DARWIN 270138
Dexterous Assembler Robot
Working with Embodied Intelligence

SEVENTH FRAMEWORK PROGRAMME
ICT Priority

**Deliverable D4.1: “A roadmap for conducting explorative learning in cognitive robots”**

Due date of deliverable: 31.01.2012
Actual submission date: 06.02.2012

Start date of project: 1st February 2011  Duration: 48 months

Coordinator:
Dr. Christian Eitzinger
Profactor GmbH

Lead Beneficiary: IIT
Contributions by: CVUT, PROF, NOCA, KCL, FORTH

<table>
<thead>
<tr>
<th>Dissemination Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PU</td>
<td>Public</td>
</tr>
<tr>
<td>PP</td>
<td>Restricted to other programme participants (including the Commission Services)</td>
</tr>
<tr>
<td>RE</td>
<td>Restricted to a group specified by the consortium (including the Commission Services)</td>
</tr>
<tr>
<td>CO</td>
<td>Confidential, only for members of the consortium (including the Commission Services)</td>
</tr>
</tbody>
</table>
Contents

1. Introduction and Purpose of this document ................................................................. 3

2. Darwin at Home vs. Darwin at Industry: requirements and areas of convergence ...... 3

2.1 Functional requirements and constraints: Darwin at Industry .................................. 4

2.2 Functional requirements and constraints: Darwin at Home ..................................... 6

3 A roadmap of Environmental scenarios to aid cumulative learning in Darwin systems ..... 9

3.1 The 'Identify-localize-plan-reach-grasp-transport' loop .......................................... 9

3.1.1 Service robot (iCub) ......................................................................................... 10

3.1.2 Application domain/Industrial scenario ......................................................... 12

3.2 The 'Perceptual Decomposition – Compositional Action' loop ............................... 12

3.2.1 Service robot (iCub) ......................................................................................... 12

3.2.2 Application domain/Industrial scenario ......................................................... 14

3.3 The ‘Goal-Affordance-Skill-Value’ loop ................................................................. 15

3.3.1 Service robot (iCub) ......................................................................................... 15

3.3.2 Application domain/Industrial scenario ........................................................... 18

3.4 The ‘Disassembly-Reparation-Assembly’ loop ....................................................... 18

3.4.1 Service robot (iCub) ......................................................................................... 18

3.4.2 Application domain/Industrial scenario ......................................................... 19

4 Towards a practical perspective on ‘Learning’ ............................................................ 20

4.1 Assumptions and Muddiness ............................................................................... 20

4.2 Guiding “constraints” and “challenges” for learning in Darwin ............................. 23

4.2.1 Learning Green: Abstraction and Compositionality ........................................ 23

4.2.2 Body schema and the territory of Action without ‘movement’ ......................... 24

4.2.3 Developmental programs vs. Learning/knowing everything .......................... 24

4.2.4 Physical Embodiment vs. Multiple embodiments .......................................... 25

4.2.5 Embodied robot vs. Socially embedded robot ................................................. 25

4.2.6 Meta exploration vs. Meta exploitation .......................................................... 26

4.2.7 Maturational Constraints, Incorporation of tools .......................................... 26

4.2.8 Memory and the generative use of experience ................................................. 27

4.2.9 Sense vs. ‘Common sense’ ................................................................................ 27

5 Learning in Darwin: Ongoing progress and future challenges ................................. 28

5.1 Drives and Motivations ....................................................................................... 28

5.2 Learning in Vision .............................................................................................. 29

5.3 Learning in Action .............................................................................................. 32

5.3.1 Body schema for embodied robots: The Passive Motion Paradigm (PMP) ....... 32

5.3.2 From PMP to Motor Skill learning, Imitation, tool use and Motor knowledge recycling... 38

5.4 Learning Affordances ........................................................................................ 47

5.5 Acquisition of ‘Spatial’ and ‘Causal’ relations ..................................................... 49

5.6 On the “Coevolution” Learning and Reasoning in Darwin, abstraction levels ........ 54

6. Conclusions: A roadmap for future, a pictorial summary ......................................... 57

References .................................................................................................................. 59
1. Introduction and Purpose of this document

The most striking feature of any cognitive system is its ability to learn. Every human child plays, interacts with objects (and people) around it and ‘continuously’ learns from its actions and the resulting consequences of its actions. From learning to control our own bodies and knowing what we can do with different objects around us to understanding causal relations and simple facts of the way the world works, play gradually helps extending our cognitive horizons and prepares us to act with ‘purpose’ and ‘reason’ in a changing, dynamic world. What is the computational/neural substrate that can endow Darwin robots with such capabilities? Broadly this is the vision of DARWIN: to develop an ‘acting, learning and reasoning’ assembler robot that is capable of functioning with some degree of dextrousness, flexibility and autonomy in both ‘industrial’ and ‘domestic’ environments. This is a difficult problem – but there are many pressures to provide a solution both from the intrinsic viewpoint of better understanding ourselves to creating artificial agents, smart devices and assistive robots that can deal autonomously with our needs and with the peculiarities of the environments we inhabit and construct.

In order to support ‘active’ learning, the creation of a robust ‘perception- action’ loop through basic integration of various sensorimotor contingencies (vision, touch, proprioception, motion, force) and their associated control schemas assumed central focus during the first year. At the same time, it was necessary to keep an eye on how the integrated sensorimotor machinery supports both short term and long term objectives of Darwin in terms of facilitating open ended learning and purposeful action generation in environments not completely in the robots ‘explicit’ control. Hence arises the need to formulate an early sketch or roadmap on how explorative learning is planned to progress in Darwin, what are the environments, what are the assumptions and challenges, what are the learning techniques in different areas like perception, action, affordances, causal relations etc and how they come together to create the Darwin cognitive control layer. The purpose of this deliverable is to offer a first sketch on these crucial topics and at the same time present an initial working hypothesis as to how we expect the architecture "to come to life" as the project progresses. We also note that this document should not be read as the final report on the DARWIN architecture but rather as a ‘fluid’ collection of our hypothesis of the necessary core components and their interactions which will suffice to implement the vision of the DARWIN project.

The rest of the draft is organized as follows. Section 2 presents a short summary of the key challenges and questions faced by the project. Section 3 presents a roadmap of ‘environmental scenarios’ of gradually increasing complexity to aid cumulative learning in Darwin, synchronized with their relation to the emerging functional capabilities of the Darwin controller. Section 4 describes the learning techniques being employed at perceptual, motor, social and cognitive levels in the Darwin architecture. The concluding section integrates ‘environmental scenarios’, ‘scientific challenges’ and ‘the resulting Darwin functionality’ in a temporal context to formulate a roadmap for future developments.

2. Darwin at Home vs. Darwin at Industry: requirements and areas of convergence

Robotics has undergone a major transformation in the last decades. In addition to the conventional application of robots in the industry to improve factory efficiency and augment human
**performance** in numerous tasks that may be well beyond our physical capabilities, there has been an emerging wave of interest in making robotic systems an integral part of *day to day human activities* ranging from housekeeping, work-place automation, education and entertainment, aid for the elderly and physically challenged, to mention a few. Such *extrinsic* utility of robots are at the same time being *supplemented and contrasted* by an *intrinsic* utility value based on the idea that robots can also be used as a *tool* to better understand our own selves, to understand how interactions between body and world shapes the mind, shapes reason and shapes action. This stems from the fact that under such conditions a robot acts as a modern day *microscope* that permits the re-enactment of the gradual process of acquisition of cognitive skills and their integration into an interacting cognitive system. The *Darwin project uniquely lies at an intersection of all these crucial enterprises: industrial assembly, domestic assistance and enhancing our understanding of our own selves*. The industrial sector is by far market-driven. The majority of the robots deployed worldwide are of this type. Typical applications are in material handling, assembly, and other labor-replacing scenarios where performing at a greater *efficiency than the ‘human* in terms of power, speed and repetitive accuracy determines the ‘value’. On the other hand, in personal and service robotics, operating around humans and in environments we inhabit with *specific emphasis on friendly and cooperative interfaces*, runtime *adaptability to human needs determines the ‘value’. In cognitive robotics research conducted in laboratories, the ability of a *robot to learn ‘cumulatively’ like an infant by merging its diverse sensory and motor experiences, its curiosity to grow in experience, its ability to interpret and anticipate the ways in which the world works and act with reason determines ‘value’.*

In this sense, a distinctive feature of Darwin project is that it aims to develop a *domain agnostic* cognitive architecture for robots of varying physical structure and sensorimotor capabilities and expected to function either in ‘domestic’ or ‘industrial’ environments. The Darwin project involves two different robots, one that is expected to perform in a purely industrial context and other i.e. the iCub humanoid that enters into subtle aspects of both understanding human cognition and application domains in relation to the domestic and service sectors. Most of the techniques for learning, perceptual, motor and cognitive development undergoing basic research and experimentation in the service robot platform (iCub) are also planned to be appropriately “recycled or transferred” in short term to the industrial robot (operating under environments of similar complexity) additionally complying to industrial requirements. Hence before formulating the various stages of “development” and “learning” in Darwin and the environments in which such developments occur there is a need to understand the constraints and specifications imposed by these different environments and determine areas of convergence in terms of “core functionality”. The following two subsections briefly summarize issues related to central functional requirements and challenges for Darwin at home and Darwin at Industry, seeking to establish meeting grounds to formulate a general developmental roadmap.

### 2.1 Functional requirements and constraints: Darwin at Industry

The principle layout of assembly lines in industry has not changed drastically since their institution by Henry Ford in the beginning of the 20th century. Even in highly automated production lines such as in the automotive industry, a surprisingly large number of assembly steps are still done manually. Clearly, there is always the pressure to automate such manual processes, but in some cases this is economically infeasible. In addition, there are a number of crucial issues that
need to be considered while developing a novel technology to modify the industrial assembly process.

**Human Factors:** Aside from obvious economic considerations, human factors have become a highly relevant topic considering that many assembly processes require steps such as lifting heavy objects, precise positioning of objects above one’s shoulders or operating in places that are hard to reach for human. Such workplaces are generally labelled “red” in industry, because they may have negative long-term effects on human health. This is becoming increasingly important also because of the aging work-force. Thus there is now a trend to automate such work-places even though it reduces or sometimes does not lead to any short-term profits.

**Cycle time:** A crucial requirement in industrial assembly is the cycle time required for the assembly step. High-performance assembly machines operate at cycle times in the range of a few seconds. If objects are large or heavy, cycle times increase to 20 or 30 seconds, sometimes up to one minute, but rarely more than that. The reason for this limit is once again is economic feasibility. *If production rates are low, then it simply does not pay off to automate the process. Rather use manual work and benefit from the flexibility of humans.* If any technology is used in such processes the target cycle time should be in the range of 15 to 30 seconds for assembly operations of medium complexity.

**Safety and Coexistence:** In order to coexist and learn from a human, the robot and the human have to operate in the same workspace. Existing safety systems in industry do not deal with such complex interactions, because they lack the capability to interpret what is going on. The second problem is more fundamental. A cognitive robot is an experience-based device. During the training phase it is subject to changes in behaviour that is a function of its current training state, which makes it unpredictable to a large extent (for a cooperating human). At the same time a good GUI depicting the internal state of the robot, compliant physical interactions during action generation, user friendly communication protocols may be needed to ensure safety. There are several European research projects that specifically deal with safe cooperation of humans and robots in the same workspace. Darwin does not address these issues, but we should be aware that the industrial exploitation of the Darwin robot fully depends on the existence of such systems. Based on these constraints, the preferred solution for several industrial work-places is a ‘semi autonomous’ robotic assistant rather than a fully automated system.

**Repair and Maintenance:** In case a ‘semi autonomous’ or ‘cognitive’ robotic assistant is deployed another industrial issue is its repair and maintenance. The complexity of a cognitive robot caused substantial concerns about how to repair a robot when it does not perform the task as it should be. There have to be ways to deal with situations in which the robot fails, e.g. if the assembly task is not successfully completed in – say – 10% of the cases. *How to repair and maintain a device that largely depends on experience and that is unique to a certain extent needs to be considered seriously while installing a cognitive robot in a commercial production line.*

The value of any novel ‘learning’ technology will be measured by its ability to

1) bridge the missing connection between CAM software and a versatile assembly station, with minimal efforts related to maintenance and repair;
2) perform multiple tasks flexibly within the required cycle time and high accuracy to be economically feasible;
3) learn new tasks, particularly some of which that may be difficult for a human to learn in the same time (For example, identifying cables that look very similar and plugging the right cable to a particular type of engine, where there may up to 1000 variants)

4) coexist and cooperate with humans taking into account of stringent industrial ‘safety’ measures

Specifically the main tasks of where Darwin controller may be applied is in pick and place operations, screwing operations, inserting plugs, performing bimanual operations to create a subassembly of components and in most challenging cases ability to distinguish and handle elastic or non rigid objects.

Thus in a typical industrial scenario, clearly there is a need to augment strict requirements of speed, accuracy and dexterity (that determine the economic and short term feasibility of any novel technology) with task independent flexibility, adaptability and ability to cooperate and learn from humans (that determine the long term benefits of such a technology).

2.2 Functional requirements and constraints: Darwin at Home

Importantly, a central objective of the having the service robotics scenario in Darwin project was to use the state of the art humanoid robot iCub to gain insights into human cognition in general and test it simultaneously in applied industrial tasks (as mentioned in previous section). In this sense, iCub serves an experimental test bed in Darwin, where preliminary insights gained in terms of perceptual, motor and cognitive capabilities that iCub acquires by playing with various ‘toys and tools’ can simultaneously be transferred and tested in similar situations in industry. Hence arises the special requirement to formulate ‘scenarios’ that allow this kind of transferability into application domains. This is also the reason for timing the service robotics demonstrator well ahead of the vaguely similar industrial situation, to allow transfer of technology from lab to industry.

Of course, majority of the constraints listed in the industrial robot environment in technical terms also applies to the service robot scenario (i.e. the iCub humanoid). However the emphasis is not very much on outperforming the human but rather to gradually learn the basic skills and knowledge to assist and perform alongside humans in ordinary day to day tasks, handling a large set of objects in unstructured environmental conditions. Since the natural target applications in service robotics are in the home and assisted living market, there is an additional need to have cooperative and friendly interfaces and runtime adaptability.

