, basal ganglia, and hippocampus. In Butz, M., Sigaud, O., Pezzulo, G., and Baldassarre, G., editors, *Anticipatory Behavior in Adaptive Learning Systems: Advances in Anticipatory Processing*, LNAI 4520. Springer.
- Fodor, J. (1975). *The language of thought*. Harvard University Press, Cambridge, MA.
- Fodor, J. A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences*, 3:63–109.
- Fogassi, L., Ferrari, P., Chersi, F., Gesierich, B., Rozzi, S., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, 308:662–667.
- Friston, K. (2003). Learning and inference in the brain. *Neural Netw*, 16(9):1325–1352.
- Frith, C. (2007). *Making up the Mind. How the Brain Creates our Mental World*. Blackwell.
- Frith, C. D. and Frith, U. (2006). How we predict what other people are going to do. *Brain Research*, 1079(1):36–46.
- Gallagher, S. and Jeannerod, M. (2002). From action to interaction. *Journal of Consciousness Studies*, 9:3–26(24).
- Gallese, V. (2000). The inner sense of action: agency and motor representations. *Journal of Consciousness Studies*, 7:23–40.
- Gallese, V. (2001). The ‘shared manifold’ hypothesis. from mirror neurons to empathy. *Journal of Consciousness Studies*, 8:5–87.
- Gallese, V. and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12):493–501.
- Gallese, V. and Lakoff, G. (2005). The brains concepts: The role of the sensory-motor system in reason and language. *Cognitive Neuropsychology*, 22:455–479.
- Gallese, V. and Metzinger, T. (2003). Motor ontology: The representational reality of goals, actions, and selves. *Philosophical Psychology*, 13(3):365–388.
- Gardenfors, P. (2004). Cooperation and the evolution of symbolic communication. In Oller, K. and Griebel, U., editors, *Evolution of Communication Systems: A Comparative Approach*, pages 237–256. MIT Press, Cambridge, MA.
- Gardenfors, P. and Osvath, M. (2005). The evolution of anticipatory cognition as a precursor to symbolic communication. In Brook, S., editor, *Proceedings of the Morris Symposium on the Evolution of Language*, NY, USA.
- Gattis, M. (2001). *Spatial Schemas in Abstract Thought*. MIT Press, Cambridge.
- Gers, F. A., Schraudolph, N. N., and Schmidhuber, J. (2003). Learning precise timing with lstm recurrent networks. *J. Mach. Learn. Res.*, 3:115–143.
- Ghez, C. and Thach, W. T. (2000). *Principles of Neural Science*, chapter The cerebellum, pages 833–835. McGraw-Hill, New York, 4th edition.
- Gibson, J. (1966). *The Senses Considered as Perceptual Systems*. Houghton Mifflin, Boston, MA.

- Gibson, J. (1979). *The ecological approach to visual perception*. Lawrence Erlbaum Associates, Inc, Mahwah, NJ.
- Glenberg, A. (1997). What memory is for. *Behavioral and Brain Sciences*, 20:1–55.
- Goldman, A. (2005). Imitation, mind reading, and simulation. In Hurley and Chater, editors, *Perspectives on Imitation II*, pages 80–81. MIT Press, Cambridge, MA.
- Gorniak, P. and Roy, D. (2007). Situated language understanding as filtering perceived affordances. *Cognitive Science*, 31(2):197–231.
- Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review*, 77:73–99.
- Gross, H.-M., Volker, S., and Torsten, S. (1999). A neural architecture for sensorimotor anticipation. *Neural Networks*, 12:1101–1129.
- Grush, R. (1997). The architecture of representation. *Philosophical Psychology*, 10(1):5–25.
- Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences*, 27(3):377–96.
- Gurney, K., Prescott, T., and Redgrave, P. (2001). A computational model of action selection in the basal ganglia, i. a new functional anatomy. *Biological Cybernetics*, 84:401–410.
- Haken, H. (1988). *Information and Self-Organization, a Macroscopic Approach to Complex Systems*. Springer-Verlag, Berlin/New York.
- Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42:335–346.
- Haruno, M., Wolpert, D., and Kawato, M. (2003). Hierarchical mosaic for movement generation. In Ono, T., Matsumoto, G., Llinas, R., Berthoz, A., Norgren, H., and Tamura, R., editors, *Excepta Medica International Congress Series*. Elsevier Science, Amsterdam.
