International Journal of Humanoid Robotics
© World Scientific Publishing Company

RobotCub: An Open Framework for Research in Embodied Cognition*

GIULIO SANDINI, GIORGIO METTA, and DAVID VERNON†
DIST, University of Genova, Italy
sandini@dist.unige.it
pasa@dist.unige.it
vernon@ieee.org

We describe a research initiative in embodied cognition that will create and exploit a 54 degree-of-freedom humanoid robot. This humanoid — *RobotCub* — is currently being designed and the final system will be made freely available to the scientific community through an open systems GNU-like general public licence. In addition, we describe a research agenda in cognitive systems that is based on the co-developmental learning through embodied physical interation: exploration, manipulation, imitation, and communication. This agenda borrows heavily from experience in developmental psychology and cognitive neuroscience. All cognitive software associated with *RobotCub* will also be available under the open systems licence.

Keywords: Cognition; Co-development; Open systems; Embodiment; Exploration; Manipulation; Interaction; Imitation.

1. Introduction

This paper describes a new research initiative on the realization of embodied cognitive systems. It has the twin goals of (1) creating an open and freely-available humanoid platform —*RobotCub* — for research in embodied cognition, and (2) advancing our understanding of cognitive systems by exploiting this platform in the study of cognitive development.

To achieve this goal we plan to construct an embodied system able to learn: i) how to interact with the environment by complex manipulation and through gesture production & interpretation; and ii) how to develop its perceptual, motor and communication capabilities for the purpose of performing goal-directed manipulation tasks. *RobotCub* will have a physical size and form similar to that of a two year-old child and will achieve its cognitive capabilities through ontogenic co-development with its environment: by interactive exploration, manipulation, and

---

[*] The content of this paper is based on the work of several people: Paolo Dario, Scuola S. Anna, Pisa; Rolf Pfeifer, University of Zurich; Claes von Hofsten, University of Uppsala; Luciano Fadiga, University of Ferrara; Kerstin Dautenhahn, University of Hertfordshire; Jose Santos-Victor, IST Lisbon; Darwin Caldwell and John Gray, University of Salford; Aude Billard and Auke Ijspeert, Ecole Polytechnique Federal de Lausanne; Francesco Becchi, Telerobot S.r.l.; Emilio Bizzi, European Brain Research Institute. Their contributions are gratefully acknowledged.
[†] David Vernon is on sabbatical leave at Etisalat University, UAE.

imitation. *RobotCub* will be designed as a freely-available open system which can be shared by scientists as a common tool for research in cognitive systems.

To enable the investigation of relevant cognitive aspects of manipulation the design will be aimed at maximizing the number of degrees of freedom (DOF) of the upper part of the body (head, torso, arms, and hands). The lower body (legs) will be designed to support crawling on four legs and sitting on the ground in a stable position with smooth autonomous transition from crawling to sitting. This will allow the robot to explore the environment and to grasp and manipulate objects on the floor. The total number of degrees of freedom for the upper body will be approximately 40 (7 for each arm, 8 for each hand, 7 for the head and 3 for the torso and spine). Each leg will have a further 7 degrees of freedom. The sensory system will include a binocular vision system, touch, audition, and inertial sensors. Functionally, the system will be able to coordinate the movement of the eyes and hands, grasp and manipulate lightweight objects of reasonable size and appearance, crawl on four legs and sit.

2. The *RobotCub* Approach to Cognitive Systems

Our guiding philosophy — and the motivation for creating *RobotCub* — is that cognition cannot be hand-coded but has to be the result of a developmental process through which the system becomes progressively more skilled and acquires the ability to understand events, contexts, and actions, initially dealing with immediate situations and increasingly acquiring a predictive capability.

The *RobotCub* approach to cognition rests on three pillars:

(1) Its scientific stance on cognition: that cognition emerges through embodied co-development.
(2) Its research methodology: that cognition is best studied through a programme of progressive development.
(3) Its research strategy: that progress in the global scientific community is best served by creating an open systems platform and by exploiting consequent synergies in that community.

We will look at each pillar in turn. Before that, we begin by considering the motivation for embodied cognition.

3. Background in Cognitive Systems Research

3.1. *Overview of Approaches to Cognition*

There are several quite distinct approaches to understanding and synthesis of cognitive systems, including physical symbol computation, connectionism, artificial life, dynamical systems, and enactive systems[9,66]. Each approach makes significantly different assumptions about the nature of cognition, its purpose, and the manner in which cognition is achieved. Among these, however, we can discern two broad classes: the *cognitivist* approach based on symbolic information processing representational systems; and the *emergent systems* approach, embracing connectionist systems, dynamical systems, and enactive systems, and based to a lesser or greater extent on principles of self-organization.

3.2. *Cognitivist Models*

Cognitivism asserts that cognition involves computations defined over symbolic representations, in a process whereby information about the world is abstracted by perception, represented using some appropriate symbol set, reasoned about, and then used to plan and act in the world. This approach has also been labelled by many as the *information processing* approach to cognition[21,26,29,34,53,62,66]. Traditionally, this has been the dominant theme in cognitive science[26] but there are indications that the discipline is migrating away from its stronger interpretations[9].

For cognitivist systems, cognition is representational in a strong and particular sense: it entails the manipulation of explicit symbolic representations of the state and behaviour of an objective external world[70]. Reasoning itself is symbolic: a procedural process whereby explicit representations of an objective world are manipulated and possibly translated into language.

In most cognitivist approaches concerned with the creation of artificial cognitive systems, the symbolic representations are the product of a human designer. This is significant because it means that they can be directly accessed and understood or interpreted by humans and that semantic knowledge can be embedded directly into and extracted directly from the system. However, it has been argued that this is also the key limiting factor of cognitivist systems: these designer-dependent representations are the idealized descriptions of a human cognitive entity and, as such, they effectively bias the system (or 'blind' it[70]) and constrain it to an domain of discourse that is dependent on and, a consequence of, the cognitive artifacts of human activity. This approach works well as long as the system doesn't have to stray too far from the conditions under which these descriptions were formulated. The further one does stray, the larger the 'semantic gap'[58] between perception and possible interpretation, a gap that is normally plugged by embedding programmer knowledge or enforcing expectation-driven constraints[47] to render a system practicable in a given space of problems.