It may be worth pondering on the area of divergence, constraints and additional issues where work conducted in Darwin on the service robot scenario can be even more relevant: Table 1 provides a comparison of the sensory and motor capabilities of the industrial and service robot used in Darwin. The most important area of divergence and constraint is the is the fact that since the service robot (i.e. iCub) was originally devised for the purpose of better understanding human cognition, there has been a conscious attempt to design it within the boundaries of human sensory motor capabilities. Thus, there are no laser scanners or the presently well-known Kinect system or other enhanced perception mechanisms to add to the basic stereo vision system. However, to augment the sensorimotor capabilities more like humans there is fully functional haptics (touch sensing in the fingers, palm and skin in the body), proprioception and fully functional force control. In addition the robot is anthropomorphic and highly redundant (30 DoF in the upper body in comparison to the standard 7 DoF industrial arms). This is supplemented with highly dexterous hands (with which even objects like screws have to be grasped and not pneumatically transported to predefined locations). All this has to be coordinated in a task specific fashion, both during learning and goal directed action synthesis, with considerations of safe interactions and minimal
self damage. Having multiple sensory and motor channels further comes with two additional constraints:

1) The necessity to transmit all this massive information to appropriate places in real-time (limited by the bandwidth of the communication channel in the body). This particularly limits the resolution of visual images acquired from the cameras that can be transmitted at real time with the available bandwidth, also considering all other information (touch, force, proprioception, audition) that also needs to be sent simultaneously. Presently visual information occupies half the band width of the state of the art communication channel in iCub.

<table>
<thead>
<tr>
<th>Functionality</th>
<th>Service robot</th>
<th>Industrial Robot</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vision</strong></td>
<td>6 Degrees of freedom head sporting two cameras (640x480 pixels), color, @ 20 frames/sec</td>
<td>Two cameras (1280x960 pixels), color</td>
</tr>
<tr>
<td>Laser profile scanner for 3D reconstruction</td>
<td>Not available</td>
<td>Available</td>
</tr>
<tr>
<td><strong>Robot Body</strong></td>
<td>53 Degrees of freedom humanoid of which 32 degrees of freedom are distributed in the upper body in an anthropomorphic fashion.</td>
<td>Two industrial robots Satubli TX90 and Satubli RX130B, each having 6 degrees of freedom</td>
</tr>
<tr>
<td><strong>Grasp mechanism</strong></td>
<td>Five fingered hand, 8 degrees of freedom are distributed in the Thumb, index and middle finger thus allowing a fairly large dexterity</td>
<td>Both Pneumatic and Jaw grippers are available and will be deployed based on the nature of the task and objects being handled</td>
</tr>
<tr>
<td>Proprioception (body configuration)</td>
<td>Available</td>
<td>Available</td>
</tr>
<tr>
<td>Touch sensing, Haptics</td>
<td>Both hands are integrated with 108 tactile sensors in the fingertips and palm, hence permitting use of touch information during interaction (to augment vision)</td>
<td>Not available</td>
</tr>
</tbody>
</table>
2) The more scientific question is the requirement to integrate and make sense of such multimodal sensorimotor information to modulate subsequent behavior.

Of course, having additional sensory and motor capabilities (not typically available in humans) helps immensely from the short term industrial application viewpoint and enhances the solution developed in the ‘cognitive robot’ environments. But at the same time it is also true that such solutions just provide a short cut with which the ‘hard problems’ related to human perception, motor control and skill learning are sidelined for the time being. In Darwin we have tried to follow a middle path, and find a balance between the quest to better understand fundamental principles and exploitation of results to industrial applications.

(B). As in humans, learning is expected to be cumulative and of progressively increasing complexity, and to result from self-exploration of the world in combination with some inputs gradually coming from social interaction (though imitation and linguistic communication). There are two questions and three functional requirements that need to be addressed. The questions are

1. How can the functional capabilities of a service robot continuously develop without frequent intervention of an engineer once it is "out of the factory"?
2. How and what can it learn through natural social interactions with humans, most of which is underspecified at design time?

The value of any learning and any technological advancement in abstract terms (as compared to the industrial case) are:

1. The need to focus on the formation of cognitive capabilities rather than these capabilities themselves;
2. The need to focus on embodied and situated sensorimotor and social skills acquisition instead of assuming the capability of advanced symbolic computing capabilities focused on solving abstract symbolic problems as is prevalent in several Artificial intelligence based techniques (Oudeyer, 2011).
3. The need to focus on task independent self-determined learning rather than task-specific inference over "spoon fed human-edited sensory data" (Weng et al., 2001) as is prevalent in most machine learning techniques.
4. The need to focus on functional modeling of integrated architectures of cognitive systems, rather than excessive significance to which brain areas are activated under which conditions (as prevalent in several neuroscience studies). Though some of these studies are highly relevant, there is a need for more functional and causal interpretations of data rather than seeking simple ‘associations’.
We believe that extracting the best out of both the industrial constraints as well as the constraints related to operating in open ended environments in natural environments inhabited and created by humans will help realize Darwin architecture its full potential in the coming years. The core competences developed during the different scenarios of increasing complexity in the next section are a function of both application oriented requirements as well as the intrinsic quest to better understand our own perceptions’, actions and reasons by using a state of the art humanoid platform meant to operate in ordinary day to day environments (i.e homes, offices etc).

3 A roadmap of Environmental scenarios to aid cumulative learning in Darwin systems

Noteworthy insights from the field of developmental psychology suggests that not just functional capabilities but also the environment particularly in the zone of proximal development (Vygotsky, 1978) plays a central role in the expansion of cognitive horizons in embodied agents by triggering appropriate acquisition of new knowledge through learning. The previous section particularly emphasized on the need to formulate tasks such that insights gained in terms of perceptual, motor and cognitive capabilities that acquired by the service robot (iCub) playing with various ‘toys and tools’ can effectively be transferred and tested under situations in the applied industrial environment. Undoubtedly, this is a guiding principle for formulation of the environmental scenarios. Nevertheless, there is a need to start formulating a roadmap from the very basics to the outermost core. As we move from the basics to the outermost core the task complexity naturally increases. This has to be complemented with a ‘demonstrable’ increase in the perceptual, motor and cognitive capabilities of the two robots. We broadly categorize this increase in ‘functionality’ of the Darwin controller terms of four loops:

1. The ‘Identify-localize-plan-reach-grasp-transport’ loop
2. The ‘Decomposition of Perception-Composition of Action’ loop
3. The ‘Goal-Affordance-Skill-Value’ loop
4. The ‘Disassembly-Reparation-Assembly’ loop

Around each of these functional loops are centred the environmental ‘scenarios’ for both iCub and industrial robot, that aid development of the anticipated functional advancements. Also the scenarios associated with the different loops are not formulated in a sense that they sequentially follow each other, but have temporal overlaps to also build up the cognitive repertoires both horizontally and vertically. At this point, we admit that the task related to the final milestone (i.e. the reparation scenario) is still blurred. Though the functionality we aim at is rather well defined. In the following subsections we present a concise summary of the environmental scenarios of gradually increasing complexity related to each of the above milestones with a short description of the relevant task (among a pool of other similar ones), role of different groups, core functional advancements, preliminary results (where applicable) and scope of what can be transferred to the industrial robots (with the related application scenario).

3.1 The ‘Identify-localize-plan-reach-grasp-transport’ loop
3.1.1 Service robot (iCub)

**Task S1: Pick up the ‘top most’ object in the scene and place it at a specified location**

*Short description*

A collection of objects (wooden blocks) of various shapes and sizes will be placed at random locations in the scene. Some of the objects may be stacked on top of each other. The goal of the robot is to grasp the top most objects in the scene and transport it to a container.

*Objects: Wooden blocks of various colours, shapes and sizes*

*Functional requirements:*

In general, this scenario relates to the deployment of the first prototype of *integrated* perception-action system of Darwin. It is a fundamental requirement to have this basic loop highly robust in order to attempt more complex tasks or any closed loop real-time learning in the future. The scenario is unstructured. Specific competences related to this environment include:

1. Ability to classify, identify objects in the scene based on simple visual properties like colour and texture
2. Ability to localize and estimate 3D position of different objects in the scene, estimate simple geometric attributes like width, length (to be used for action planning). Identify the top most object in the scene.
3. Ability to generate dexterous movements using the iCub upper body (using PMP based forward inverse models) in order to reach the target object, with the appropriate hand, hand pose and pick it up. PMP stands for Passive Motion Paradigm (Mohan and Morasso, 2011).
4. Ability to predict the consequences of the generated movement in terms of success/failure and alter action plans.
5. Ability to detect contact through touch, preliminary integration of proprioception, force/torque sensing to aid safe and compliant interactions during action execution.

*Collaborative efforts:*
Figure 2. Collaborative efforts in the ‘Identify-localize-reach-grasp-transport’ loop

Timeline: Month 12
Results: See section 5.2 and 5.3

Other variants: The ‘Identify-localize-plan-reach-grasp’ loop is very basic and applies to most physical interactions, independent of the task and objects involved.

Task S2: Perform a sequence of ‘stacking’ using MECCNAO 2+ Kids play objects

Short description
This is the first scenario using the MECCANO 2+ toy kit for 2 year old children. The task involves stacking two MECCANO blocks on top of each other, such that the holes in them are aligned and then inserting a MECCANO screw inside the aligned holes. Since the toy kit is designed for 2+ children, the objects are of appropriate dimensions to be handled by iCub and recognized by the vision systems.

Objects: MECCANO 2+ TOYS: Mainly coloured blocks with holes at the center, MECCANO screws (3cms in diameter)

Functional requirements:
The basic functional requirements are similar to scenario S1. However the task is extremely challenging in comparison to S1 in terms of accuracy required in perceiving and acting on MECCANO objects, task of achieving alignment of the first two objects (visuo-motor coordination), Grasping the screw (assumption is that screws are ‘standing’ with head upwards) and successfully inserting it in the hole. Even though the task appears simple, it is one of the basic ‘sequence’ of actions in several assembly operations (most of which is till date done manually in the industry).

Collaborative efforts: Even though broadly similar to the previous scenario, this scenario requires much tighter ‘real-time’ and ‘closed loop’ integration of vision, haptics and force sensing (for example, when hand occludes the object of interest like the screw during insertion) and motor control. Novel control schemas to integrate all this perceptual and motor information through techniques developed by CVUT-FORTH-IIT may be envisaged.

Timeline: Month 10-18
Results: See section 5.3.1 for preliminary results on manipulating MECCANO objects

Other variants: There are several variants like aligning appropriate shapes, inserting appropriately shaped objects in appropriate holes, picking up other relatively small objects similar to screws etc.
3.1.2 Application domain/ Industrial scenario

**Task I1: Industrial Pick and Place task belonging to “red” work-zone**

**Short description**

There are many industrial scenarios that require a more flexible (and task independent) ‘Identify-localize-plan-reach-grasp-transport’ loop, especially in the work places labelled ‘red’ which require handling of heavy objects that may have long term impact on the health of humans performing such tasks. These are areas where more flexible technologies can be extremely beneficial both in terms of human health and economic returns. A typical scenario involves removing the cover of a box picking up the object (industrial part) inside the box and placing it on a conveyor belt, then picking up the empty box and placing it on the stack of empty boxes. An example of a typical industrial exploitation scenario of this functionality is the ‘Test case 1’ explained in detail in D.2.2 (“Exploitation plan and industrial assembly task selection”).

**Timeline:** Month18.

3.2 The ‘Perceptual Decomposition – Compositional Action’ loop

The scenarios in this section are designed to exploit the basic perception-action loop developed for the service robot in the previous test case and take the Darwin controller to next level of complexity.

3.2.1 Service robot (iCub)

**Task S3: Assemble the ‘final configuration’ in the scene, given a set of building blocks**

**Short description**

A simple assembly of objects is initially created by the teacher and shown to the robot. Then a set of constituent parts (along with some irrelevant parts) are placed in front of the robot. The goal of the robot is to try and reassemble the final configuration shown to it (and learn from its experiences).

**Objects:** Wooden Blocks as in S1

*Toy clown assembly kit:* There is a central stand into which wooden blocks (of various sizes and having different surface markings) have to be inserted in proper sequence.

**Functional requirements:**

Building up on the sensory and motor capabilities developed previously, this task additionally requires:

1. Visual perception of more complex object properties mainly shape and surface markings (in the toy clown scenario);
2. Ability to perform more complex motor actions (including bimanual ones) on various objects based on their affordances. Dynamically take into account multiple constraints like
the appropriate wrist orientation to grasp a particular object, obstacles caused by other objects in the workspace, internal constraints (like joint and torque limits) during the motion planning process; Highly accurate positioning is also desirable to be able to insert a block into the central stand (as in the toy clown);

3. Ability to perceptually decompose a compound object into a set of constituent parts,

4. Ability to generate action sequences to assemble a compound object given a set of constituent parts;

5. Preliminary acquaintance with ‘part-whole’ relationships, spatial concepts (like top, down);

6. This task also requires creation of a working memory and long term memory structures that store various experiences related to actions and their consequences

Results: NA

Collaborative efforts:
Collaborative efforts may be summarized by figure 5 (that also summarizes what new contributions may be needed in addition to the previously existing functionality).

Timeline: Month 24
Results: NA

Other variants: An interesting variant that is rather well known as a test for planning capabilities i.e. Tower of London task. Note the similarity in terms of assembling a final configuration of objects starting from an initial configuration, with the service robot also requiring complex manipulation capabilities. In
humans, planning capabilities are frequently tested with the Tower of London (TOL) task (Shallice et al, 1982, Duff et al, 2011). Deficits in performing the TOL-task have been shown to correlate with specific lesions in the prefrontal cortex. In the TOL-task the subject is asked to sequentially move three beads of different colors from an initial configuration to a goal configuration (see figure). The number of moves required to go from the initial position to the goal position can vary. The task gets more and more demanding with the number of steps involved. For a one-step problem the necessary move is directly linked to the goal configuration. A five-step task does however comprises of moves that lead to configurations that are not directly related to the goal-configuration (Unterrainer, 2005). Note that the TOL-task cannot be solved in a systematic iterative way, opposed to the Tower of Hanoi task (at task similar to the TOL-task) for which a recursive algorithmic strategy exists (Ward et al, 1997). Solving the task is both a challenge for the reasoning system (to perform mental simulations of sequences of actions to arrive at the goal) and for the action generation system (to perform sequences of assembly-disassembly actions) to construct the goal configuration.