- Heidegger, M. (1977). *Sein und Zeit*. Niemeyer, Tbingen.
- Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences*, 6:242–247.
- Hoffmann, H. (2007). Perception through visuomotor anticipation in a mobile robot. *Neural Networks*, 20:22–33.
- Hoffmann, H. and Moller, R. (2004). Action selection and mental transformation based on a chain of forward models. In Schaal, S., Ijspeert, A., Billard, A., Vijayakumar, S., Hallam, J., and Meyer, J.-A., editors, *From Animals to Animats 8, Proceedings of the Eighth International Conference on the Simulation of Adaptive Behavior*, pages 213–222, Los Angeles, CA. MIT Press.
- Hoffmann, J. (1993). *Vorhersage und Erkenntnis: Die Funktion von Antizipationen in der menschlichen Verhaltenssteuerung und Wahrnehmung [Anticipation and cognition: The function of anticipations in human behavioral control and perception]*. Hogrefe, Goettingen, Germany.
- Hoffmann, J., Stöcker, C., and Kunde, W. (2004). Anticipatory control of actions. *International Journal of Sport and Exercise Psychology*, 2:346–361.
- Hommel, B. (2003). Planning and representing intentional action. *TheScientificWorld JOURNAL*, 3:593–608.
- Hommel, B., Musseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (tec): a framework for perception and action planning. *Behavioral and Brain Science*, 24(5):849–78.

- Houk, J. C. and Wise, S. P. (1995). Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. *Cereb Cortex*, 5(2):95–110.
- Hurley, S. and Chater, N. (2004). *Perspectives on Imitation: From Cognitive Neuroscience to Social Science*. MIT Press, Cambridge, MA.
- Hurley, S. L. (2005). Active perception and perceiving action: the shared circuits hypothesis. In Gendler, T. and Hawthorne, J., editors, *Perceptual Experience*. Oxford University Press, Oxford.
- Husserl, E. (1982). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy*. The Hague, Nijhoff.
- Iacoboni, M. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cereb Cortex*, 12:847–855.
- Iacoboni, M. (2003). Understanding others: Imitation, language, empathy. In Hurley, S. and Chater, N., editors, *Perspectives on Imitation: From Cognitive Neuroscience to Social Science*. MIT Press, Cambridge, MA.
- Jacobs, R., Jordan, M., Nowlan, S., and Hinton, G. (1991). Adaptive mixtures of local experts. *Neural Computation*, 3:79–87.
- James, M., Singh, S., and Littman, M. (2004). Planning with predictive state representations. In *Proceedings of the International Conference on Machine Learning and Applications*, pages 304–311.
- James, W. (1890). *The Principles of Psychology*. Dover Publications, New York.
- Jeannerod, M. (1979). Visuomotor experiments: failure to replicate, or failure to match the theory? *Behavioral and Brain Sciences*, 2:71.
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behav Brain Sci*, 17:187–245.
- Jeannerod, M. (1997). *The Cognitive Neuroscience of Action*. Blackwell, Oxford.
- Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage*, 14:S103–S109.
- Jeannerod, M. and Decety, J. (1995). Mental motor imagery: a window into the representational stages of action. *Curr Opin Neurobiol*, 5(6):727–732.
- Jeannerod, M. and Pacherie, E. (2004). Agency, simulation, and self-identification. *Mind and Language*, 19(2):113–46.
- Johnson, M. and Demiris, Y. (2005a). Hierarchies of coupled inverse and forward models for abstraction in robot action planning, recognition and imitation. In *Proceedings of the AISB 2005 Symposium on Imitation in Animals and Artifacts*, pages 69–76, Hertfordshire, UK.
- Johnson, M. and Demiris, Y. (2005b). Perceptual perspective taking and action recognition. *International Journal of Advanced Robotic Systems*, 2(4):301–309.
- Johnson-Laird, P. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge University Press and Harvard University Press, Cambridge.
- Jordan, M. I. and Rumelhart, D. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science*, 16:307–354.
- Kant, I. (1781/1998). *Critique of Pure Reason*. Cambridge University Press, Cambridge.

- Kaplan, F. and Hafner, V. (2006). The challenges of joint attention. *Interaction Studies*, 7(2).
- Kauffman, S. A. (1993). *The origins of order: self-organization and selection in evolution*. Oxford University Press, New York.