This approach usually then goes hand-in-hand with the fundamental assumption that 'the world we perceive is isomorphic with our perceptions of it as a geometric environment'[57]. The goal of cognition, for a cognitivist, is to reason symbolically about these representations in order to effect intelligent, adaptive, anticipatory, goal-directed, behaviour.

3.3. *Emergent Systems*

Emergent systems, embracing connectionist, dynamical, and enactive systems, take a very different view of cognition. Here, cognition is a process of self-organization whereby the system is continually re-constituting itself in real-time to maintain its operational identity through moderation of mutual system-environment interactions and co-determination[37]. Co-determination implies that the cognitive agent is specified by its environment and at the same time that the cognitive process determines what is real or meaningful for the agent. In a sense, co-determination means that the agent constructs its reality (its world) as a result of its operation in that world.

Co-determination is one of the key differences between the emergent paradigm and the cognitivist paradigm. For emergent systems, perception provides appropriate sensory data to enable effective action[37] but it does so as a consequence of the system's actions. In the emergent paradigm, cognition and perception is functionally-dependent on the richness of the action interface[20].

Dynamical systems theory is one of the most promising approaches to the realization of emergent cognitive systems. Advocates of the dynamical systems approach to cognition (*e.g.* [26,62,64]) argue that motoric and perceptual systems, as well as perception-action coordination, are dynamical systems, that self-organize into meta-stable patterns of behaviour.

Proponents of dynamical systems point to the fact that they directly provide many of the characteristics inherent in natural cognitive systems such as multistability, adaptability, pattern formation and recognition, intentionality, and learning. These are achieved purely as a function of dynamical laws and consequent self-organization. They require no recourse to symbolic representations, especially those that are the result of human design.

It has been argued that dynamical systems allow for the development of higher order cognitive functions such as intentionality and learning in a straightforward manner, at least in principle[26]. Although dynamical models can account for several non-trivial behaviours that require the integration of visual stimuli and motoric control, including the perception of affordances, perception of time to contact, and figure-ground bi-stability[18,19,26,31,69], the principled feasibility of higher-order cognitive faculties has yet to be validated.

Enactive systems take the emergent paradigm a little further. In contradistinction to cognitivism, which involves a view of cognition that requires the representation of a given objective pre-determined world[64,66], enaction[35,36,38,37,65,66,70] asserts that cognition is a process whereby the issues that are important for the continued existence of the cognitive entity are brought out or enacted: co-determined by the entity as it interacts with the environment in which it is embedded. Thus, nothing is 'pre-given', and hence there is no need for symbolic representations. Instead there is an enactive interpretation: a real-time context-based choosing of relevance. The advantage is that it focusses on the dynamics by which robust interpretation and adaptability arise.

Theoretical support for the emergent position can be found in recent studies which have shown that an organism can learn the dimensionality and geometry of the space in which it is embedded from an analysis of the dependencies between motoric commands and consequent sensory data, without any knowledge or reference to an external model of the world or the physical structure of the organism[50,51]. The conceptions of space, geometry, and the world that the body distinguishes itself from arises from the sensorimotor interaction of the system, exactly the position advocated in developmental psychology[62].

3.4. *Perception, Cognition, and Time*

It can also be useful to look at cognition from an another perspective; specifically one that distinguishes it from perception in a temporal context. From this perspective, we can define cognition as the 'complementary set' of perception, in which perception only deals with the immediate, and cognition deals with the longer time frame. Thus, cognition reflects the mechanism by which an agent compensates for the immediate nature of perception and can therefore adapt to and anticipate environmental interaction that occurs over much longer time-scales. We will see the relevance of this viewpoint shortly when we come to consider cognitive development.

3.5. *The Necessity of Embodiment in Cognitive Systems*

The cognitivist and the emergent approaches adopt diametrically opposed stances on the issue of embodiment. Cognitivist systems don't necessarily have to be embodied. The very essence of the cognitivist approach is that cognition comprises computational operations defined over symbolic representations and these computational operations are not tied to any given instantiation. They are abstract in principle. It is for this reason that it has been noted that cognitivism exhibits a form of mind-body dualism[61,62]. Symbolic knowledge, framed in the concepts of the designer, can be programmed in directly and doesn't have to be developed by the system itself through exploration of the environment. Some cognitivist systems do exploit learning to augment or even supplant the *a priori* designed-in knowledge and thereby achieve a greater degree of adaptiveness, reconfigurability, and robustness. Embodiment may therefore offer an additional degree of freedom to facilitate this learning, but it is by no means necessary.

On the other hand, emergent systems, by definition, must be embodied and embedded in their environment in a situated historical developmental context[62].

Cognition is the process whereby an autonomous system becomes viable and effective in its environment. In this, there are two complementary things going on: one is the self-organization[a] of the system as distinct entity, and the second is the coupling of that entity with its environment. 'Perception, action, and cognition form a single process'[61] of self-organization *in the specific context of environmental perturbations of the system.* This gives rise to the co-development of the cognitive

---

[a] The self-organization is typically achieved through an operationally-closed network of activities characterized by circular causality[26] and possibly modelled by a dynamical system defined over space of order parameters and control parameters.

system and its environment and thereby to the ontogenic development of the system itself over its lifetime. This development is identically the cognitive process of establishing the space of mutually-consistent couplings. Put simply, the system's actions define its perceptions but subject to the strong constraints of continued dynamic self-organization. The space of perceptual possibilities is predicated not on an objective environment, but on the space of possible actions that the system can engage in whilst still maintaining the consistency of the coupling with the environment. These environmental perturbations don't control the system since they are not components of the system (and, by definition, don't play a part in the self-organization) but they do play a part in the ontogenic development of the system. Through this ontogenic development, the cognitive system develops its own epistemology, *i.e.* its own system-specific knowledge of its world, knowledge that has meaning exactly because it captures the consistency and invariance that emerges from the dynamic self-organization in the face of environmental coupling. Thus, we can see that, from this perspective, cognition is inseparable from 'bodily action'[61]: *without physical embodied exploration, a cognitive system has no basis for development.*