3.2.2 Application domain/ Industrial scenario

Scenario I2: Inserting ‘screws’ into an industrial work piece and create a sub-assembly of low complexity

The scenario is a typical assembly task of low to medium complexity that is still done manually and could possibly be automated, given a sufficiently intelligent robotic system. The task firstly involves picking up a wheel from a stack of wheels and placing it on a carrier system (similar to inserting constituent blocks into a stand in the toy clown task). Then the robot is required to pick up
6 screws from a box and insert the screws at appropriate position on the wheel. To add to the complexity of pick and place, inserting and precise positioning of screw, cycle time of 20 to 30 seconds may be challenging for the robot, especially if only one screw is taken at a time.

![Image of wheel, carrier, screws, and final object](image)

*Figure 6: Industrial scenario for the ‘Perceptual Decomposition-Compositional Action’ loop*

### 3.3 The ‘Goal-Affordance-Skill-Value’ loop

The scenarios in this section are devised to both build up on the functionality developed in the previous task environments, with a special emphasis on affordance learning, motor skill learning and sensorimotor concepts related to spatial and physical causality (that are most relevant for a dexterous assembler robot working with embodied intelligence).

#### 3.3.1 Service robot (iCub)

**Task S4: Learn to assemble the ‘tallest’ stack given a random set of 4-5 objects (with varying sensory motor affordances)**

*Short description*

Here the focus on learning properties of objects through goal directed and curiosity driven physical interactions with them (both of which can be rewarding to the robot). Particularly, this scenario builds up on the previous scenarios but with a twist. While scenario S1 was to pick up the top most objects in the scene, scenario S3 was to create a goal configuration shown out of constituent parts, *in this scenario goal is to create the highest possible stack, given a random set of 5 objects*. Even though the goal is to make a tallest stack, in this simple scenario, there is a lot to learn about the properties of the objects and actions that can be performed on them. For example, you cannot place an object on top of a pyramidal object or a ball or a cylinder oriented with its circular edge on the top. Similarly, putting something inside a container does not increase the net height. The base has to be object with greater geometric dimensions. Acting on all these different objects require complex motor control (some times bimanual), to change the orientations, pick and place a large cylinder etc. So even though the scenario seems trivial for humans, there is a treasure of ‘common sense’ knowledge related to affordances and physical causality that can be learnt from this task.

*Objects:*

Wooden blocks of various shapes, colors and sizes, large cylinders made of sponge to reduce weight (see figure)
**Functional requirements:**

1. Co-ordinate movements of head/eyes and two arms/hands to grasp manipulate and experiment with more complex, sometimes novel objects.
2. Learn and store abstract perceptual properties (colour, shape, texture, geometric attributes) of various objects and their motor affordances (how to reach, grasp, transport);
3. Learn the relationship between different objects based on sensorimotor interactions between them;
4. Learn to reason and generate actions that are most valuable (based on the sensory and motor properties of objects at hand, past experiences with them and the goal).

**Collaborative efforts:** The collaborative efforts are shown in the figure that covers this complete subsection and not restricted to this scenario.

---

**Figure 7. Collaborative efforts in the Goal-Affordance-Skill-Value’ loop**

*Timeline: Month 18-36*

*Results: NA*

*Other variants:* This task has as many variants as the combinations of the objects involved, in this sense the task is still fluid.

**Task S5: To learn to Push ‘intelligently’ to support various ‘goal directed’ actions**

*Short description*
This task is related to the previous task in terms of learning object properties through physical interactions but focuses on the acquisition of force related and spatial concepts that are relevant for a dexterous assembler robot to operate in varying environments. ‘Pushing’ itself is a complex action that has different levels of complexity and utility in the context of the objects, environments and goals under which this action is employed. From ‘pushability’ to consequences that ‘pushing’ (and pulling) ensues, the task has been investigated under several interesting conditions in the field of animal and infant cognition. It is further, a fundamental motor skill after reaching and grasping. Acquiring the abstract motor skill of ‘pushing’ in a task independent fashion is the first challenge posed by this task. Reasoning to deploy the skill in a ‘task specific’ fashion (with appropriate body effector) and based on environmental conditions in order to realize a high level goal is the second challenge posed by this task. The third challenge is related to the previous two i.e. executing the movement, perceiving the consequences and adapting the next ‘micro movement’ in real time to successful realize the planned strategy.

**Functional requirements:**

The task is aimed at enhancing the previous knowledge gained regarding properties of various objects in terms of how to ‘reach-grasp-assemble’ them, to how to make them ‘reachable, graspable, assemblable’. This requires learning and abstraction of core ‘spatial, contact related and force related’ concepts:

1. Relation of Pushability with object affordances (shape, size, etc)
2. Relation between ‘contact’, virtual contact and ‘pushability’
3. Relation between force applied and consequence of pushing (directionality, trajectory of motion)
4. Effects of ‘counter force’ (i.e. obstacles) and how to generate pushing actions that counter them
5. Chunking ‘pushing’ action as a ‘sub-goal’ dynamically in order to realize a ‘goal’

**Collaborative efforts:** Same as scenario S5.

**Results:** see section 5.5.2 for preliminary exploration in this direction

**Timeline:** M 8-M36

**Other variants:** A number of scenarios of increasing complexity inspired from several works on animals and infants can be created. Some of the tasks related to ‘pushing-pulling’ may also involve human robot cooperation to solve the goal.

**Task S6: Learning to use common tools and their resulting consequences (in terms of extension of reach, amplification force and task specific coupling of body and tools)**

Short description: While the previous two scenarios focussed on learning concepts and affordances through physical interactions with different objects, this task focuses on tool use. In addition to the ‘cognitive’ and ‘functional’ implications of tool use in cognitive robots (Mohan and Morasso, 2011), the scenarios place special emphasis on:

1. Learning motor skills through a combination of imitation, physical interaction and motor imagery (Mohan and Morasso, 2011, Mohan et al, 2011a);
2) Recycling the knowledge of movement across ‘skills’: motor equivalence, procedural memory and task specific compositionality of motor knowledge (see Mohan and Morasso 2011 for a summary).

Objects: Sticks of various shapes (Use: extension of reach, push-pull operations)
   5 DoF toy crane (Use: Bimanual control of a tool to transport useful objects to work space)
   MECCANO Screw driver (Coupling of two objects)
   Toy Lever (Amplification of force)

Functional requirements:
   1. Perception of Action (of the teacher)
   2. Motor Skill Learning through Imitation and Interaction
   3. Motor Knowledge Abstraction
   4. Motor Knowledge ‘recycling’
   5. Tool use

Collaborative efforts:

This scenario results from ongoing work of IIT on motor skill learning with anticipated future collaborations mainly with KCL and CVUT and involvement of PROF for the industrial robot environment.

Timeline: M10-M40

Results: See section 5.3.2 for basic architecture and results in this direction

Other variants: New tools/tasks may be considered in the future, at present none.

3.3.2 Application domain/ Industrial scenario

Task I4: Industrial assembly scenario that involves bin-picking, creating a sub-assembly (using bimanual actions) and fixing the sub-assembly on a larger part

There are two specific industrial scenarios where the functional capabilities can be exploited that are described detail as ‘task 4’ and ‘task 5’ in D.2.2 (“Exploitation plan and industrial assembly task selection”).

3.4 The ‘Disassembly-Reparation-Assembly’ loop

3.4.1 Service robot (iCub)

Task S7: Complete MECCANO 2+ Assembly task

Short description
Perform a complete MECCANO 2+ assembly task (like assembly of a ‘dog’ as shown in the example in figure or something similar), given a collection of constituent parts and tools for assembly. Even though the original scenario was designed for two to three year old infants, it is worth noting that it is a nontrivial task requiring complex visual perception, motor coordination, tool use, some understanding of spatial and physical causality, well developed reasoning capabilities to plan the action sequence for assembly (with prediction of consequences) and runtime adaptability to unforeseen circumstances.

![MECCANO assembly scenario: Assembly of a Dog](image)

**Figure 8. Full MECCANO assembly scenario: Assembly of a Dog (several other assembly operations can be performed with the same kit and tools)**

**Functional requirements:**

The task requires goal specific integration of all the major functionality listed in the previous scenarios. At the same time, the task requires a full maturation of the Darwin controller in terms of its functional capability to:

1. Develop ‘assembly’ plans autonomously based on incomplete or ambiguous information, re-plan when current plans fail, generalize and reuse previously learnt plans to new settings
2. Know a lot about objects; predict how they behave in the world, what they are good for and how they can be exploited to realize otherwise unrealizable goals and to create new affordances

**Timeline: M36-M48**

**Other variants:** The example shown in figure is just one of the several scenarios that can be devised using the MECCANO 2+ toy set. Moreover the complexity of the perceptual, motor and cognitive capabilities and the actions involved are a function of the environmental conditions in which the assembly operation has to be performed.

### 3.4.2 Application domain/Industrial scenario

**Task I5: Complete Industrial Assembly scenario**

An example of a related industrial scenario is test cases 5-6 described in detail in D 2.2. Both these tasks require precise bimanual operations, a degree of spatial reasoning, need to distinguish
between very similar objects (appropriate cables and plugs that need to be attached to an engine) and handling of elastic components. The tasks require an additional training phase with demonstrations by an expert human.

4 Towards a practical perspective on ‘Learning’

The word ‘learning’ itself has a multitude of interpretations ranging from generalization/inference over “spoon fed, human edited data” (quoting, Weng et al, 2001) to actively acquiring knowledge through intentional interactions with the world (Oudeyer et al, 2011). While some interpretations risk the danger of being ‘over-simplistic’ others risk the danger of being ‘over idealistic’ (for example, it is a fact that everything cannot be human edited and it is also a fact that physical and social spaces are highly complex to learn/know everything). This calls for a need to have a ‘practical’ perspective that is focused on:

A) Delivering the core functionality required in the scenarios described in the previous section;
B) Without compromises on issues that may hamper the overall development and future extensions of the architecture;
C) A clear measure to evaluate how complex the scenario, associated learning and the demonstrated skill;
D) A set of pointers as to what the underlying assumptions and guiding principles are;

We briefly discuss the four points in the two subsections that follow specifically in the context of Darwin. The section on Muddiness deals with (B-C) while the section on pointers deals with (A-D).

4.1 Assumptions and Muddiness

In the early 90’s Howard Gardner (Gardner, 1993) suggested the necessity to look at intelligence not only as indicated by logical-mathematical reasoning aspects (IQ) or emotional aspects (EQ) but also through performance of tasks that require bodily-kinesthetic and spatial skills in an ‘end-to-end’ nature (i.e. from raw sensors to actuators & back). He argued that the latter metric may further help appreciate what humans do in their day to day lives (often in a seemingly effortless fashion) and what embodied machines envisaged to assist them in future may be expected to do. It is this metric that is most relevant for Darwin and other architectures that are expected to acquire the sensorimotor intelligence to achieve end to end functionalities in day to day tasks. **However, formulating what is good to assume, how the complexity of learning and intelligence varies as a function of assumptions made, how assumptions made translate into a guiding principles for the learning architecture are very difficult to estimate for such requirements.**

The term ‘muddy’ is often used in psychology to informally describe tasks that cannot be clearly defined or specified (for example: tying a shoe lace). In a seminal article (Weng, 2009) proposed the idea of muddiness as a possible metric to find a practical balance between ‘what is the goal’, ‘what assumptions are allowed’, ‘what is learnt’, ‘how complex is the learning’ and ‘how complex is the acquired intelligence’ of the agent (human, animal, robot) executing the goal. Basically ‘muddiness’ allows intuitive quantification of the difficulty level of tasks (across domains) and hence facilitates comparison between complexities posed by any two tasks (like, chess playing vs. learning to make the tallest stack). It is independent of species or technology performing the task: a muddy task for a cat is muddy for a human as well and for iCub, may be cats and humans are
better in some cases! Hence ability to solve tasks with greater muddiness is indicative of the intelligence of the performer, helps compare state of the art intelligent machines and appreciate what humans do on a day to day basis.