- Kawato, M. (1990). Computational schemes and neural network models for formulation and control of multijoint arm trajectory. pages 197–228.
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9:718–27.
- Keijzer, F. (2001). *Representation and behavior*. MIT Press, Cambridge, MA.
- Kelso, J. A. (1991). Anticipatory dynamical systems, intrinsic pattern dynamics and skill learning. *Human Movement Science*, 10(1):93–111.
- Kelso, J. A. S. (1995). *Dynamic patterns: the self-organization of brain and behavior*. MIT Press, Cambridge, Mass.
- Knoblich, G. and Prinz, W. (2005). *Higher-order motor disorders*, chapter Linking perception and action: An ideomotor approach., pages 79–104. Oxford University Press., Oxford, UK.
- Kondo, T. and Ito, K. (2006). A design principle of adaptive neural controllers for realizing anticipatory behavior in reaching movement under unexperienced environments. In Butz, M., Sigaud, O., Pezzulo, G., and Baldassarre, G., editors, *Proceedings of the Third Workshop on Anticipatory Behavior in Adaptive Learning Systems (ABiALS 2006)*.
- Kording, K. and Wolpert, D. (2006). Bayesian decision theory in sensorimotor control. *Trends Cogn. Sci.*, 10:319–326.
- Kosslyn, S. M. and Sussman, A. (1994). Roles of imagery in perception: Or, there is no such thing as immaculate perception. In Gazzaniga, M., editor, *The cognitive neurosciences*, pages 1035–1042. Cambridge, MA: MIT Press.
- Kuipers, B., Beeson, P., Modayil, J., and Provost, J. (2006). Bootstrap learning of foundational representations. *Connection Science (accepted for publication)*, 18(2).
- Kunde, W., Koch, I., and Hoffmann, J. (2004). Anticipated action effects affect the selection, initiation and execution of actions. *The Quarterly Journal of Experimental Psychology. Section A: Human Experimental Psychology*, 57(1):87–106.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press, Chicago.
- Lakoff, G. and Johnson, M. (1999). *Philosophy In The Flesh: The Embodied Mind and Its Challenge to Western Thought*. Basic Books.
- Landauer, T. K. (1962). Rate of implicit speech. *Perceptual and Motor Skills*, 15:646.
- Littman, M., Sutton, R., and Singh, S. (2001). Predictive representations of state. In *Proc. NIPS-02*, Vancouver.
- Lu, X. and Ashe, J. (2005). Anticipatory activity in primary motor cortex codes memorized movement sequences. *Neuron*, 45(6):967–973.
- Lungarella, M., Metta, G., Pfeifer, R., and Sandini, G. (2004). Developmental robotics: a survey. *Connection Science*, 0(0):1–40.

- Maravita, A., Spence, C., and Driver, J. (2003). Multisensory integration and the body schema: close to hand and within reach. *Curr Biol*, 13(13):R531–R539.
- Markman, A. B. and Dietrich, E. (2000). Extending the classical view of representation. *Trends in Cognitive Sciences*, 4(12):470–475.
- Maturana, H. R. and Varela, F. J. (1980). *Autopoiesis and Cognition: the Realization of Living*. D. Reidel Pub., Dordrecht, Holland.
- Mehta, B. and Schaal, S. (2002). Forward models in visuomotor control. *Journal of Neurophysiology*, 88:942–53.
- Mel, B. W. (1990). Vision-based robot motion planning. *Neural networks for control*, pages 229–253.
- Meltzoff, A. and Prinz, W. (2002). *The imitative mind: Development, evolution and brain bases*. Cambridge University Press, Cambridge.
- Merleau-Ponty, M. (1945). *Phnomnologie de la perception*. Gallimard, Paris.
- Miall, R. C. and Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, 9(8):1265–1279.
- Middleton, F. A. and Strick, P. L. (2000). Basal ganglia output and cognition: evidence from anatomical, behavioral, and clinical studies. *Brain Cogn*, 42(2):183–200.
- Miller, G. A., Galanter, E., and Pribram, K. H. (1960). *Plans and the Structure of Behavior*. Holt, Rinehart and Winston, New York.
- Mohan, V. and Morasso, P. (2006). A forward / inverse motor controller for cognitive robotics. In *Proceedings of ICANN*, pages 602–611.
- Neisser, U. (1976). *Cognition and reality*. San Francisco, CA: Freeman.