In the last 10 years or so, an ever growing number of cognitive scientists[15,54] have begun to appreciate the possibility of instantiating (admittedly crude) cognitive models in robotic systems. The space of research spanned is quite wide, starting from the locomotion and organizational behaviors of insects and early vertebrates[25,44] through models of high order cognitive skills in humans such as social behaviors,[8] imitation,[1,4,56] communication, and language.[4,6,7,60,67] More recently a new strain of research explicitly included developmental aspects and the modeling of development,[71,33] and epigenetic robotics.[71] Examples are the work of Metta and Sandini,[40,41,42,43] of the group of Pfeifer,[32,48,49] of Dautenhahn *et al.*[11,12]

4. The *RobotCub* Stance on Cognition

The *RobotCub* initiative coincides with the emergent systems approach to cognition. In our view, therefore, cognition must always be embodied: cognition involves development, development requires experience, and experience can only be gathered autonomously by acting. Therefore, the cognitive system must be embodied. Cognition also serves to supply the predictive ability that perception cannot provide. Indeed, the process of cognitive development can be viewed as the gradual acquisition of powerful predictive skills. Consequently, there are two complementary (and strongly co-dependent) components of the *RobotCub* platform. The first is the *RobotCub* humanoid; the second is the software that drives the cognitive process. We will address each in turn. Note first, however, that the co-dependence of the two components is quite critical and we will return to this issue in Section 5.2.

4.1. *The RobotCub Humanoid*

4.1.1. *Overview*

As noted already, the dimensions of the *RobotCub* humanoid will be similar to that of a two-year old child. It will have a head, torso, two arms and hands, and two legs. The legs will be used for crawling, not bipedal walking. Although the design is far from complete and it is still possible to contribute ideas and technologies, we expect that *RobotCub* will have about 54 degrees of freedom organized as follows:

- 7 for each arm
- 8 for each hand
- 7 for the head
- 3 for the torso and spine
- 7 for each leg

The eye-head sub-system will include basic visual processing primitives as well as low-level oculomotor control, visual, inertial and proprioceptive sensors.

RobotCub will have two arms with the motor skills and sensory components required for dextrous manipulation. From the control point of view, reaching and grasping primitives will be implemented as well as primitives to acquire tactile and proprioceptive information. It is expected that most of the actuators of the hand will be located in the forearm. The hands will be underactuated. Underactuation is implemented by means of mechanical coupling either rigidly, such as using a single tendon to bend two joints of a finger alike, or elastically coupling the joints. Overall, the main advantage of using elastic elements lies in that the fingers acquire some passive compliance that helps them in ‘adapting’ to the shape of the grasped object. Underactuation also saves on space, power consumption, and cost.

RobotCub will be able to crawl ‘on four legs’ and sit. This is to allow the system to explore the environment not only by manipulating objects but through locomotion as well. For this reason we consider it particularly important to equip *RobotCub* not only with legs but also with a ‘spine’ to allow bending the torso during transition between sitting and walking posture as well as to look down while manipulating objects lying on the floor.

4.1.2. *Compliance*

Aspects of compliance are also important to the design of the legs of the *RobotCub*. In this respect, traditional robotic systems design has concentrated on developing high speed and accurate mechanisms. This typically results in heavy manipulator arms, with large power requirements presenting relatively low payloads and with limited capacity for human interaction. Advances in computational power and materials have opened the possibility of using lightweight and highly flexible structures, similar to those found in nature to design a new family of robot. This has led to the development of biomimetics and bio-robotic designs where the trend is to try to emulate the 'soft' physically compliant structure of muscle, bone, tendons and skin. This trend is particularly relevant when the robot systems under construction are child-sized and child-like. Key features must be the lightness and resilience to damage of the robot units since the learning (cognition, understanding, and behavior) of the robot will potentially involve the many 'falls' and 'accidents' experienced by any child as it learns to cope with the world.

The system will provide the range of motion and flexibility of a child (which can be much greater than for an adult) with degrees of freedom that permit experimentation and development of the full range of human child motions. The structure will have the functional capacity to emulate the locomotion, motion and motion transition behaviors of a young child. We intend adopting a 'soft' biomimetic approach to the actuation structure. Issues that we see as fundamental are the power/weight and power/volume performance, compliance and stiffness regulation, robustness, control behaviors and more biological issues such as self repair and tolerance of injury and adaptability during the repair process. Issues relating to the control during these repair stages will be important as the overall behavior will be modified during the healing period, and learning and updating control strategies will be essential at this time, as with a child or human. Control of the locomotion of *RobotCub* will require a novel approach particularly when it is remembered that the robot is following a continuous learning and cognitive and behavioral development strategy.

4.1.3. *Hardware Architecture*

The initial version of *RobotCub* will have a tether link to provide both high-level control and power. Low-level control, signal conditioning, and interface functionalities will be carried out by on-board control cards, but for complicated visual and signal processing tasks more computational power is required and will be effected by remote systems. The use of a remote computer is important since different groups may be already developing on different platforms. The idea is to specify the interface and to allow the main processing unit to run on different operating systems and/or processors.

4.1.4. *Software Architecture*

The goal of the software interface is to flatten the user's learning curve as much as possible. The software layer will be open source to allow researchers to modify and customize their specific system. The library will include testing, diagnostic software, and debugging facilities at the lowest level. The intention here is to allow researchers not interested in motor control to start using the robot without going through the burden of re-implementing yet another controller. At the same time, there are many researchers that might really need to tweak the controller at the single joint or the parameters of the multi-joint synchronization, and we will allow for this. This software architecture will comprise three layers: i) DSP code for low-level control, ii) relay station code for signal conditioning and interfacing, and iii) remotely-executed cognitive software. All software will be available as part of the open systems licence. This will allow researchers to work on the aspects that interest them most. For example researchers not interested on studying the exact mechanisms of reaching could rely on the initial open source implementation, while others might re-implement it. Eventually new versions will be available open source (because of the licensing system itself) and might end up substituting the initial implementation.