There are a number of metrics for estimating what is ‘muddy’, the cumulative ‘muddiness’ a measure of what is the challenge in terms of learning and autonomy of the agent executing the task. In the table below, we list some parameters of muddiness that is relevant to Darwin and compare it with scenario S1 and scenario S7. As in the original categorization proposed by Weng (2009), the universe is split into: the external environment in which the agent's brain works, including the agent's body; the input is the information that the agent's brain receives from the external environment; the internal environment is the agent's brain and does not include the body; the output is the motor actions commanded from the agent’s brain. Finally, the goal is the objectives of the tasks that the agent performs (or asked to perform). Each metric indicates its definition (and relevance in terms of learning techniques employed in the literature in general), what are the allowed underlying assumptions made in the context of the scenarios (S1 and S7, other scenarios in between have gradual increments, but follow basic measures set in S1). The cumulative sum of components allows an intuitive quantification of how assumptions are gradually narrowed, how complexity of performance and associated Darwin intelligence is expected to increase (or add up) from S1 to S7. Details are described in table 1.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
<th>Assumptions in Scenario S1</th>
<th>Assumptions in Scenario S7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rawness</td>
<td>Rawness is a measure of how much information is sensed (and not edited by the human). Inputs directly derived from the sensor without human editing have higher rawness.</td>
<td>Real sensor (Vision, proprioception)</td>
<td>Real sensor (Vision, proprioception, force/torque and Touch)</td>
</tr>
<tr>
<td>Size</td>
<td>Indicates the number of possible different values of input that the robot has to consider while performing the task. An image of 360x240 with each input taking values from 0 to 255 (further dependent on external factors) gives an astronomical number.</td>
<td>Moderate</td>
<td>Highly Complex</td>
</tr>
<tr>
<td>Background and Occlusion</td>
<td>Indicates whether there are inputs that are not related to the task (additional objects)</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Variation</td>
<td>Indicates the complexity of variation among inputs that require the same output</td>
<td>Nil to very low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Activeness</td>
<td>Indicates whether the robot must actively acquire input in order to perform the task</td>
<td>Low and almost restricted to head and eye movements (if necessary, that too automated)</td>
<td>High, includes creation of additional affordances</td>
</tr>
<tr>
<td>Multimodality</td>
<td>How many distinct sensory channels are employed and the need to integrate them</td>
<td>Moderate</td>
<td>Complex</td>
</tr>
<tr>
<td>External environmental factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Awareness</td>
<td>How much the user/programmer knows about the environment in which the robot is supposed to realize the goal</td>
<td>Partially known in terms of the allowed workspace, objects present</td>
<td>Unknown</td>
</tr>
<tr>
<td>Control Level</td>
<td>In an uncontrolled environment, the complexity of the environment is not bounded. In a controlled case several circumstances, objects, actions are disallowed by programming.</td>
<td>Moderate</td>
<td>High</td>
</tr>
<tr>
<td>Variability</td>
<td>Indicates if the environment changes constantly.</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Forseeability</td>
<td>Indicates the need to foresee future environments through internal simulation to realize the goal at hand</td>
<td>No, the goal is just to pick up the top most object in front of the robot</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Internal Factors**

| Size and Representation | The size of the state space including the present context (state vector), and the range of experiences stored in long term memory that needs to be interpreted to perform the task. | Moderately complex | Highly complex |
| Observability | Observability of internal environment means the degree to which the internal representation of the agent is observable by the outside world. As the cognitive architecture becomes more complex internal states become more autonomous and can only be observed by the resulting behavior. | Fully observable | Moderately complex (there will be a need of an excellent GUI to interpret the internal states in real-time) |
| Imposability | How easily a user can set the value of the internal representation of the agent. Note that this is related to observability because if the lesser the state is observable the greater is the difficulty to impose a state (on the part of the user.) | Imposable effortlessly | Imposable but with some efforts (in any other case than termination) |
| Temporal convergence | This relates to both time (of the past) and temporal order (of the future). How much the robot has to extrapolate forward and backward relates to complexity in terms of temporal convergence. | Simple | Complex |

**Output**

| Size | Size of output is similar to size of input. Generating a temporally ordered sequence of skilled actions with 30 degrees of freedom (and other tools that needs to be coordiantes) is more complex than reaching with an arm. | Simple | Highly complex |
| Terminality | Terminality may be more relevant for the industrial robot, but also in some ways to the service robot. It is a measure of how independent the generated output is in terms of further human interventions /processing. | Nil | Moderate |
| Modality and multimodality | How many distinct effectors are used (arm, 2 arms, bimanual plus tool, whole body and text/speech based messages, gestures etc) | Simple | Complex |

**Goal**

| Richness | This is a tricky but highly relevant criterion. | Moderately | Complex |
Richness of a goal implies how easily the goal that can be fully described in mathematical terms. If it can be described easily, it can be converted into an algorithm with little ambiguity.

Variability indicates if the goal may change, this includes opportunistic scheduling of goals autonomously based on evolving environmental conditions in the quest to attain rewarding situations.

This deals with the issue of whether the goal is given or created by the robot as a sub goal. This is a measure of autonomous intentionality of the robot.

How the goal is conveyed (text based or spoken language etc)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Possible Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richness of goal</td>
<td>Implies how easily the goal can be fully described in mathematical terms.</td>
<td>complex</td>
</tr>
<tr>
<td>Variability</td>
<td>Indicates if the goal may change, this includes opportunistic scheduling of</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>goals autonomously based on evolving environmental conditions.</td>
<td>Possible</td>
</tr>
<tr>
<td>Availability</td>
<td>Deals with the issue of whether the goal is given or created by the robot</td>
<td>Yes, the ability to</td>
</tr>
<tr>
<td></td>
<td>as a sub goal. This is a measure of autonomous intentionality of the robot.</td>
<td>instantiate sub goals</td>
</tr>
<tr>
<td>Mode of instruction</td>
<td>How the goal is conveyed (text based or spoken language etc)</td>
<td>Text based,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Most probably text based,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>though spoken language</td>
</tr>
<tr>
<td></td>
<td></td>
<td>interaction is under</td>
</tr>
<tr>
<td></td>
<td></td>
<td>development for the iCub.</td>
</tr>
</tbody>
</table>

Cumulative Muddiness = Intuitive measure of the sum of individual factors

Table 1. Central assumptions in learning and cumulative muddiness between scenario 1 and scenario 7

4.2 Guiding “constraints” and “challenges” for learning in Darwin

The fundamental assumption concerning learning in Darwin is that it happens as the self-organized results of dynamical interactions among brains, bodies and their physical and social environments. Trying to understand how this self organization can be harnessed to provide task-independent, open ended learning of skills is a challenge, especially in ‘muddy’ environments and basic criteria (in terms of constraints and functionality) listed in the previous sections. Below we list nine pointers and open challenges that we wish to give emphasis while creating the final Darwin architecture. There may be changes to this list in the future since at year 1 we are still swimming in shallow waters, surely with a vague idea of what lies ahead. The assumptions and pointers below at the same time may not be just relevant for Darwin but also for other adventures with similar goals.

4.2.1 Learning Green: Abstraction and Compositionality

Task independent ‘Abstraction’ and task specific ‘Compositionality’ go hand in hand. Task-specific learning signals and ad hoc experiments focused on accomplishing a single task have a strong limit for cumulative learning in that they can only drive the acquisition of new skills strictly related to the task(s) decided by the researcher. We believe that explorative processes should not guide the learning of a single task but a more general learning process that simultaneously acquires sensory-motor representations, skills, and solutions (that may be independent of the task itself). This ‘task independent’ learning and abstraction of learnt knowledge is a fundamental necessity for the evolution of next generation general purpose cognitive machines. Even tough manual programming has successfully created impressive machines; they do not perform
well in muddy tasks simply because a hand designed representation cannot account for changing environments. Further, tasks as well as what new could be learnt, experienced are generally unknown at the time of programming. **Hence, attempting to create a system with a ‘potential’ to learn ‘when needed’ in order to sustain interrelated constraints of both its own ‘survival’ and its ‘utility value’ from the perspective of the user is more practical for survival of Darwin itself.** Undoubtedly, the challenge is tough, but it’s nevertheless serves as an underlying criterion for measuring the success of Darwin. **Complementing this requirement is the parallel need to ‘assemble’ and ‘exploit’ such ‘task independent’ knowledge in a ‘task specific’ contexts.** An attempt to coherently unify explorative learning through physical and/or social interactions with mechanisms of abstraction and compositionality is crucial for a ‘domain agnostic’ architecture like Darwin. This may in turn pave way for Darwin systems that learn Green and actively recycle their past experiences in new contexts. *(See section 5.3.2 for preliminary results in this regards on skill learning).*

### 4.2.2 Body schema and the territory of Action without ‘movement’

Picking up from the last section and moving from “task” to “body” itself, the explosive growth in the complexity of the body in a limited number of species is typically concentrated in two body parts (the hand and the vocal tract) that do not have a specific function but are general-purpose tools (for manipulation and communication, respectively). Exploiting this general purpose tool (i.e. one’s body) in a task specific fashion brings in both complex coordination/control problems and multimodal perceptual problems. Further, in organisms with complex bodies surviving in open environments, acting is not just limited to reactive, albeit adaptive, stimulus-response mechanisms. There is increasing support for the view that action system in biological organisms effectively participates both as a generative model in perceptual processes and in structuring knowledge about the world. This view point closely resonates with the internal simulation theory that assumes that the brain simulates the environment and reasons on the simulation before committing to a given course of action. In this sense, overt actions are just the tip of an iceberg: under the surface it is hidden a vast territory of actions without movements (covert actions) which are the essence of motor cognition. A multireferential, dynamic representation of one’s own body (i.e. the body schema) plays a fundamental role in various contexts related to ‘action’. In simple terms, an iCub needs a malleable ‘body schema’ as much as a human or a chimp, in order to use its “complex body”, take advantage of it, and ultimately survive. Mounting evidence accumulated from different directions such as brain imaging studies (Grafton 2009, Kranczioch et al, 2009; Munzert at al 2009), mirror neuron systems (Rizzolatti et al, 2010) and embodied cognition (Gallese et al, 2011) generally support the view that action ‘generation, observation, imagination and understanding’ share similar underlying functional networks in the brain. Since these capabilities are highly relevant for efficient learning and reasoning capabilities in Darwin, **there is a fundamental requirement to develop sophisticated computational machinery for ‘Motor cognition’ that unifies the different aspects of purposive actions: execution, planning, reasoning, observation, imitation etc. (See section 5.3.1 for results).**

### 4.2.3 Developmental programs vs. Learning/knowing everything

Recently, a new “paradigm” *(autonomous mental development)*, has been proposed (Weng et al, 2001; Di Paolo, 2002; Prince and Demiris, 2003; Blank and Meeden, 2006) that stresses on the
fact that instead of directly training or evolving a set of functionalities in robots, we should endow them with developmental programs that allow them to autonomously develop the needed skills on the basis of prolonged periods of interactions with the environment and intrinsic motivations. It is well known that human children while playing carry out several activities driven only by intrinsic motivations like curiosity. These activities allow them to acquire knowledge and skills exploited at a later time to pursue useful goals. Further, it is a fact that the sensorimotor and social spaces inhabited by humans are so vast and complex that only a small fraction of it is actually ‘explored and learnt’ within an individual’s lifetime. We do not learn everything or anything. The gap between ‘potential to learn’ and ‘what is learnt’ is constrained by a wide range of ‘internal’ and ‘external’ forces that also play a role in limiting the space of action, exploration at the same time maximizing learning. Different authors have claimed that typical criteria used in active learning approaches, such as the search for maximal uncertainty or prediction errors, might get trapped or become inefficient in situations that are common in open-ended robotic environments (Lopes and Oudeyer 2010). Only very recently have these approaches been applied to robotic problems, and even more recently if we consider examples with real robots. Further, developmental constraints themselves cannot be fully specified at design time as both the future environment and user may bias the learning curve of the system. Undoubtedly, Darwin needs a developmental program but the issue is subtle. To enable sustained and meaningful interactions with the world requires a novel approach that takes into account several factors like maximizing learning within the zone of proximal development (Vygotsky, 1978), constraining unbounded exploration, retaining autonomy, allowing modulation by user with appropriate sensory interfaces, realizing the priority of user requirements. Finding the right balance may make Darwin a very interesting system to watch in the future. (No results in the user stories attempted so far, theoretical development is ongoing, see section 5.1).

4.2.4 Physical Embodiment vs. Multiple embodiments

The properties of embodiment, including geometry, sensory inputs, or innate motor primitives/synergies often encoded as dynamical systems, can considerably simplify the acquisition of sensorimotor or social skills, and is sometimes referred as morphological computation (Pfiefer et al, 2005). The interaction of the constraints of embodiment with other constraints is an important axis of investigation, especially in Darwin where there are multiple embodiments (industrial vs. service robots). At an abstract level Darwin is envisaged as a domain agnostic architecture. Hence despite the similarities and dissimilarities in embodiment and there is a challenge here in the requirement to ‘reuse’ and ‘transfer’ learnt knowledge in an efficient fashion from one embodiment to another (at an abstract level). This is a subtle issue within the axis of ‘Darwin’ architecture. There are no results so far but this constraint is in our mind while developing the necessary theoretical and software modules.

4.2.5 Embodied robot vs. Socially embedded robot

We all learn from each other. Like humans, even for embodied robots operating in an unconstrained environment: the sensorimotor spaces, including the body dynamics and its interactions with the external world, are simply much too large to be learned entirely through exploration. Unconstrained or infinitely long babbling and exploration is a myth that works only in computer programs, but not on real world systems that operate under a wide range of constraining
‘forces’ that ensure both learning and learning that does not take so long that ‘survival’ is sabotaged. Thus, there is a crucial requirement to ‘blend’ information acquired through explorative interventions with information gained through social interactions in order to both ‘constrain’ the space of exploration and effectively speed up convergence. Hence learning continuously through interactions with a teacher/user and hence the additional requirement of creation of appropriate interfaces to aid social learning assumes significance in Darwin (more importantly for the service robot). A practical perspective is something that generally a good teacher understands very well: finding a fine balance between spoon feeding too much (that results in poor learning) and spoon feeding too little (that results in lack of motivation to learn). Appropriate social embedding is crucial for Darwin like systems. (A preliminary example is in the area of motor skill learning by imitation see section 5.3.2).

4.2.6 Meta exploration vs. Meta exploitation

Like physical embodiment, social embedding itself has other indirect advantages that play a crucial role in speeding up learning (all by oneself). There are two practical consequences: The first relates to the issue of meta-exploration. All active learning systems are generally based on a measure of “interest” in particular zones of the sensorimotor/state space, and hence pushed to focus on zones of maximal interest. But the computation of this measure of interest for a given zone of the sensorimotor space requires that this zone be at least explored / sampled a little bit. In other words, we have a meta-exploration/meta-exploitation dilemma for achieving efficiently the “smart” exploration process. How this meta-exploration/ meta-exploitation dilemma could be solved (or cannot be solved without particular constraints) in high-dimensional spaces still remains to be understood. One suggested solution is that to augment the push (i.e. the measure of interest of the cognitive agent) with a parallel a pulling mechanism (that arises through social embedding) or influence of other individuals (like the user). This practically reduces the ‘areas’ where both physical exploration and learning is focused. But appropriate interfaces and integration mechanisms are necessary so as to allow appropriate ‘brain washing’ by a user but not necessitating any additional programming by an engineer. This topic is subtle and nontrivial but will nevertheless modulate the overall functionality of the Darwin robot. However more brainstorming is necessary in this direction.