- Newell, A. (1990). *Unified Theories of Cognition*. Harvard University Press, Cambridge, MA.
- Nolfi, S. (2002). Power and limits of reactive agents. *Neurocomputing*, 49:119–145.
- Nolfi, S. (2005). Behaviour as a complex adaptive system: On the role of self-organization in the development of individual and collective behaviour. *ComplexUs*, 2:195–203.
- O’Regan, J. and Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5):883–917.
- Oudeyer, P.-Y., Kaplan, F., Hafner, V. V., and Whyte, A. (2005). The playground experiment: Task-independent development of a curious robot. In Bank, D. and Meeden, L., editors, *Proceedings of the AAAI Spring Symposium on Developmental Robotics, 2005*, pages 42–47, Stanford, California.
- Pacherie, E. and Dokic, J. (2006). From mirror neurons to joint actions. *Cognitive Systems Research*, 7(2-3):101–112.
- Paine, R. W. and Tani, J. (2005). How hierarchical control self-organizes in artificial adaptive systems. *Adaptive Behavior*, 13(3):211–225.
- Peirce, C. S. (1897 / 1940). *Philosophical writings of Peirce*, chapter Logic as semiotic: The theory of signs. Dover.
- Pezzulo, G. (2006). How can a massively modular mind be context-sensitive? a computational approach. In *Proceedings of ICCM 2006*.

- Pezzulo, G., Baldassarre, G., Butz, M. V., Castelfranchi, C., and Hoffmann, J. (2007). From actions to goals and vice-versa: Theoretical analysis and models of the ideomotor principle and tote. In Butz, M., Sigaud, O., Pezzulo, G., and Baldassarre, G., editors, *Anticipatory Behavior in Adaptive Learning Systems: Advances in Anticipatory Processing*, LNAI 4520. Springer.
- Pezzulo, G. and Calvi, G. (2006a). A schema based model of the praying mantis. In Nolfi, S., Baldassarre, G., Calabretta, R., Hallam, J., Marocco, D., Miglino, O., Meyer, J.-A., and Parisi, D., editors, *From animals to animats 9: Proceedings of the Ninth International Conference on Simulation of Adaptive Behaviour*, volume LNAI 4095, pages 211–223, Berlin, Germany. Springer Verlag.
- Pezzulo, G. and Calvi, G. (2006b). Toward a perceptual symbol system. In *Proceedings of the Sixth International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*. Lund University Cognitive Science Studies 118.
- Pezzulo, G. and Calvi, G. (2007). Schema-based design and the akira schema language: An overview. In Butz, M., Sigaud, O., Pezzulo, G., and Baldassarre, G., editors, *Anticipatory Behavior in Adaptive Learning Systems: Advances in Anticipatory Processing*, LNAI 4520. Springer.
- Pezzulo, G., Calvi, G., Ognibene, D., and Lalia, D. (2005). Fuzzy-based schema mechanisms in akira. In *CIMCA '05: Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce Vol-2*, pages 146–152, Washington, DC, USA. IEEE Computer Society.
- Pezzulo, G. and Castelfranchi, C. (2007). The symbol detachment problem. *Cognitive Processing*, 8(2):115–131.
- Pfeifer, R. and Scheier, C. (1999). *Understanding Intelligence*. MIT Press, Cambridge, MA.
- Piaget, J. (1954). *The Construction of Reality in the Child*. Ballentine.
- Piaget, J. (1985). *Equilibration of cognitive structures*. University of Chicago Press.
- Pierce, D. M. and Kuipers, B. J. (1997). Map learning with uninterpreted sensors and effectors. *Artificial Intelligence*, 92:169–227.
- Popper, K. R. (1996). *Alles Leben ist Problemlösen. über Erkenntnis, Geschichte und Politik*. R. Piper-Verlag, München.
- Port, R. and van Gelder, T. (1995). *Mind as motion: Explorations in the dynamics of cognition*. MIT Press, Cambridge MA.
- Prinz, W. (1990). A common coding approach to perception and action. In Neumann, O. and Prinz, W., editors, *Relationships between perception and action*, pages 167–201. Springer Verlag, Berlin.
- Prinz, W. (2005). An ideomotor approach to imitation. In Hurley, S. and Chater, N., editors, *Perspectives on imitation: From neuroscience to social science*, volume 1, pages 141–156. MIT Press, Cambridge, MA.