4.1.5. *The Cost of RobotCub*

The cost of the *RobotCub* will be approximately 50,000 euro. This is divided equally between the cost of materials and fabrication of the electronics, on the one hand, and the cost of machining the mechanical parts, on the other. Assembly costs are not considered in this estimate.

4.2. ***RobotCub Cognition***

Our stance on cognition is based on the paradigm of developmental learning, in which cognitive capabilities evolve as a consequence of the system's exploration, manipulation, and interaction with its surroundings. We believe that one can't achieve cognitive capability any other way except through closed-loop developmental learning achieved through the experience of exploratory physical interaction, ranging from basic hand-eye coordination and reaching actions to imitation and gestural communication. In the following, we will focus on the two key themes of interaction and imitiation.

Note first that development occurs in a very special way. Action, perception, and cognition are tightly coupled in development: not only does action organize perception and cognition, but perception and cognition are also essential for organizing action. Actions systems do not appear ready-made. Neither are they primarily determined by experience. They are the result of a process with two foci, one in the central nervous system and one in the subject's dynamic interactions with the environment. Perception, cognition, and motivations develop at the interface between brain processes and actions. Biology has prepared the infant for action by investing in certain primitives and making those unfold in specific ways that optimize the developmental process. Thus, development takes place in the context of a circular causality of action and perception[26], each a function of the other as the system manages its mutual interaction with the world. We refer to this as co-development.

4.2.1. *Interaction*

We adopt the definition in Ogden *et al.*[46] of interaction as a shared activity in which the actions of each agent influence the actions of the other agents engaged in the same interaction, resulting in a mutually constructed pattern of shared behavior.

This definition is consistent with the emergent cognition paradigm discussed above, especially the co-constructed nature of the interaction, inspired by concepts of autopoiesis and structural coupling.[38] This aspect of mutually constructed patterns of complementary behavior is also emphasized in Clark's notion of joint action.[10] According to this definition explicit meaning is not necessary for anything to be communicated in an interaction, it is simply important that the agents are mutually engaged in a sequence of actions. Meaning emerges through shared consensual experience mediated by interation. The *RobotCub* research programme is based on this foundational principle of interaction. As we will see later, we intend to begin with basic action-perception skills and build on them to develop more sophisticated and complex cognitive behaviour.

4.2.2. *Imitation*

The ability to imitate has for long been used as a means to measure the infant's stages of development.[52] Piaget's theories have by now been largely revisited and the age markers of the imitation stages redefined.[39] Imitation in children goes from being immediate to being deferred, when the replication occurs within a short (few minutes) or long (hours, days) time after the demonstration. It progresses from being partial or selective (when only part of the imitative behavior is replicated), goal-directed (when only the means-end of the demonstration is perfectly reproduced) to being exact.[3] The developmental progress of imitation follows tightly that of the development of other interactive and communicative skills, such as joint attention, turn taking and language.[45,59,63] Imitation is at the basis of many social interactions and children who are impaired in their imitative skills (*e.g.*, children with autism) also show general impairments in other social skills.[13] While the study of the ability of infants and adults to imitate has remained foremost a field of the psychological literature, recently, it has found a ground in the neurological literature with the discovery of the mirror neuron system in monkeys.[55] The mirror neuron system is formed by premotor neurons discharging both when the animal acts and when it sees similar actions performed by other individuals. A system, similar to that found in monkeys, has been indirectly shown to exist also in humans by transcranial magnetic stimulation studies of the motor cortex during action observation.[17] Further investigations have shown that the mirror system can be activated not only by visually perceived actions but also by listening to action-related sounds[30] and, in humans, by listening to speech.[16] In addition to these electrophysiological data, in humans, a number of brain imaging studies all point to a network of brain areas responsible for the visuo-motor transformation mechanism underlying action recognition.[14,24] It is plausible that the motor resonant system formed by mirror neurons is involved in understanding someone else's action and, at least in humans, imitation. The principle of motor resonance will be extensively taken into account during the design of the robot's mindware.

The developmental psychology and neurological literature on imitation has led the way to novel computational and robotics models of imitation.[5,41]

Imitation is one of the key stages in the development of more advanced cognitive capabilities. The *RobotCub* platform will be used to further these studies and create a communicative cognitive system.

5. The *RobotCub* Research Methodology — A Programme of Progressive Development

5.1. *The Importance of Development*

The issue of development is crucial to *RobotCub* for two quite distinct but related reasons. First, it is, in our view, the manner in which cognition itself emerges: it is the ontogenic development of a cognitive system as it learns to make sense of its world through exploration, manipulation, imitation, and gestural communication. As noted above, this development takes place in the context of a circular causality of action and perception, each a function of the other as the system manages its mutual interaction with the world. Again, we refer to this as co-development.

However, development is also crucial in a more pragmatic sense in that it provides us with a research strategy. Since development is a temporally-extended event, it allows us to study cognition in an incremental way without having to understand the complete scheme *ab initio*. It gives us a way to choose an early point of departure in the endeavour to understand cognition and then to make progress as the system itself develops. The legitimacy of this methodology is supported by experience in studying newborn infants.[23,2,68]

There are essentially two processes of perceptual development. The first one is a spontaneous perceptual learning process that has to do with the detection of structure in the sensory flow. As long as there is variability and change in the sensory flow, the perceptual system will spontaneously learn to detect structure and differentiate invariants that correspond to relatively stable and predictable properties of the world. The second process is one of selecting information relevant for guiding action. This is a question of learning about affordances and in this process perception and action are mutually dependent on each other. However, before infants can learn about affordances they must already have detected that structure in the sensory flow.