4.2.7 Maturational Constraints, Incorporation of tools

In human infants the body, its capabilities and the neural system encoding such capabilities grow progressively rather than being full-fledged already at birth. This implies that new degrees of freedom, as well as increase in volume and resolution of available sensorimotor signals, may appear as learning and development unfolds. Transporting these mechanisms to embodied robots and understanding the acquisition of new knowledge is a central question in developmental robotics. A highly relevant example related to incorporation of new degrees of freedom as far as Darwin is concerned is the case of ‘Tool Use’. In addition to its relevance in terms of assembly process, investigation on acquisition of tool use skills in both primates and cognitive robots could open a new window to understand the several fundamental issues like the emergence of mind, the sense of self, the continuity of self in time, ‘other selves’ in other bodies and the horizontal spread of skills through culture through social interactions, human-human, human –humanoid (Iriki et al, 2008, Umilta et al, 2009, Mohan and Morasso, 2011). Hence, in Darwin there is an emphasis on the need to go beyond ‘reaching’, ‘grasping’ and ‘walking’ robots, and rather focus on the ‘grey
area’ of motor skills that begin once an object of interest is ‘reached’ and ‘grasped’. See section 5.3.2 for first results in this direction.

4.2.8 Memory and the generative use of experience

Open-ended learning and organization of memory in relation to what is learnt are two sides of the same coin. We don’t cook/store the food needed for all our lives. We just store the ingredients and compose the recipe instantaneously based on the situational requirements. Similarly, in muddy tasks, robots are naturally expected to generate goal directed actions retrieving ingredients’ from memory of their multitude of previously learnt sensory motor experiences (related to skill, consequences, value etc). Thus, independent of the sensory, motor and cognitive skills learnt, in Darwin there is an underlying need for generative use of experience both to realize ones goals and to learn further. This requires efficient organization of memory in order to store the learnt information in a format that is retrieved quickly based on context/value, can be causally combined with other information to generate new knowledge and creative outputs. Most importantly, being able to build generative models endows the system with the ability to imagine new solutions via generalization, to predict future events from past experiences, and to explain ambiguous events (by imagining a causal chain leading to that event). Hence creation of a growing, compliant, organized memory structure that enables storage, forgetting and swift task specific retrieval of information must go hand in hand with ‘learning’ in order to allow generative exploitation of experience (and hence cognition). This issue is still in the phase of theoretical development, first prototype scheduled around month 18, so there are no demonstrable results so far.

4.2.9 Sense vs. ‘Common sense’

The central challenge faced by Darwin system is to learn to perform dexterous motor skills and learn to autonomously reason about space, force, the associated physical causality related to objects and their behaviour. The scenarios developed in section 1 are aimed at creating environments that when supported by an appropriate developmental program aids acquisition of such knowledge. In other words, the net effect of the scenarios and the developmental program is towards driving iCub to acquire internal models that ‘abstract’ recurring patterns of ‘sensorimotor’ experience arising out its interactions with the world. (For example, learning that all objects fall when dropped, you cannot place a cube on top of a pyramid while making a stack (scenario S4), learning part-whole relationships (scenario S3), learning the effects of ‘force’ and ‘directionality’ when you push something (scenario S5), the effects of containment (in/out) with which you infer that if ‘B’ is inside ‘A’ and ‘C’ is inside ‘B’ then ‘C’ is inside ‘A’). At the same time playful interactions of the robot occurs in an ‘end to end’ nature (i.e. from raw sensors to actuators & back) and is not symbolically coded. This may entail two crucial considerations while developing the learning architecture for Darwin robot: 1) need to develop integrated representations arising from multiple sensory streams (proprioception, force/torque sensing, touch and vision); 2) Need to filter out task dependent factors and pick up invariants; Practically, there are strong reasons to believe that ‘common sense’ arises through ‘sense’ related to the robots actions with the world. Some preliminary work in this direction is already ongoing in relation with both motor skill learning and teaching iCub to ‘push’ objects intelligently. Though this is just the incubatory phase for these works, we are aware of their relevance for the future. Practically, bridging the gap between sense and commonsense in an analog, sub symbolic fashion is central to learning in Darwin.
5 Learning in Darwin: Ongoing progress and future challenges

In the previous sections, firstly, we devised a range of tasks organized into four basic groups to conduct explorative learning directed towards acquisition of various sensory, motor, conceptual and cognitive knowledge in Darwin robots. The central assumptions and practical guidelines we intend to follow in the future while creating the Darwin learning architecture was summarized. This section both presents progress in different directions achieved so far in the various subareas and key challenges ahead. Of course, in year I we are still swimming in shallow waters, in some areas there are significant results (skill learning, body schema, tool use), some have preliminary results that aid development of other components (mainly vision, and some primitive reasoning), some are under software development or undergoing integration with the robot (experiments on teaching iCub to push, affordance modules, touch), some topics are still under theoretical development and deliberations (motivation system, embodied concepts). Nevertheless, we present what is achieved so far with a perspective on how we will extend it.

5.1 Drives and Motivations

No significantly complex cognitive system operates in a motivational vacuum. For a cognitive system to develop and expand its knowledge of its environment it needs to support a developmental program which will allow it to be developed in terms of skills, capabilities, and knowledge in a life-long manner.

The DARWIN architecture suggests a Motivational component which will play the role of the driver of world exploration and accumulated knowledge exploitation. Modeling of motivational systems (or alternatively of drive systems) attracted in the last ten years a lot of interest in the field of artificial and robotic agents (Breazeal, 2002, Oudeyer et al, 2006, Kasderidis 2006) with specific conferences appearing on the field, typically under the name of ‘developmental robotics’ or ‘autonomous mental development’. These efforts are based on the increasingly detailed information arising from various neuro-imaging studies and also from psychological studies (Franken et al, 1997, Carver et al, 1998) or from explorations with animals and robots (McFarland, 1993) which took place during the last twenty years.

In the DARWIN project a motivational system model is being developed, which is inspired by brain and psychological studies, in order to provide the cognitive system with a generating mechanism for new explorative behaviors leading to new or refined skills, capabilities and knowledge acquisition. Even though the exact details are not yet fully determined, we can assume that the system will be based, at least, on the following core drives (and possibly more):

- Survival
- Pleasure
- Development

These are high-level processes which all contribute to the overall ‘well-being’ of a DARWIN agent. It is assumed that a Core drive is capturing the whole agent vs. environment interaction and it is
decomposed on the three aforementioned drives. In essence the Core drive is a mechanism, which tries to maintain an internal homeostatic variable to a range of suitable values. This variable could be called 'well-being'. Similarly each of the other drives is a mechanism that tries to keep inside its homeostatic range corresponding variables. For example the Survival drive monitors the agent’s energy, structural integrity, etc; Pleasure monitors the pleasure potential influenced by external & internal rewards gained by an action/plan, etc; Development monitors the learning potential including seeking novel situations, exploiting known ones, avoiding overly complex ones, etc. The (numerical) value of a drive is a function of the corresponding monitored homeostatic variable(s). This function could be Identity or otherwise. For example while energy can range from zero to a maximum value $E_{\text{max}}$ the corresponding drive’s value lies inside the $[0, 1]$ interval. In this way we achieve an independence from the underlying physical mechanisms that implement the agent / environment interaction and we focus instead on the internal mechanisms needed for controlling these physical variables. Thus we abstract the agent-environment interactions and we convert them to an internal abstract representation. In a similar way the internal state of the agent can be transformed to the same abstract representation and thus the global motivational dynamics takes place in a state space generated by the high-level drives. However the exact mechanism, which introduces transitions to the motivational dynamics, can be modeled in many ways. In previous literature approaches a hierarchical control system of drives (Carver C. & Scheier M. 1998), including high-level ones (as above) and lower level ones, was suggested; such an approach has limitations in its dynamics and in exploration of the possible state space generated by the high-level drives. Other approaches include modeling based on neuro-transmitters and homeostatic variables (Breazeal, 2002) as well as other ideas. Thus the exact details for the DARWIN’s motivational system is under theoretical development, mainly in the level of the global motivational dynamics and the exact mappings of homeostatic variables to drives and the result of external and internal environment in these variables. We expect the first prototype to be deployed along with the first prototype of Darwin reasoning layer scheduled around month 18. Further, one needs to balance the need for computational requirements, in a real robotic platform, against the range of generated behaviour. What we seek is a motivational system which can re-produce interesting and non-trivial behaviour albeit many orders of magnitudes simpler than the one observed in humans.

Part of a motivational system is an agent goal-(self) generating system which will provide new goals to a robotic agent to pursue when user goals are absent. In this way we will develop an agent which will not stay inactive with the lack of human commands, but instead it will be driven to pursue its own development and well-being. This abstract internal behavior, due to motivational dynamics, will generate explicit action plans which hopefully will lead to progressively acquire more refined skills, capabilities and knowledge or develop new ones. The exact details of the internal goal generation system need also to be defined during the DARWIN project. Part of the challenge includes the creation of a uniform, dynamic and global representation of both user and internally generated goals and the associated dynamics that enables the agent to autonomously/opportunistically execute one among them based on what the environment affords.

**5.2 Learning in Vision**

Primarily, vision is the main source of sensory information for both robots in Darwin. While there are ongoing plans to have a preliminary level of visuo-haptic integration (especially, for the service robot) in the near future, to start with there is a basic need to close the perception-action loop. While this was the prime objective of loop 1, at the same time any realistic ‘end to end’ learning is
not possible unless the basic loop between visual perception and motor coordination is closed in a robust fashion. To begin with, a robust, real-time detection of scene parts with uniform color was assumed to be a basic competence of the DARWIN visual recognition system. Color based recognition is also a reasonable starting point because most toys of the service robot (basically suited for 2+ children) come with vivid combinations of colors.

Although color-based recognition may seem easy, it is actually not. An object with a constant bidirectional reflectance function can give rise to pixels with very different colors caused by varying illumination and surface normal. Coping with varying illumination is especially challenging. It is widely believed that the human visual system includes a complex algorithm to reliably estimate illumination from many clues, which allows it to recognize colors to a great extent invariant to illumination (i.e. color constancy). Unfortunately, despite a lot of effort none of the proposed computer algorithms for color constancy seems to be good enough for reliable recognition of colors (Funt, Barnard and Martin 1998). A popular recent approach (Rother, Kolmogorov and Blake 2004) based on Markov random fields were employed to create the basic color segmentation module (see deliverable 3.1 for details). This module estimates identity of objects in the scene, computes the bounding boxes that feeds to the 3D reconstruction module (Mohan et al, 2011a) responsible for estimating the 3D location of the object of interest in the egocentric frame of reference of the robot (to trigger action). Figure presents the first results of the closed loop interaction in the ‘identify-localize-reach’ loop, to reach the top most object in the with a hand pose to afford grasping.

Parallelly work is ongoing to combine the basic color based object identification module other crucial information like texture and shape (as detailed in deliverable 3.1). With regards to shape perception other approaches are also being followed by IIT mainly for action perception and imitation (Mohan et al, 2011a, see section on skill learning) and neural network based approach of KCL in relation to affordance learning (see section  ). To add, color information can also be combined with other modalities such as shading or edges, or more complex objects could be modeled as consisting of a small number of parts each with uniform color, with additional constraints such as which colors can be neighbors to which. This seems to be particularly useful starting point for indentifying MECCANO toys (relevant to the initial scenarios using them). To add, specifically in relation to more complex assembly tasks related to industrial objects (and some MECCANO objects), recognition of elongated thin objects (various tools, pipes, pen-like objects, cables, etc.) and objects without surface markings is crucial. There is no established methodology to recognize such objects. A two pronged approach is being implemented. The first method for the wire-like object recognition is based on identification of curvilinear features [1] and other channel representation of the input.
signal that can be separated according to local orientation and scale. Objects will be described as co-occurring events in orientation-scale-position space. The second approach exploits advances in optimization methods based on the Markov cost function [2] and their combination with shape priors to segment parts of wire-like objects on a broad class of background (the non-compactness of wire-like objects, with background visible in every neighborhood that includes the foreground object is a major challenge). For objects without surface markings and distinctive shape, modified shape-from-shading methods will be deployed. As the name suggests, shape from shading methods have been developed in the context of reconstruction of 3D shape. For object recognition, it is not necessary to reconstruct the object, only to detect distinguishing, characteristic parts with distinctive shape and to use such parts as “local patches”. The approach will thus connect together the extensive research on shape-from-shading and the methodology of the “local patch” method.

Coming to concept development, some concepts are obvious to invent, and more or less correspond towards in an ordinary human language, examples are ‘green’, ‘cylinder’, ‘left of’, ‘slowly’. However, as far as visual perception is concerned, it is more and more realized that these obvious and high-level concepts are hard to detect reliably in such a way and richness that would allow us to build reliable and non-trivial reasoning over them. Instead, more important may be
low- or mid-level visual concepts that do not directly correspond to words of the ordinary language and perhaps are used by humans unconsciously (and therefore information about them is much harder to obtain by direct introspection). Examples are various features and descriptors nowadays successful in computer vision: Harris corners, external regions, cascades of weak classifiers (used in AdaBoost or face detection), the features used in Video Google approaches (Sivic et al, 2003), SIFT descriptors, and others.

These can also form hierarchies, e.g. such as in (Fidler et al, 2007) learned automatically in a semi-supervised manner. In terms of cognitive science, this is related to Gestalt psychology, and concept/symbol grounding. In fact, while some concepts are directly 'grounded' in percepts, others are formed from these low-level concepts. This is sometimes called 'symbol tethering' (Sloman et al, 2004). Along these lines, a part of the research (somewhat overlapping with WP3) is to develop new low- or mid-level visual concepts. While much effort in computer vision has been on highly textured regions, features and descriptors in texture-free objects are considerably harder. At the first sight, obvious such concepts are edges and line segments, however, it turns out that these are not reliable and invariant enough. We need concepts invariant to pose, scale, blur, and illumination, and which would enable robust visual recognition, matching, indexing, or reconstruction. In more detail, the current line of research being pursued is bottom-up grouping, similarly as (Gioi et al, 2010), but using global grouping techniques such as Markov random fields.