- Rainer, G., Rao, S. C., and Miller, E. K. (1999). Prospective coding for objects in primate prefrontal cortex. *The Journal of Neuroscience*, 19(13):5493–5505.
- Rao, R. P. and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci*, 2(1):79–87.
- Riesenhuber, M. and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat Neurosci*, 2(11):1019–1025.
- Rizzolatti, G. and Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21(5):188–194.

- Rizzolatti, G. and Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27:169–192.
- Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3.
- Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2:661–670.
- Roitblat, H. (1980). Codes and coding processes in pigeon short-term memory. *Anim Learn Behav*, 8:341–351.
- Rosen, R. (1985). *Anticipatory Systems*. Pergamon Press.
- Rosenblueth, A., Wiener, N., and Bigelow, J. (1943). Behavior, purpose and teleology. *Philosophy of Science*, 10(1):18–24.
- Roy, D. (2005). Semiotic schemas: a framework for grounding language in action and perception. *Artificial Intelligence*, 167(1-2):170–205.
- Roy, D., yuh Hsiao, K., Mavridis, N., and Gorniak, P. (2006). Ripley, hand me the cup: Sensorimotor representations for grounding word meaning. In *Int. Conf. of Automatic Speech Recognition and Understanding*.
- Ruby, P. and Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A pet investigation of agency. *Nature Neuroscience*, 4(5):546–550.
- Ryle, G. (1949). *The Concept of Mind*. Barnes and Noble, New York.
- Scassellati, B. (1999). Imitation and mechanisms of joint attention: A developmental structure for building social skills on a humanoid robot. *Lecture Notes in Computer Science*, 1562:176–195.
- Schenck, W. and Moller, R. (2006). Learning a visual forward model for a robot camera head. In Butz, M., Sigaud, O., Pezzulo, G., and Baldassarre, G., editors, *Proceedings of the Third Workshop on Anticipatory Behavior in Adaptive Learning Systems (ABiALS 2006)*.
- Schmidhuber, J. (1991). Adaptive confidence and adaptive curiosity. Technical Report FKI-149-91, Institut für Informatik, Technische Universität.
- Schmidhuber, J. (2002). Exploring the predictable. In Ghosh, S. and Tsutsui, S., editors, *Advances in Evolutionary Computing*, pages 579–612. Springer.
- Schomaker, L. (2004). Anticipation in cybernetic systems: A case against mindless anti-representationalism. In *IEEE International Conference on Systems, Man and Cybernetics*.
- Schöner, G. (2002). Dynamical systems approaches to neural systems and behavior. In Smelser, N. J. and Baltes, P. B., editors, *International Encyclopedia of the Social & Behavioral Sciences*, pages 10571–10575, Oxford. Pergamon.
- Schöner, G. and Kelso, J. A. S. (1988). Dynamic pattern generation in behavioral and neural systems. *Science*, 239:1513–1520.
- Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, 11(5):211–218.
- Schultz, W., Dayan, P., and Montague, P. (1997). A neural substrate of prediction and reward. *Science*, 275:1593–1599.
- Shanahan, M. (2005). Cognition, action selection, and inner rehearsal. In *Proceedings IJCAI 2005 Workshop on Modelling Natural Action Selection*, pages 92–99.

- Simons, D. J. and Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, 1(7):261–267.
- Sloman, A. (2006). Orthogonal recombinable competences acquired by altricial species. Technical report, University of Birmingham.
- Steels, L. (2003). Intelligence with representation. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 361(1811):2381–2395.
- Stephan, V. and Gross, H.-M. (2003). Visuomotor anticipation - a powerful approach to behavior-driven perception. *Künstliche Intelligenz*, 17(2):12–17.
- Stern, D. (2004). *The Present Moment*. W.W. Norton & Company, New York, London.
- Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*, 3:9–44.
- Sutton, R. S. and Barto, A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychol Rev*, 88(2):135–170.
- Tani, J. (2003). Learning to generate articulated behavior through the bottom-up and the top-down interaction processes. *Neural Netw*, 16(1):11–23.
- Tani, J., Ito, M., and Sugita, Y. (2004). Self-organization of distributedly represented multiple behavior schemata in a mirror system: reviews of robot experiments using rnnpb. *Neural Netw*, 17(8-9):1273–1289.
- Tani, J. and Nolfi, S. (1999). Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems. *Neural Netw*, 12(7-8):1131–1141.