The study of human development demonstrates that the presence of action clearly emerges as a gluing principle in understanding cognition. Cognition emerges in humans because of the requirement of controlling actions, from reaching movements, to grasping or more complex social interactions, and speech. Our methodology is therefore to explain this incremental emergent cognitive behaviour and to create a cognitive agent through its development over time and across visuo-motor skill sets.

Our goal is to model and synthesize cognitive (adaptive, anticipatory, interactive, goal-achieving, and social) capabilities, emulating how humans learn to use their hands and arms not just for manipulation but also to convey information, to express their emotional status, to interact socially. The intention is to start from 'behavioural-driven' research and to move beyond it to address, for example, the role of manipulation as a source of knowledge and new experience, as a way to communicate socially, as a tool to teach and learn, as a means to explore and control the environment.

5.2. *Co-Development and Co-Design*

One of the principal guiding philosophies of the *RobotCub* approach is that the software that drives the cognitive capabilities of the system must be developed synergistically with the development of the physical humanoid platform, and *vice versa*; the two aspects are complementary but inseparable. This follows directly from our working hypothesis that cognition emerges from the co-development of the system with its environment. A consequence of co-development therefore is the necessity for the complete system to be co-designed. For example, a system that has to learn how to build a visuo-haptic representation of an object requires dexterous hands providing postural as well as haptic information with relatively high accuracy. Without such performance, the resultant cognitive impact may be negligible: rich cognition requires a rich space of interaction. The co-design philosophy also applies in the configuration of the physical system. Our view is that the complexity of the complete system can best be handled not by dealing with each part in isolation but by designing the system as a whole and thereby exploiting the mutual constraints that one design decision imposes on others. For example designing the arm and the hand together will allow us to choose the best positioning of the actuators as well as the most appropriate transmission mechanism for each degree of freedom (gears, tendons, belts). The same would apply if jointly designing the head-shoulder and torso (since they share the same physical space).

5.3. *The Research Agenda*

As noted already, the research methodology is to exploit the developmental nature of embodied cognition and to conduct a programme of experimental investigation at many different levels, beginning with the immediate time scale (*e.g.* motor control, sensory mapping, *etc.*), aspects of prospective control (*e.g.* reaching/grasping moving objects, tracking/eye movement, anticipation), followed by the more elaborate predictions required for manipulating objects (*e.g.* grasping according to shape and use), and finally, towards skills requiring deliberation and prediction such as communication, imitation, and complex manipulation involving tools.

Thus, the research programme is centred on the study of developmental issues, borrowing from the development of cognitive faculties in humans, and using empirical research to guide the creation of an artificial system.

The experimental scenarios we plan on adopting include the following.

(1) Discovering the manipulation abilities of its own body:

  - Learning to control one's upper and lower body (crawling, bending the torso) to reach for targets.
  - Learning to reach static targets.
  - Learning to reach moving targets.
  - Learning to balance in order to perform stable object manipulations when crawling or sitting.

(2) Discovering and representing the shape of objects:

  - Learning to recognize and track visually static and moving targets.
  - Discovering and representing object affordances (e.g. the use of tools).

(3) Recognizing manipulation abilities of others and relating those to one's own manipulation abilities:

  - Learning to interpret and predict the gestures of others.
  - Learning new motor skills and new object affordances by imitating manipulation tasks performed by others.
  - Learning what to imitate and when to imitate others' gestures.

(4) Learning regulating interaction dynamics:

  - Approach, avoidance, turn-taking, and social spaces.
  - Learning to use gesture as a means of communication.

(5) Developing robot 'personalities' *via* autobiographic memory based on interaction histories:

  - Learning about meaningful events in the lifetime of the robot.
  - Sharing memory (events) during interaction.

These experimental scenarios will be used in a coordinated programme of experimental research of five skill-based categories of behaviour, each of which exhibits a greater degree of cognition. These five categories are set out in the following sections.

5.3.1. *Eye-head-hand Coordination*

We plan to study and implement the sensorimotor primitives at the basis of eye-head-hand coordination. The need to include this level stems from the fact that the implementation of cognitive skills, particularly in an embodied system, has to take into consideration and possibly exploit the peculiarities of the body and this has natural implications at the level of sensorimotor coordination. For example, a control system able to adapt to lengthening and shortening of the muscles, growing and aging of sensors and actuators, *etc.*, forms the basis of the adaptability that is a crucial aspect of cognition.

5.3.2. *Bimanual Cooperation*

This strand of research will be devoted to the study of how two hands can cooperate to achieve higher-level manipulation skills. By this we mean for example:

(1) The ability to use the two arms-hands as a scaled up version of a single hand. This will enable the system to do things that cannot be done with one hand (such as grasping larger objects).
(2) The ability to use the two hands in a cooperative and similar way. By this we mean, for example, grasping two extremes of a rope or lifting and extending soft materials.
(3) The ability to use two hands in a cooperative but dissimilar way. For example holding a glass with one hand and pouring water from a bottle with the other.

From some aspects, the use of two hands is not qualitatively different from the use of a single hand. For example, finding a stable grasp with two fingers requires the same skills needed to grasp a large object with two hands. However the role of two-arm manipulation in the process of developing manipulation skills is very significant in humans and may represent a necessary step to learn how to approach a graspable object and how to properly pre-shape the hand. The symmetries in controlling the arms observed in young infants may in fact facilitate learning by filtering out some of the noise present in motor commands during early stages of the developmental process (infants have difficulties in performing independent movement of the arms and tend to perform symmetrical movements). Therefore, besides the increased manipulation skills allowed by two-arm manipulation and the consequent increase in system's complexity, we will investigate the role of two-arm coordination in learning manipulation skills.