**5.3 Learning in Action**

The ability to perform/learn dexterous motor skills is a fundamental requirement to enable Darwin robots to play, interact, transform and learn from their world (and people around it). In general, for a complex body interacting with a complex world, the goal of any movement is to position the body in more rewarding states. In order to do so, it is now well known in that action related networks play an important role not only in producing overt movements but also in simulating the environment and reasoning on the simulation before committing to a given course of action. Hence, as advocated in section 4.2.2, there is a basic need in Darwin to develop sophisticated computational machinery for 'Motor cognition' that unifies the different aspects of purposive actions: execution, planning, reasoning, observation, imitation etc. A significant effort was invested in this direction in year 1 to get the first prototype of such an ‘action’ system functional in the service robot (iCub). While a detailed description can be found in a recent article (Mohan and Morasso, 2011c), in the sections that follow we summarize the main theoretical insights in different contexts related to ‘action’, results obtained and some pointers to developments envisaged in the near future.

**5.3.1 Body schema for embodied robots: The Passive Motion Paradigm (PMP)**

*The need for a body schema*

Why does a complex robot need a body schema? For the same reason for which a human or a chimp needs it: simply put, without one, it would be unable to use its “complex body”, take advantage of it, and ultimately survive. The crucial consequence of such complexity, as formulated by Bernstein about 80 years ago, is the Degrees of Freedom Problem (Bernstein, 1943). Along with a complex body comes complex coordination problems, the need to effectively integrate intrinsic space (the body’s spatial properties, hierarchical arrangement of limbs, shape of body
surface etc) with the extrinsic (peripersonal) space, the need to dynamically simulate spatiotemporal organization of action and their ensuing perceptual consequences in the context of the pursued ‘goal’. Clearly, having a complex body without such active mechanisms would be a drawback in natural or artificial phylogenesis since purely reactive, adaptive architectures would suffer from the curse of dimensionality (as the number of useful/necessary coordination mechanisms among DoF’s grow up exponentially). The neuronal basis of the notion of body schema has been the topic of a number of studies (e.g. Paillard 1999, Gallagher 2005, de Vignemont 2010), focusing on subtle distinctions such as the differences between motor body schema and perceptual body schema (or body image) which are associated with different neuronal subsystems, such the ventral and the dorsal pathways. From a more functional point of view, focused on the artificial physiology of an embodied robot, such distinctions are irrelevant. In contrast, direct inspiration for a biomimetic approach to motor cognition in embodied robots comes from experimental evidence accumulated from different directions such as brain imaging studies (Grafton 2009, Frey and Gerry, 2006), mirror neuron systems (Rizzolatti et al, 2010) and embodied cognition (Gallese et al, 2011) that generally support the idea of common underlying functional networks subserving both the execution and imagination of movements. This leads to the formulation of the Passive Motion Paradigm.

**PMP: The Central hypothesis**

Considering the mounting evidence from neuroscience in support of common neural substrates being activated during both ‘real and imagined’ movements, it is not unreasonable to posit that also real, overt actions are the results of an ‘internal simulation’. This is the central idea behind Passive Motion Paradigm (PMP) i.e. actions (overt as well as covert) are the consequences of an internal simulation process that ‘animates’ the body schema with the attractor dynamics of force fields induced by the goal and task specific constraints. The internal simulation is analogous to the coordination of the movement of a marionette by means of attached strings: as the puppeteer pulls the task relevant tip to a target, the rest of its body elastically reconfigures so as to allow the tip to reach the target. As we demonstrate, such internal simulation offers the brain a ‘soft’ mechanism to dynamically link motor redundancy with task oriented constraints ‘at runtime’, hence solving the ‘degrees of freedom problem’ without explicit kinematic inversion and cost function computation (as in the well known optimal control framework). Further, the function of such computational machinery is not only restricted to shaping motor output during action execution but also to provide the self with information on the feasibility, consequence, understanding and meaning of ‘potential actions’. Running such internal simulation process on an interconnected set of neuronal networks is, in our view, the main function of the body schema. Therefore, the body schema is not a static representation, like the Penfield’s homunculus, but a dynamical systems that generates spatio-temporal patterns both for action generation and active inference. In this sense, simulation using PMP based action networks have a central function in linking higher level cognitive layers (like reasoning, understanding) with lower level execution layers while abstracting the complexity of the body.

The PMP framework derives inspiration from both equilibrium point hypothesis () and the theory of impedance control (). In the classical view of EPH, the attractor dynamics that underlies production of movement is based on the elastic properties of the skeletal neuromuscular system and its ability to store/release mechanical energy. PMP on the other hand posits that this may not be the only possibility. The discovery of motor imagery and the strong similarity of the recorded neural patterns in overt and covert movements, suggests that attractor dynamics and the associated force fields may not be uniquely determined by physical properties of the
neuromuscular system but may arise as well from ‘similar’ neural dynamics due to interaction among brain areas that are active in both situations. This could explain the similarity of real and imagined movements because, although in the latter case the attractor dynamics associated with the neuromuscular system is not operant, the dynamics due to the interaction among other brain areas are still at play. In this sense, PMP extends the EPH viewpoint by positing that cortico-cortical, cortico-subcortical, and cortico-cerebellar circuits associated with synergy formation may also be characterized by similar attractor mechanisms that cooperate in shaping flexible behaviors of the body schema in the context of ever-changing environmental interactions. This is the crucial difference between PMP and EPH i.e PMP is a generalization of EPH in a way to incorporate coordination of both covert and overt actions.

**Purpose of PMP in terms of development of learning and reasoning**

One may wonder what role does the PMP system play in a deliverable related to learning. First of all let us consider the *Emulation theory of representation* (Grush 2004), which is based on the use of forward/inverse models, i.e. neural circuits that model the interaction with the body and environment and are constructed by the brain. The idea is that during overt sensorimotor engagement, these models at the same time provide efferent signals for the neuromuscular system and process efference copies in parallel with the body and environment, in order to provide expectations of the sensory feedback and enhance/process sensory information. These models can also be run off-line in order to produce imagery, estimate outcomes of different actions, and evaluate and develop motor plans. Holland et al (2003), Hesslow (2007), Shanahan (2006, 2005), Gallese et al (2011) among others have investigated in the same direction, arguing in favour of a mechanism of *functional imagination* that allows an embodied agent to simulate its own actions and their sensory consequences internally, with significant benefits from a behavioral point of view. In particular, there are atleast five necessary and sufficient requirements for the implementation of functional imagination:

1) Representation of alternative sensory states;
2) Sensory state based predictions;
3) Goal-driven actions;
4) Evaluation of the degree of success;
5) Action selection before acting.

As depicted in figure PMP networks play a central role not only in action generation, but also in task specific action planning (and compositional actions), motor skill learning, and goal directed reasoning. In following sections, we present initial results of how such networks give rise to competences crucial for the development of Darwin (starting from simple control of a single kinematic chain, full upper body coordination in iCub as in scenario S1, skill learning and tool use (scenario S6), incorporating multiple constraints (S1-S7), performing covert reasoning (S3-S7).
PMP is a task specific model and such networks have to be spontaneously composed based on the nature of the task, the environment in which it has to be performed and the body segment (or tool) chosen for its execution. We believe that runtime creation/modification of such networks is a fundamental operation in motor planning and goal directed action synthesis. While the general principles underlying creation of such networks is described in detail in (Mohan and Morasso, 2011d), here we give a glimpse of how the model is gradually expanded while preserving the modularity, local and distributed nature of computation. Figures 11-13 show the gradual expansion of the PMP relaxation network form a single kinematic chain (arm), to the upper body of iCub (mostly relevant of Darwin), the whole body PMP with scope of extensions to coordinated tools. As seen, the computational model is a fully connected network of nodes either representing forces (shown in pink) or displacements (shown in blue) in different motor spaces (end-effector space, joint space, muscle space, tool space etc). There are two kinds of connections 1) Vertical: between each force and displacement node that describes the elastic causality of the coordinated system (stiffness K and admittance matrices A) and 2) between two different motor spaces that describes the geometric causality of the coordinated system (Jacobian matrices J). While skill learning is the topic of the next section, while just mention here that in the context of PMP, when we learn a motor skill, we basically learn the connecting links in the PMP network associated with the task (i.e vertical links or impedances, horizontal links or Jacobians and the timing). Note that, once we learn to control our own bodies, further learning occurs at the interface between body and the tool/task being learnt. In complex kinematic structures, there are three additional nodes: Ground node that determines the parts of the body schema through which the induced force fields propagate; Sum and Assignment nodes to add or assign displacements and forces to different connecting elements of the kinematic chain.

The activation of the ‘goal’ induces elastic reconfiguration of the network analogous to the reconfiguration of the body of a marionette when pulled by puppeteer (for example, as the finger tip
is pulled to a specific spatial location, through a specific trajectory the rest of the body complies to the pull and takes configurations that allow the finger tip to track the pull). The mechanism is labeled "passive" because the final solution is not actively computed. Instead, it gradually emerges due to dynamics of the interactions between the internal body model the task relevant force fields just 'switched on' by the brain. When motor commands obtained by this process of passive simulation of the body schema are actively fed to the robot, it reproduces the motion. Formally, the motion of the kinematic chain evoked by the activation of a target is equivalent to integrating non-linear differential equations that, in the simplest case in which takes the following form: 

\[ x = J A J^T K (x_r - x) \].

As a simple example, if we deactivate the left arm and the waist space in figure and enter the network at the right arm end effector \( dx \) and exit at right arm joint space \( dq \), we get the following rule for computing incremental joint angles: 

\[ dq = A R J^T K_R dx \].

The rules become more complex as additional motor spaces participate in the PMP relaxation. We refer the reader to Mohan and Morasso 2011d for a detailed formal analysis.

![Figure 11. Basic kinematic network that implements the Passive Motion Paradigm for a simple kinematic chain (as the arm). In this simple case, the network is grouped into two motor spaces (extrinsic or end effector space and intrinsic or arm joint space). Each motor space consists of a generalized displacement node (blue) and a generalized force node (pink). Vertical connections (purple) denote impedances (K: Stiffness, A: Admittance) in the respective motor spaces and horizontal connections denote the geometric relation between the two motor spaces represented by the Jacobian (Green). The goal induces a force field that causes incremental elastic configurations in the network analogous to the coordination of a marionette with attached strings. The network also includes a time base generator which endows the system with terminal attractor dynamics: this means that equilibrium is not achieved asymptotically but in finite time. External and internal constraints (represented as other task-dependent force/torque fields) bias the path to equilibrium in order to take into account suitable 'penalty functions'. This is a multi-referential system of action representation and synergy formation, which integrates a Forward and an Inverse Internal Model.

Multiple task specific constraints can be added in a task specific fashion by switching on other constraint related force fields. A constraint in the extrinsic space could be an obstacle to avoid, an appropriate hand pose with which to reach an object so as to allow further manipulation actions to be performed (like grasp or push or other necessary actions using MECCANO toys). In the intrinsic space a constraint could take into account the limited range of motion of a joint, the saturation power or torque of an actuator etc. Figure 14 shows a composite PMP network for the right arm kinematic chain, for reaching an object (Goal) with an appropriate wrist orientation/ hand pose to support further manipulations (constraint 1) and generating a solution such that the joint

![Figure 11 Diagram](image-url)
Figure 12. Bimanual coordination task of reaching and picking up two objects at the same time. This scenario integrates the ‘identify-localize-reach-grasp-transport’ loop (mainly S1). Panel A: PMP network for the upper body with two target goals and a single time base generator. Similar to figure the expanded network now includes three modules: 1) Right arm, 2) Left arm, 3) Waist. The dimensionality of $J_R$ & $J_L$ is $3 \times 10$ (this includes the 7 DoF’s of the arms and the 3 DoF’s of the waist). The dimensionality of $A_j$ is
7×7 and of $A_T$ is 3×3. The three sub-networks interact through a pair of nodes (‘assignment’ and ‘sum’) that allow the spread of the goal-related activation patterns. Panels B & C show the initial and the final posture of the robot and the two target objects. Panels D & E show the trajectories of the two end-effectors and the corresponding speed profiles (together with the output $\Gamma(t)$ of the time base generator). Panel F clarifies the intrinsic degrees of freedom in the right arm-torso chain. Panel G shows the time course of the right-arm joint rotation patterns: J0-J2: joint angles of the Waist (yaw, roll, pitch); J3-J9: joint angles of the Right Arm (shoulder pitch/yaw/roll; elbow flexion/extension; wrist pronation/supination pitch/yaw).

Angles are well within the permitted range of motion (constraint 2). Hence, in the PMP network of 14 there are three weighted, superimposed force fields that modulate the spatiotemporal behavior of the system: 1) the end-effector field (to reach the target); 2) the wrist field (to achieve the specified hand pose); 3) the force field in joint space for joint limit avoidance. Note that the same timing signal $\Gamma(t)$ synchronizes all the three relaxation processes. Figure 5B shows results of iCub performing different manipulation tasks driven by such a network. More importantly, this aspect is related to synthesis of ‘compositional actions’ (section).

Figure 13. Full kinematic network of the iCub robot (53 DoFs, excluding the DoFs of tools).

In the previous section we outlined the core of the action generation mechanism in Darwin that is responsible for coordinating the complex body of iCub to initiate exploration, imitation, reasoning and understanding. In this section, we outline the approach for skill learning (and tool use) a central functional requirement in Darwin robots, using this computational framework as a core module. As has been pointed out in section, the idea behind the skill learning system is to exploit simultaneously multiple learning streams like 1) motor babbling (self exploration), 2) imitative action learning (social
Figure 14 A. Composite PMP network with three force fields applied to the right arm of iCub: a field $F_1$ that identifies the desired position of the hand/fingertip (Goal); a field $F_2$ that helps achieving a desired pose of the hand via an attractor applied to the wrist (Constraint 1). Here $J_{RW}$ is the Jacobian matrix of the subset of the kinematic chain, up to the wrist; an elastic force field $F_3$ in the joint space for generating a solution such that the joint angles are well within the permitted range of motion (constraint 2). Note that the same timing signal synchronizes all three relaxation processes, hence allowing the hand to reach the target with a specific pose and posture.

Figure 14 B. shows three examples of iCub performing manipulation tasks driven by the composite PMP net of figure 5A. Note that in all these cases, reaching the goal object with specified hand pose is obligatory for successful realization of the goal.
interaction) and 3) reuse of past motor experiences. The purpose of this combination is to both counter the curse of dimensionality and give rise to sensorimotor knowledge endowed with seamless compositionality, generalization capability and body-effectors/task independence. The central objectives of the skill learning system are pictorially depicted in figure 15.