- Taylor, C. (1971). What is involved in a genetic psychology? In Mischel, T., editor, *Cognitive development and epistemology*, pages 393–416. New York: Academic Press.
- Thelen, E. and Smith, L. B. (1994). *A Dynamic Systems Approach to the Development of Perception and Action*. MIT Press.
- Tinbergen, N. (1951). *The study of instinct*. Oxford University Press, New York.
- Tolman, E. (1932). *Purposive Behavior in Animals and Men*. Appleton-Century-Crofts, New York.
- Tomasello, M., Kruger, A., and Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences*, 16:495–552.
- Toussaint, M. (2006). A sensorimotor map: Modulating lateral interactions for anticipation and planning. *Neural Comput.*, 18(5):1132–1155.
- Tsoukalas, L. H. (1998). Neurofuzzy approaches to anticipation: a new paradigm for intelligent systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 28(4):573–582.
- van Gelder, T. (1995). What might cognition be, if not computation? *Journal of Philosophy*, 92(7):345–381.
- Varela, F. J., Thompson, E. T., and Rosch, E. (1992). *The Embodied Mind: Cognitive Science and Human Experience*. The MIT Press.
- Vaughan, R. T. and Zuluaga, M. (2006). Use your illusion: Sensorimotor self-simulation allows complex agents to plan with incomplete self-knowledge. In *Proceedings of the International Conference on Simulation of Adaptive Behaviour (SAB)*, Rome, Italy.
- Vijayakumar, S., D’souza, A., and Schaal, S. (2005). Incremental online learning in high dimensions. *Neural Comput.*, 17(12):2602–2634.

- von Holst, E. and Mittelstaedt, H. (1950). Das reafferenzprinzip. *Naturwissenschaften*, 37:464–476.
- von Uexküll, J. (1985). Environment [umwelt] and inner world of animals. In Burghardt, G. M., editor, *Foundations of Comparative Ethology*, pages 222–245. Van Nostrand Reinhold, New York.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press, Cambridge, MA.
- Weng, J. and Zhang, Y. (2002). Developmental robots: A new paradigm. In *Proceedings Second International Workshop on Epigenetic Robotics*.
- Wickens, J. (1997). Basal ganglia: Structure and computations. *Network: Computation in Neural Systems*, 8:R77–R109.
- Witkowski, M. (2003). Towards a four factor theory of anticipatory learning. In Butz, M., Sigaud, O., and Gérard, P., editors, *Anticipatory Behavior in Adaptive Learning Systems*, LNAI 2684, pages 66–85. Springer.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Basil Blackwell, Oxford.
- Wohlschläger, A., Gattis, M., and Bekkering, H. (2003). Action generation and action perception in imitation: An instance of the ideomotor principle. *Philosophical Transactions of the Royal Society of London*, 358:501–515.
- Wolfe, B., James, M. R., and Singh, S. (2005). Learning predictive state representations in dynamical systems without reset. In *ICML '05: Proceedings of the 22nd international conference on Machine learning*, pages 980–987, New York, NY, USA. ACM Press.
- Wolfe, B. and Singh, S. (2006). Predictive state representations with options. In *ICML '06: Proceedings of the 23rd international conference on Machine learning*, pages 1025–1032, New York, NY, USA. ACM Press.
- Wolpert, D. and Flanagan, J. (2001). Motor prediction. *Current Biology*, 11:R729–32.
- Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, 1:209–16.
- Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philos Trans R Soc Lond B Biol Sci*, 358(1431):593–602.
- Wolpert, D. M., Gharamani, Z., and Jordan, M. (1995). An internal model for sensorimotor integration. *Science*, 269:1179–1182.
- Wolpert, D. M. and Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 11(7-8):1317–1329.
- Ziemke, T., Jirnhed, D.-A., and Hesslow, G. (2005). Internal simulation of perception: a minimal neuro-robotic model. *Neurocomputing*, 68:85–104.
- Zlatev, J. and Balkenius, C. (2001). Introduction: Why epigenetic robotics? In *Proceedings of the First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, volume 85, pages 1–4. Lund University Cognitive Studies.
- Zlatev, J., Persson, T., and Gardenfors, P. (2005). Bodily mimesis as 'the missing link' in human cognitive evolution. Technical Report LUCS121, Lund University Cognitive Science, Lund, Sweden.