5.3.3. *Interaction and Affordance*

The goal of this research theme is to study how, by interacting with the environment, it is possible to discover/learn the use of objects. By this we mean being able to manipulate objects according to use as well as shape. An example is the grasping of a rod either with a palm grasp when it is used like a hammer or with a pinch grasp when it is used like a pen. These aspects cannot be separated from those described in the previous section but they address more abstract properties of objects, *i.e.* properties that cannot be derived by contingent sensory perception alone. At the same time, the acquisition of these abilities cannot be separated by aspects addressed in other research strands in the sense that the 'affordant' use of objects can be learned by self-experimentation as well as through imitation and social interaction with other animate agents (see next two sections). For instance, various experimental observations suggest that object knowledge (and therefore an object's affordances) grows in parallel with the development of motor capabilities. Thus, actions at the same time are used to explore objects and to describe and categorize them. In other words, an object is not an entity pertaining only to the external world, but can be described in terms of physical properties and motor interactions.

5.3.4. *Interaction and Imitation*

Within the *RobotCub* initiative we view imitiation as one of the most crucial issues. Our idea of imitation is, in fact, intimately linked to understanding. We do not see imitation only as a kinematic playback aimed at reproducing the temporal evolution of an action but we plan to give *RobotCub* the ability to imitate in order to comprehend two essential cognitive abilities: i) learning how to act upon objects (with all the implications due to the need of a unified sensory/motor/affordant representation of tools); ii) the ability to communicate through imitation. The displayed action could be kinematically similar to the one being imitated (imitating the action) but the level of understanding reached should eventually allow the system to 'imagine' new behaviors producing the same effect (imitating the effect). Data coming from neurophysiological investigation of primates' mirror neurons, as well as from brain imaging and electrophysiological studies in humans, will be used in defining the framework for imitation and in guiding the implementation.

5.3.5. *Interaction and communication*

Interaction and communication subsumes all previous activities in the sense of requiring the system to learn how to use knowledge to communicate through gestures and to interact socially with other agents. To some extent this is the core problem of 'cognitive interface'. What we want to stress is that these skills require all the levels of analysis and implementation described in all other cognitive skills. In order to be able to interact socially, the system has to understand how to exploit the physical structure of the body, its manipulation skills, the affordances of objects, the difference between animate agents and inanimate objects, how to imitate and how to understand when it is being imitated, the self from the others. Specific techniques, such Conversation Analysis,[22,27,28] are being considered for analyzing and designing the structure of the interaction. In particular, this strand of research should eventually design the cognitive software structure that allows the robot to engage in meaningful communication with humans in order to gather the structure of tasks/games.

6. Open Systems Development

The *RobotCub* platform, including both the humanoid robot and all embedded cognitive software, will be a freely-available open system. This openness will be facilitated making available the design of the *RobotCub* humanoid and all embedded software (controller software, interface software, and cognition software) to anybody interested in studying and developing cognitive behaviours on terms similar to the GNU General Public Licence. Thus, the scientific community can use it, copy it, and alter it, provided that all alterations to the humanoid design and the embedded software are also made available under the *RobotCub* open licence. We are actively encouraging the community to use *RobotCub* in this manner. Our intention is to build a strong international base of users, beginning with the core group of European Union partners and collaborators in the USA and Japan, and adding more collaborators as the initiative advances.

To help ensure that the research community can actually exploit the openness of *RobotCub*, parts will be competitively priced and, insofar as it practicable, standard off-the-shelf components will be used. The use of special purpose electronics will be minimized (but will be included in the GPL, in any case). The mechanical structural parts can be manufactured by any engineering workshop using the provided documentation (CNC files or mechanical drawings).

RobotCub will be easy to maintain in a standard robotic laboratory or, given proper training, in any laboratory. For example, psychology laboratories wishing to do research into cognitive robotics or robot-human interaction should be able to do so without difficulty.

Additional research projects will also be launched to help enlarge the user base and enable a second wave of research on embodied cognition using the *RobotCub* platform. Several copies of *RobotCub* will be constructed and distributed as part of this second wave. This initiative will also provide feedback on the platform, thereby bootstrapping the first cycle of continuous improvement of *RobotCub*.

In the near future, we will launch a research and training site (RTS) where several fully functional copies of the *RobotCub* will be maintained. The RTS will be the reference site for the open system and will be responsible for the preparation of suitable server computers, www access, updates, and management of releases. The RTS will also act as a gathering place where new software or upgrades will be attached to the system and tested before being released. It will also provide training facilities. Since the target users of *RobotCub* will be drawn from several disciplines, many of which won't necessarily be skilled in robotics, some training may be required. The RTS will fulfil this need, providing fully-trained personnel to assist in the preparation, utilization, and development of new components for the *RobotCub*, including the assembly of complete systems. Finally, the RTS will also act as an open research centre where any researcher with a sound research programme can come to do research on embodied cognition without having to commit to the purchase of a dedicated system.

7. Concluding Note

Finally, we wish to emphasize again that the principal motivation for this initiative is to help foster the study of embodied cognition throughout the global research community by making the *RobotCub* humanoid and cognitive software freely available. Representatives of this international community have been involved with *RobotCub* from the outset. Our goal is to increase this involvement as much as possible over the coming years and we welcome potential collaborators.

Acknowledgements

This work is funded by the European Commission's newly-established Cognition Unit, Directorate-General Information Society, as part of project no. IST-2004-004370: *RobotCub* — ROBotic Open-architecture Technology for Cognition, Understanding, and Behaviour.