In the context of PMP, when we learn a motor skill, we basically learn the connecting links in the PMP network associated with the task (i.e. vertical links or impedances, horizontal links or Jacobians and the timing of the time base generators). We will describe central ideas using a preliminary scenario of tool use: learning to bimanually control a toy crane in order to position its magnetized tip at a goal target. In general, while learning to control the toy crane, iCub has to learn:

1. the appropriate impedance and timing to execute the required ‘spatiotemporal’ trajectories using the body+tool chain (for example, performing synchronized quasi circular trajectories with both hands while turning the toy crane) and
2. while performing such coordinated movements with the tool, learn the Jacobians that map the relationship between the movements of the body effectors and the corresponding consequence on the tool effector (the magnetized tip).
3. The third issue is of course related to using this learnt knowledge to generate ‘goal directed’ body+tool movements (given a goal to reach/pick up an otherwise ‘unreachable’ environmental object using the toy crane).

The fourth issue of opportunistically using this tool as a ‘sub goal’ in a broader context through reasoning is still under development.

Note that using a toy crane is a task that goes beyond reach-grasp movements as it not only requires iCub to reach (and grasp) the tool but also perform coordinated spatiotemporal movements with the tool (both during exploration and performing goal directed movements using the tool). Part of the information as to what kind of movements can be performed with the tool can be acquired by observing a teachers demonstration. The teacher’s demonstration basically constrains the space of explorative actions when iCub practices with the new toy to learn the consequences of its actions. The basic PMP system on the iCub is presently being extended to incorporate these capabilities with preliminary results we present. With the help of Figure 16, we outline central features of the extended skill learning architecture.
Learning through Imitation, Exploration and Motor Imagery: Three streams of learning i.e. learning through teacher’s demonstration (information flow in black arrow), learning through physical interaction (blue arrow) and learning through motor imagery (loop 1-5) are integrated into the architecture. The imitation loop initiates with the teachers demonstration and ends with iCub reproducing the observed action. The motor imagery loop is a sub part of the imitation loop, the only difference being that the motor commands synthesized by the PMP are not transmitted to the actuators. Instead, the forward model output is used to close the learning loop. This loop hence allows iCub to internally simulate a range of motor actions and only execute the ones that have high performance score ‘R’.

![Diagram of Motor Skill learning and Action generation architecture of iCub](image-url)

From Trajectory to Shape, towards ‘Context Independent’ motor knowledge: Most skilled actions involve synthesis of spatio-temporal trajectories of varying complexity. A central feature in our architecture is the departure from the well known concept of ‘trajectory formation’ by introduction of the notion of ‘Shape’ in the domain of movement. The idea is to conduct motor learning at an abstract level and thus speed up learning by exploiting the power of ‘compositionality’ and motor knowledge ‘reuse’. In general, a trajectory may be thought as a sequence of points in space, from a starting position to an ending position. ‘Shape’ is a more abstract description of a trajectory, which captures only the critical events in it. By extracting the ‘shape’ of a trajectory, it is possible to liberate the trajectory from task specific details like scale, location, coordinate frames and body effectors that underlie its creation and make it ‘context independent’. Using Catastrophe theory (Thom 1975), Chakravarthy et al (2003) have derived a set of 12 primitive shape features (figure 6, bottom right panel) sufficient to describe the shape of any trajectory in general. As an example, the critical events in a trajectory like ‘U’ is the presence of a minima (or Bump ‘B’ critical point) in between two end points (‘E’). Thus, the shape is
represented as a graph ‘E-B-E’ (see figure 16). If the ‘U’ was drawn on a paper or if someone runs a ‘U’ in a playground, the shape representation is ‘invariant’ (there is always a minima in between two end points). More complex shapes can be described as ‘combinations’ of the basic primitives, like a circular trajectory is a composition of 4 bumps. In short, using the shape extraction system it is possible to move from the visual observation of the end effector trajectory of the teacher to its more abstract ‘shape’ representation.

Figure 17: Top left panel describes the task: i.e bimanually coordinating the toy crane to pick up otherwise unreachable objects in the environment. Top right Panel shows the PMP network for goal directed body+ toy crane coordination. Note that both the body and the tool are represented in the same fashion and as far as goal directed coordination is concerned the tool is as much a part of the body, actively exchanging information “bidirectionally” with each other in order to realize the goal at hand. Middle panel shows the how the knowledge of past motor experience can be recycled to learn new skills quickly:in other words how learning one motor skill in a abstract fashion can evoke the implicit potential to perceive, execute and imitate several other skill that share a similar structure. Bottom panel shows preliminary results of goal directed coordination to the toy carane to reach targets otherwise unreachable with either end effectors of iCub.
**Imposing ‘context’ while creating the Motor Goal:** The extracted ‘shape’ representation may be thought of as an ‘abstract’ visual goal created by iCub after perceiving the teacher’s demonstration. To facilitate any action generation/learning to begin, this ‘visual’ goal must be transformed into an appropriate ‘motor’ goal in iCub’s egocentric space. To achieve this, we have to transform the location of the shape critical point computed in the image planes of the two cameras \((U_{left}, V_{left}, U_{right}, V_{right})\) into corresponding point in the iCub’s egocentric space \((x,y,z)\) through a process of 3D reconstruction (see figure 6, top left box). Of course the ‘shape’ is conserved by this transformation i.e a bump still remains a bump, a cross is still a cross in any coordinate frame. Reconstruction is achieved using Direct Linear Transform (Shapiro, 1978) based stereo camera calibration and 3D reconstruction system already functional in iCub (implementation details of this technique are summarized in the appendix of Mohan et al 2011b). At this point, other task related constraints like the scale of the shape, end effector/body chain performing the action can be added to the goal description. So the motor goal for iCub, is an abstract shape representation of the teachers movement (transformed into the egocentric space) and other task related parameters that needs to be considered while generating the motor action. An example of a motor goal description is like: ‘use the left arm-torso chain coupled to the toy crane, generate a trajectory that starts from point 1, ends at point 2 and has a ‘bump’ at point 3 (and observe the consequence through visual and proprioceptive information).

‘Virtual trajectories’-Motor equivalent action representation: The motor goal basically consists of a discrete set of shape critical points (their spatial location in iCub’s ego centric space and type), that describe in abstract terms the ‘shape’ of the spatiotemporal trajectory that iCub must now generate (with the task relevant body chain). Given a set of points in space an infinite number of trajectories can be shaped through them. How can iCub learn to synthesize a continuous trajectory similar to the teacher’s demonstration using a discrete set of shape descriptors in the Motor goal? The virtual trajectory generation system (VTGS) performs this inverse operation. It transforms the discrete shape representation (in the motor goal) into a continuous set of equilibrium points that act as moving point attractor to the PMP system. **Intuitively, one may imagine that the puppeteer is pulling the strings in a specific fashion in order to generate specific spatiotemporal trajectories with the puppet.** VTGS preserves the same ‘force field’ based structure as in PMP. As seen in figure 6, the VTGS depends on two parameters virtual stiffness and timing of the time base generator, different values of which generate different moving point attractors. Goal of learning is to learn the correct values of \(K\) and TBG such that the such that the ‘Shape’ of the spatio temporal trajectory created by iCub correlates with the shape description in motor goal.

So ‘how difficult and how long’ does it take to learn these parameters given the demonstration of a specific movement by the teacher? It is here we reap the advantage of moving from ‘trajectory’ to ‘shape’, since compositionality in the domain of shapes can be exploited to speed up learning. In other words, the amount of exploration in the space of ‘\(K\)‘ and \(\gamma\) is constrained by the fact that once iCub learns to generate the 12 movement shape primitives, any motion trajectory can be expressed as a composition of these primitive features. The main idea is that since more complex trajectories can be ‘decomposed’ into combinations of these primitive shapes, inversely the actions needed to synthesize them can ‘composed’ using combinations of the corresponding ‘learnt’ primitive actions. Regarding learning the primitives, it has been demonstrated in (Mohan et al, 2011), that they can be learnt very quickly by just exploring the space of the virtual stiffness ‘\(K\)’ in a finite range of 1-10 , followed by an evaluation of how closely the shape of the synthesized trajectory (using equation 4) matches the shape described in the goal. Thus effort in terms of motor exploration is required during the initial phases to learn the basics (i.e primitives). During the synthesis of more complex spatiotemporal trajectories, composition and recycling of previous
knowledge takes the front stage (considering that the correct parameters to generate the primitives already exist in the shape library). Finally, we note that ‘virtual trajectories’ must not be interpreted as the real trajectories generated by iCub. Instead, the evolving virtual trajectory acts as moving point attractor to the PMP system that in turn generates the motor commands necessary for iCub to actually execute the motion trajectory (it observed).

**Using past motor ‘experience’ to generate virtual trajectories on the fly:** When iCub learnt to draw trajectories like ‘U’, ‘C’ etc (Mohan et al, 2011b), it acquired the correct parameters (K and γ) to synthesize virtual trajectories for shapes that result in ‘Bump’ critical points. When the teacher demonstrates iCub to bimanually steer the toy crane by performing quasi circular trajectories, the movement ‘shape’ of the teachers ‘effectors’ gives rise to ‘bump’ critical points, which iCub already knows to generate, from its previous drawing experience. Using the previously learnt parameters of K and γ from the shape library, iCub is able to instantaneously generate virtual trajectories (or attractors) that feed the PMP network of the iCub upper body. Here, we reap the benefit of moving from ‘trajectory’ to ‘shape’ and learning actions in a ‘context independent’ fashion thus allowing past experience to be exploited in new contexts. At the same time, only being able to synthesize a ‘virtual trajectory’ is not sufficient. What is needed is a system that transforms the ‘virtual trajectory’ into motor commands for the actuators, taking into account task specific

![Figure 18. Row 1. Tool use under Normal Conditions (Tool is compliant A_T=0.1, Both arms are equally active to generate force K_e=0.01); Row 2: Tool use when right arm functionality is compromised (Tool is compliant A_T=0.1, but K_e, Right Arm=0.001, K_e, Left Arm=0.01); Row 3. What if: ‘toy crane’ is replaced by a ‘human’ who chooses not to ‘Comply’?](image-url)
constraints and redundancy of the system (body-tool network) that is generating the action. Further it is necessary to learn the consequences of the generated action in this new ‘context’. For this we have to rely on the PMP system that comes next in the information flow.

**From virtual trajectory to Motor Commands using PMP: Linking redundancy to Task dynamics, Timing & synchronization:** The PMP system transforms every point in the virtual trajectory into motor commands in the intrinsic space (upper body chain), hence enabling iCub to mimic the teacher’s action of bimanually steering the toy crane. Of course, this is just the starting point. iCub has to now learn the consequence and utility of the action in this new context (from drawing a ‘U’ shape, to using the toy crane). As the virtual trajectory pulls the relevant end effector in a specific fashion, the rest of the body (arm and waist joints) elastically reconfigure to allow the end effector track the evolving virtual trajectory. When motor commands synthesized by this process are actively fed to the robot, it reproduces the movement, hence enabling iCub to maneuver the toy crane as demonstrated by the teacher. These coordinated movements of iCub with the toy crane now generate sequences of sensorimotor data:

1) the instantaneous position of the two hands \( Q \in (x_R, y_R, z_R, x_L, y_L, z_L) \) coming from proprioception (and cross-validated by forward model output of PMP i.e position node in end effector space)
2) the resulting consequence i.e the location of the tool effector \( X : (x, y, z)_{\text{Tool}} \) perceived through vision and reconstructed to Cartesian space (using the same technique to reconstruct teachers movement)

*This data is used to train a neural network that sub symbolically represents the relation between the motion of the body and its consequence (Mohan and Morasso 2011d).*

**Generating goal directed “body + tool” actions:** Once the tool Jacobians \( J_T \) are learnt by iCub, the PMP network that includes both ‘body and tool’ as a fully connected chain can be generated. Figure 17 top panel shows the basic “upper body +toy crane” network to generate goal directed actions with the tool. During goal directed movements with the toy crane, the goal now acts on the ‘tool effector’ which is the most distal part of the PMP chain. The pull of the goal acting on the tool tip is incrementally circulated to the proximal spaces (end effector, joints etc) according to information flow in figure 17. An interesting point to observe is that PMP framework does not make any special distinction between the ‘body’ and a ‘tool’. The tool space is represented exactly in the same manner as any other motor space and during coordination the body and the tool act as one cohesive unit (to realize a goal). The process whereby a tool becomes an extension of the hand to perform a specific task can be related to the flexible view of body schema offered by Head and Holmes (1911). Moving to row 1 of figure 18, we refer to this as the normal condition i.e., a situation where both the tool is compliant \( (A_T = 0.01) \) and both arms equally functional to generate force \( (K_e = 0.01) \) for both arms. As seen in panel 1, the tool angle (green) faithfully tracks the moving point attractor or the goal (red). Panel B-D; show the resulting trajectory of the tool tip from the initial condition to the goal, the X-Y components of the force exerted by the two hands and the resulting torque in the tool space. Since the tool torque is bell shaped, the tool velocity is also bell shaped [6]. Now, let us consider that for some reason (injury, etc), that the right arm functionality is compromised \( (K_e = 0.0.01 \text{ for the right hand and } K_v = 0.0.01 \text{ for the left hand}) \). Panels 5-8 of figure 2 show the resulting behavior. Observe that the actual tool trajectory still follows the goal (panel 5) and the tool tip almost reaches the target (panel 6), despite the fact that right arm functionality was compromised ‘ten times’ that of the left arm. The reason is that the left arm covers up the disability of the right arm by exerting a greater force (Panel 7), compared to the normal condition (Panel 3). *No learning is needed to accomplish this; it is in fact the property of the attractor dynamics of the ‘elastic’ PMP system to take into account such situations (be it injury,
external disturbance, etc) and yet do the best in achieving the goal. Also note that the tool torque is almost bell shaped, like the normal situation.