References

1. M. Arbib, A. Billard, M. Iacoboni, and E. Oztop. Mirror neurons, imitation and (synthetic) brain imaging. *Neural Networks*, 13(8/9):953–973, 2000.
2. J. Atkinson. *The developing visual brain.* Oxford University press, Oxford, UK, 2000.
3. H. Bekkering and W. Prinz. Goal representations in imitative actions. In K. Dautenhahn and C. Nehaniv, editors, *Imitation in Animals and Artifacts.* MIT Press, Cambridge, MA, 2002.
4. A. Billard. Imitation. In M.A. Arbib, editor, *The Handbook of Brain Theory and Neural Networks.* MIT Press, Cambridge, MA, 2002.
5. A. Billard. Imitation: a means to enhance learning of a synthetic proto-language in an autonomous robot. In K. Dautenhahn and C.L. Nehaniv, editors, *Imitation in Animals and Artifacts*, pages 281–311. MIT Press, Cambridge, MA, 2002.
6. A. Billard and K. Dautenhahn. Grounding communication in autonomous robots: an experimental study. *Robotics and Autonomous Systems*, 24(1/2 "Scientific methods in mobile robotics"):71–79, 1998.
7. A. Billard and K. Dautenhahn. Experiments in social robotics: grounding and use of communication in autonomous agents. *Adaptive Behavior*, 7(3/4):415–438, 2000.
8. C. Breazeal. *Sociable Machines: Expressive Social Exchange Between Humans and Robots.* Unpublished Doctoral Dissertation. MIT, Cambridge, MA., 2000.
9. A. Clark. *Mindware – An Introduction to the Philosophy of Cognitive Science.* Oxford University Press, New York, 2001.
10. H. H. Clark. Managing problems in speaking. *Speech Communication*, 15:243–250, 1994.
11. K. Dautenhahn and S. Coles. Narrative intelligence from the bottom up: A computational framework for the study of story-telling in autonomous agents. *JASSS, The Journal of Artificial Societies and Social Simulation*, 2000.
12. K. Dautenhahn, B. Ogden, and T. Quick. From embodied to socially embedded agents - implications for interaction-aware robots. *Cognitive Systems Research*, 3(3):397–428, 2002.
13. G. Dawson and A. Adams. Imitation and social responsiveness in autistic children. *Abnormal Child Psychology*, 12:209–225, 1984.
14. J. Decety, T. Chaminade, J. Grezes, and A.N. Meltzoff. A pet exploration of the neural mechanisms involved in reciprocal imitation. *NeuroImage*, 15:265–272, 2002.
15. G. M. Edelman. *Neural Darwinism: The Theory of Neuronal Group Selection.* Oxford University Press, Oxford, 1988.
16. L. Fadiga, L. Craighero, G. Buccino, and G. Rizzolatti. Speech listening specifically modulates the excitability of tongue muscles: a tms study. *European Journal of Neuroscience*, 15:399, 2002.
17. L. Fadiga, L. Fogassi, G. Pavesi, and G. Rizzolatti. Motor facilitation during action

observation: a magnetic stimulation study. *Journal of Neurophysiology*, 73(6):2608–2611, 1995.

18. J. J. Gibson. *The Perception of the Visual World*. Houghton Mifflin, Boston, 1950.
19. J. J. Gibson. *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, 1979.
20. G. H. Granlund. The complexity of vision. *Signal Processing*, 74:101–126, 1999.
21. J. Haugland. Semantic engines: An introduction to mind design. In J. Haugland, editor, *Mind Design: Philosophy, Psychology, Artificial Intelligence*, Cambridge, Massachusetts, 1982. Bradford Books, MIT Press.
22. I. Hutchby and R. Wooffitt. *Conversantion Analysis*. Polity Press, Cambridge, UK, 1998.
23. P.R. Huttenlocher. Morphometric study of human cerebral cortex development. *Neuropsychologia*, 28:517–527, 1990.
24. M. Iacoboni, R.P. Woods, M. Brass, H. Bekkering, J.C. Mazziotta, and G. Rizzolatti. Cortical mechanisms of human imitation. *Science*, 286:2526–2528, 1999.
25. A. J. Ijspeert. Synthetic approaches to neurobiology: review and case study in the control of anguiliform locomotion. In *Fifth European Conference on Artificial Life (ECAL99)*, 1999.
26. J. A. S. Kelso. *Dynamic Patterns – The Self-Organization of Brain and Behaviour*. M.I.T. Press, 3rd edition, 1995.
27. A. Kendon. Features of the structural analysis of human communicational behavior. In W. von Raffler-Engel, editor, *Aspects of Nonverbal Communication*. Swets and Zeitlinger, Lisse, Holland, 1980.
28. A. Kendon. *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge University Press, Cambridge, UK, 1990.
29. J. F. Kihlstrom. The cognitive unconscious. *Science*, 237:1445–1452, September 1987.
30. E. Kohler, C. Keysers, M. A. Umilta, L. Fogassi, V. Gallese, and G. Rizzolatti. Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297:846–848, 2002.
31. W. Köhler. *Dynamics in Psychology*. Liveright, New York, 1940.
32. M. Lungarella and L. Berthouze. Adaptivity through physical immaturity. In *2nd International Workshop on Epigenetic Robotics (EPIROB'02)*, pages 79–86, Edinburgh, Scotland, 2002.
33. M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini. Developmental robotics: A survey. *Connection Science*, Forthcoming, 2003.
34. D. Marr. Artificial intelligence – a personal view. *Artificial Intelligence*, 9:37–48, 1977.
35. H. Maturana. Biology of cognition. Research Report BCL 9.0, University of Illinois, Urbana, Illinois, 1970.
36. H. Maturana. The organization of the living: a theory of the living organization. *Int. Journal of Man-Machine Studies*, 7:313, 1975.
37. H. Maturana and F. Varela. *The Tree of Knowledge – The Biological Roots of Human Understanding*. New Science Library, Boston & London, 1987.
38. H. R. Maturana and F. J. Varela. *Autopoiesis and Cognition — The Realization of the Living*. Boston Studies on the Philosophy of Science. D. Reidel Publishing Company, Dordrecht, Holland, 1980.
39. A.N. Meltzoff. The human infant as imitative generalist: a 20-year progress report on infant imitation with implications for comparative psychology. In C.M. Heyes and B.G. Galef, editors, *Social Learning in Animals: The Roots of Culture*. Academic Press, New York, 1990.
40. G. Metta. *Babybot: a Study on Sensori-motor Development*. Ph.D. Thesis, University

of Genova, 2000.

41. G. Metta and P. Fitzpatrick. Early integration of vision and manipulation. *Adaptive Behavior*, 11(2):109–128, 2003.
42. G. Metta, G. Sandini, and J. Konczak. A developmental approach to visually-guided reaching in artificial systems. *Neural Networks*, 12(10):1413–1427, 1999.
43. G. Metta, G. Sandini, L. Natale, and F. Panerai. Development and robotics. In *IEEE-RAS International Conference on Humanoid Robots*, Tokyo, Japan, 2001.
44. R. Moller, M. Marinus, and D. Lambrinos. A neural model of landmark navigation in insects. In *Computational Neuroscience Meeting CNS'98*, 1998.
45. J. Nadel, C. Guerini, A. Peze, and C. Rivet. The evolving nature of imitation as a format for communication. In J. Nadel and G. Butterworth, editors, *Imitation in Infancy*, pages 209–234. Cambridge University Press, Cambridge, 1999.
46. B. Ogden, K. Dautenhahn, and P. Stribling. Interactional structure applied to the identification and generation of visual interactive behaviour: Robots that (usually) follow the rules. In I. Wachsmuth and T. Sowa, editors, *Gesture and Sign Languages in Human-Computer Interaction*, volume LNAI 2298 of *Lecture Notes LNAI*, pages 254–268. Springer, 2002.
47. J. Pauli and G. Sommer. Perceptual organization with image formation compatibilities. *Pattern Recognition Letters*, 23:803–817, 2002.
48. R. Pfeifer and C. Scheier. Sensory-motor coordination: The metaphor and beyond. *Robotics and Autonomous Systems*, 20:157–178, 1997.
49. R. Pfeifer and C. Scheier. Representation in natural and artificial agents: an embodied cognitive science perspective. In *Natural Organisms, Artificial Organisms, and Their Brains*, pages 480–503, Bielefeld, Germany, 1998. Springer Verlag.
50. D. Philipona, J. K. O'Regan, and J.-P. Nadal. Is there something out there? inferring space from sensorimotor dependencies. *Neural Computation*, 15(9), 2003.
51. D. Philipona, J.K. O'Regan, J.-P. Nadal, and Olivier Coenen. Perception of the structure of the physical world using unknown multimodal sensors and effectors. In Sebastian Thrun, Lawrence Saul, and Bernhard Schölkopf, editors, *Advances in Neural Information Processing Systems 16.* MIT Press, Cambridge, MA, 2004.
52. J. Piaget. *Play, Dreams and Imitation in Childhood.* Norton, NewYork, 1962.
53. S. Pinker. Visual cognition: An introduction. *Cognition*, 18:1–63, 1984.
54. M. I. Posner, editor. *Foundations of Cognitive Science.* The MIT Press, Cambridge, MA, fifth edition, 1996.
55. G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3:131–141, 1996.
56. S. Schaal. Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences*, 3(6):233–242, 1999.
57. R. N. Shepard and S. Hurwitz. Upward direction, mental rotation, and discrimination of left and right turns in maps. *Cognition*, 18:161–193, 1984.
58. A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, December 2000.
59. G. S. Speidel. Imitation: a bootstrap for learning to speak. In G.E. Speidel and K.E. Nelson, editors, *The many faces of imitation in language learning*, pages 151–180. Springer Verlag, 1989.
60. L. Steels and P. Vogt. Grounding adaptive language games in robotic agents. In P. Husbands and I. Harvey, editors, *Fourth European Conference on Artificial Life, ECAL97*, pages 473–484. MIT Press, 1997.
61. E. Thelen. Time-scale dynamics and the development of embodied cognition. In R. F.

Port and T. van Gelder, editors, *Mind as Motion – Explorations in the Dynamics of Cognition*, pages 69–100, Cambridge, Massachusetts, 1995. Bradford Books, MIT Press.

62. E. Thelen and L. B. Smith. *A Dynamic Systems Approach to the Development of Cognition and Action.* MIT Press / Bradford Books Series in Cognitive Psychology. MIT Press, Cambridge, Massachusetts, 1994.
63. C. Trevarthen, T. Kokkinaki, and G. A. Fiamenghi Jr. What infants' imitations communicate: with mothers, with fathers and with peers. In J. Nadel and G. Butterworth, editors, *Imitation in Infancy*, pages 61–124. Cambridge University Press, Cambridge, 1999.
64. T. van Gelder and R. F. Port. It's about time: An overview of the dynamical approach to cognition. In R. F. Port and T. van Gelder, editors, *Mind as Motion – Explorations in the Dynamics of Cognition*, Cambridge, Massachusetts, 1995. Bradford Books, MIT Press.
65. F. Varela. *Principles of Biological Autonomy.* Elsevier North Holland, New York, 1979.
66. F. J. Varela. Whence perceptual meaning? a cartography of current ideas. In F. J. Varela and J.-P. Dupuy, editors, *Understanding Origins – Contemporary Views on the Origin of Life, Mind and Society*, volume 130 of *Boston Studies in the Philosophy of Science*, pages 235–263. Kluwer Academic Publishers, 1992.
67. P. Varshavskaya. Behavior-based early language development on a humanoid robot. In *Second International Workshop on Epigenetic Robotics*, number 94 in Cognitive Studies, pages 149–158. Lund University, 2002.
68. C. von Hofsten and K. Rosander. Development of smooth pursuit tracking in young infants. *Vision Research*, 37:1799–1810, 1997.
69. W. H. Warren. Perceiving affordances: Visual guidance of stairclimbing. *Journal of Experimental Psychology: Human Perception and Performance*, 10:683–703, 1984.
70. T. Winograd and F. Flores. *Understanding Computers and Cognition – A New Foundation for Design.* Addison-Wesley Publishing Company, Inc.,, Reading, Massachusetts, 1986.
71. J. Zlatev and C. Balkenius. Introduction: Why "epigenetic robotics"? In *First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, volume 85, pages 1–4. Lund University Cognitive Studies, 2001.