Now let us consider the scenario where the ‘Toy crane’ was an animate organism instead. The only difference is that instead of being a lifeless ‘Toy Crane’, the human has a choice to ‘comply’ to the external force exerted by some other system, by controlling one’s own admittance. A scenario where the human chooses ‘not’ to comply is shown in column 3 of figure 2 (where \( A_{\tau}=0.01 \)). Note that even though the external system connected to the body is not ‘compliant’, still an attempt to achieve the goal is done by the PMP network, by increasing the force exerted by both the arms (Panel 11), rather unsuccessfully. But still, the network does its best to ‘persuade’. Also note that as a consequence of not realizing the goal, there is energy remaining in the network (since energy is an inner product of force and displacement, both of which are not zero in this case).

In a seminal paper, Umiltà et al (2009) have shown that the essence of tool use lies in the capacity to transfer proximal goals to distal goals. Recording from monkeys trained to use pliers to grasp otherwise unreachable food rewards, they demonstrated that the end effect of tool use training was the transfer of the temporal discharge pattern that controls ‘hand grasping’ (area F5) to the tool, as if the tool was the hand of the monkey and its tips were the monkey’s fingers. This of course is reminiscent of the results of Iriki and colleagues (Maravita et al, 2004, Iriki et al, 2008), who showed that, with practice, a rake becomes a part of the acting monkey body schema.

However, what Umiltà et al demonstrated was that in addition to being incorporated into the body schema, the tool, after learning, is coded in the motor system as if it were an artificial hand able to interact with the external objects, exactly as the natural hand is able to do. In the PMP network for coordinating the toy crane (figure 2), as the magnetized tip is being pulled towards the goal target, iCub’s end effectors are simultaneously being pulled towards the required positions so as to allow the tool tip to reach the goal. These positions are the goals for the end effector space. As a consequence, the joints are concurrently pulled so as to allow the end effectors to reach the position that allows the tool tip to reach the goal. These are the goals for the intrinsic space. If motor commands derived through this incremental internal simulation of action are transmitted to the robot, it will reproduce the motion, hence allowing iCub to perform goal directed movements using the ‘body + toy crane’ network. It is this kind of goal-centered functional organization of cortical motor areas for which Umiltà et al provide evidence through their tool use experiments with monkeys.

**Summary**

In the skill learning section, we presented how the basic PMP framework can be extended to support experiments related to motor skill learning, tool use and imitation in embodied robots. We outlined a scheme through which both observing a ‘conspecific’ as well as previously acquired motor knowledge (stored in an abstract manner) can speed up the acquisition of a new motor skill. To avoid open ended motor exploration, it is important to ‘combine and exploit’ multiple learning streams mainly imitation, physical interaction and motor imagery into the skill learning architecture. In the demonstrated example, while the teachers demonstration showed iCub the kind of spatiotemporal trajectories it should perform on the tool, iCub’s past experience of learning to draw (and the compositionality in the domain of shapes) gave iCub the correct parameters to generate the required spatiotemporal trajectories using the ‘body+tool’ network. Of course, in addition iCub had to learn the context specific consequences (Tool Jacobian), to complete the PMP network to perform goal directed actions with the new toy. Note that the learnt tool Jacobian is further represented in a subsymbolic ‘distributed’ fashion using neural networks. At the same time, through the PMP relaxation there is a way to systematically go down to the directly controlled
elements of the body (actuators in the robot) both during exploration and goal directed action. In this sense, our approach is quite different from other known attempts of tool use in robotics like those of Stoychev (2008), that start with a predefined set of actions (extend arm 2 inches, 5 inches, forward, backward, right and left), create a look table of the observations and conduct iterations of greedy heuristic search in the look up table, to obtain goal oriented behavior.

5.4 Learning Affordances

Affordances are the seeds of action. Being able to identify and exploit them opportunistically in the ‘context’ of a goal is a sign of cognition. Being able to do this in the mind by performing virtual actions, further allows an agent to simulate ‘what additional affordances’ it can create in its world as a consequence of its actions, hence most importantly enabling it to reason about how the world must ‘change’ such that it becomes a little bit more ‘affording’ towards realization of its internal goals (Mohan et al, 2011a). Of course, this requires the perceptual, motor and conceptual systems to work as an integrated, synchronized system. Further, we note that affordances are not intrinsic properties of objects, but rather relational properties: an object may provide different affordances depending on the features of the organism body and the context. Internalizing this relational property is the real challenge posed by affordance learning. Jeannerod (1994) and Arbib (1997) proposed that objects tend to evoke actions which are appropriate to them and suggested that this process involves motor representations and their distal goals. Many behavioural and brain imaging studies on humans and animals have provided evidence showing that observation of objects activates their associated affordances. Rizzolatti, Fogassi, and Gallese (1997) conducted a study with monkeys illustrating that the sight of objects tends to automatically evoke the activation of canonical neurons in the premotor cortex believed to underlie action preparation (a simple example is precision or power grip actions). However, the extent to which this activation occurs in an automatic bottom-up way or is modulated by the task itself is subject of debate (Borghi et al., 2007).

In general, an affordance can be described as a set of properties of an object which suggest possible actions and uses to an organism mainly,

1. The ways to couple one’s body to the object, grasp it etc (this is related to both the perceptual properties of the object and the action related constraints on the body, for example grasping a toy crane);
2. The ways in which it can be used once coupling is established (this relates to the motor skill available, for example, knowing to use a toy crane once it is grasped);
3. What are the consequences of the actions initiated?
4. The ways in which any purpose is served in the context of the goal at hand?

All four components listed above need learning (both explorative and interactive). This enters the core of the ‘goal-affordance-skill-value’ loop as formulated in section.

The theoretical framework for affordance learning being developed by KCL is based on well known principles related to known organization in the brain:
(a) the dual-route organization of the visual system relying upon a ventral and a dorsal neural pathway (where and what);
(b) action selection on the basis of prefrontal cortex inputs (PFC) and value systems (what is useful);
(c) the selection of actions on the basis of a competition between different affordances based on the bias of PFC;
(d) the capability to trigger internal simulations of actions and virtually perceive their consequence situations (as in the forward/inverse models initiated by PMP in the previous section).

Even though the approach is brain inspired, we do believe that forming “functional interfaces” between work done by other partners using other computational methods (nevertheless providing necessary information at the interface) is certainly feasible. While a detailed presentation of the neural architecture under development is out of scope for this deliverable, a concise summary is provided in the connectivity table 2 that summarizes the main modules being implemented, the functionality performed, interfaces i.e. inputs and outputs and the associated learning rules, and their relation with the associated brain areas. The neural architecture is also correlated approximately with the known functional architecture in the brain.

<table>
<thead>
<tr>
<th>Module name</th>
<th>Functionality</th>
<th>Modules feeding input</th>
<th>Learning Rules for Input synapses</th>
<th>Output modules</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>Edge detection</td>
<td>Image</td>
<td>Hard-wired</td>
<td>V2, V4</td>
</tr>
<tr>
<td>V2</td>
<td>Angle calculation</td>
<td>V1</td>
<td>SOM/LISSOM</td>
<td>V4, IT</td>
</tr>
<tr>
<td>V4</td>
<td>Partial boundaries</td>
<td>a) V1, b) V2</td>
<td>a) SOM/LISSOM b) SOM/LISSOM</td>
<td>IT</td>
</tr>
<tr>
<td></td>
<td>a) IT (object identity) b) IT (arm/hand identity)</td>
<td>a) object identity a) V2, V4, AIP (object shape), PPC (object location) b) V2, V4, AIP (hand shape), PPC (arm present position vector)</td>
<td>a) SOM/LISSOM (V2, V4), Hebbian (AIP, PPC) b) SOM/LISSOM (V2, V4), Hebbian (AIP, PPC)</td>
<td>a) AIP (object shape), PPC (object location), PFC b) AIP (hand shape), PPC (arm present position vector), PFC</td>
</tr>
<tr>
<td></td>
<td>a) AIP (object shape) b) AIP (hand shape) c) AIP (affordances)</td>
<td>a) Object shape a) V1, IT (object identity) b) V1, IT (arm/hand identity) c) AIP (object identity)</td>
<td>a) Hardwired (V1), Hebbian (IT) b) Hardwired (V1), Hebbian (IT) c) Hardwired (AIP)</td>
<td>a) AIP (affordances) b) DVv c) DVd</td>
</tr>
<tr>
<td></td>
<td>a) PPC (object location) b) PPC (arm present position vector)</td>
<td>a) target location a) V1, IT (object identity) b) V1, IT (arm/hand identity)</td>
<td>a) Hardwired (V1), Hebbian (IT) b) Hardwired (V1), Hebbian (IT)</td>
<td>a) DVd b) DVd</td>
</tr>
<tr>
<td>DVv</td>
<td>Difference vector (ventral)</td>
<td>AIP (affordances), AIP (hand shape)</td>
<td>Hardwired</td>
<td>PMv multiplied by GO signal</td>
</tr>
<tr>
<td>DVd</td>
<td>Difference vector (dorsal)</td>
<td>PPC (object location), PPC (arm present position vector)</td>
<td>Hardwired</td>
<td>PMd multiplied by GO signal</td>
</tr>
<tr>
<td>PMv</td>
<td>Preshaping motor commands</td>
<td>PFC, DVv, Cerebellum, SI (finger pressure sensor feedback)</td>
<td>Hardwired</td>
<td>Robot hand</td>
</tr>
<tr>
<td>PMd</td>
<td>Arm joint angle commands</td>
<td>PFC, Cerebellum,</td>
<td>Hardwired</td>
<td>Robot arm</td>
</tr>
<tr>
<td></td>
<td>DVd, SI (joint sensor feedback)</td>
<td>SOM (IT, PPC), hardwired (external command, AAC)</td>
<td>PMd, PMv</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>PFC</td>
<td>Goals, bias, working memory</td>
<td>External command, IT (object/arm/hand identities), AAC (reset), PPC (object/arm location vectors)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO</td>
<td>Gates DV to PM, scales movement commands</td>
<td>Cortex (not included here)</td>
<td>Hardwired</td>
<td>PMd, PMv</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>Fine motor control of movement, online correction of movement</td>
<td>Most likely receiving input from arm joint sensors (not included in the architecture)</td>
<td>Hardwired</td>
<td>PMd, PMv</td>
</tr>
<tr>
<td>SI</td>
<td>feedback from pressure sensors and joint sensors</td>
<td>Sensors on i-Cub fingers and arm joints</td>
<td>Hardwired</td>
<td>PMv, PMd</td>
</tr>
<tr>
<td>External Commands</td>
<td>Set the order of goals and how they will be executed by PMv and PMd</td>
<td>External commands (user specified)</td>
<td>Hardwired</td>
<td>PFC</td>
</tr>
</tbody>
</table>

Table 2. Brain guided neural architecture for learning affordances in Darwin: Key functional (input-output) interfaces and learning methods deployed. Expansion of acronyms used in the table: AIP: anterior intraparietal area; FEF: frontal eye field; PFC: Prefrontal cortex; PPC: posterior parietal cortex; DVv: ventral difference vector; DVd: dorsal DV; PMv: ventral premotor cortex; PMd: dorsal PM; AAC: anterior cingulated; IT: infero-temporal cortex; SI: primary somatosensory cortex.

### 5.5 Acquisition of ‘Spatial’ and ‘Causal’ relations

Exploring and learning about properties of objects, how to handle them and use them also presents a parallel opportunity to understand how the world usually works. This goes into the problem of development of core knowledge systems in Darwin. Based on research on infants and primates, Speke and Kinzler (2007) list at least four core knowledge systems of central importance: spatial, physical, psychological and numeric. The four systems serve to represent knowledge about inanimate objects and their mechanical interactions, knowledge of spatial relations, direction, shape and geometric relationships, knowledge about other agents and their goal-directed actions, knowledge about sets and numerical relationships of ordering.

As far as the functional capabilities of Darwin are concerned learning about spatial and physical causality assumes significance. Several experimental scenarios in section have been formulated to aid learning of the underlying spatial and physical causality. In general, all actions on objects involve motion of manipulated objects and the body through space (and time); learning to generate purposeful actions requires the robot to direct its eyes, hands, and body to the right places. While learning to act, it is necessary to endow robots the capability to exploit the spatial information in...
their actions and perceptions to abstract the general spatial and physical properties underlying their interactions. \textit{How recurring patterns of sensorimotor experience are abstracted to give rise to core knowledge systems is the challenge for learning.} Importantly, endowing Darwin systems with an abstract sense of ‘direction’, ‘distance’, ‘shape’, ‘force’, counterforce/obstacle’, ‘contact’ and ‘containment’ in our opinion is necessary to ensure that it comes to its full potential in the future. These factors come to play in any complex scenario involving assembly-disassembly.

\textbf{Internalization of “Shape” knowledge:} “Shapes” are ubiquitous in our perceptions and actions. All basic ‘sensorimotor’ interactions require ‘shaping’ one’s body to the shape of the world with which we are interacting (be it a monkey clinging to a branch of a tree, a humanoid grasping a ball or a couple dancing). Simply stated, seeing and doing meet at the boundaries of a shape. It is widely agreed that objects are best recognized by their shapes, beginning in early childhood (Smith et al, 2002). Object shape is processed by dedicated regions in the lateral occipital and temporal cortex of the brain (Grill-Spector, Golarai, & Gabrieli, 2008; Reddy & Kanwisher, 2006). These regions respond to the shapes both of 3D objects and of 2D forms, in humans (Kourtzi & Kanwisher, 2001) and in nonhuman primates (Kriegeskorte et al., 2008; Tanaka, 1996; Yamane, Carlson, Bowman, Wang, & Connor, 2008), further suggesting that common cognitive mechanisms underlie perception of the shapes of 2D visual forms and of 3D manipulable objects. Surprisingly, it is not easy to give a precise mathematical or quantitative definition of ‘shape’ or even express it in measurable quantities like length, angles or topological structures. In general terms, shape is the core information in any object/action that survives the effects of changes in location, scale, orientation, end-effectors/bodies used in its creation, noise, and even minor structural “injury”. We posit that it is this invariance that makes ‘shape’ a crucial piece of information for sensorimotor interaction.

A first hand attempt to “internalize” and “exploit” shape knowledge has been attempted in the context of motor skill learning (for details see Mohan et al 2011a, d). The core idea was the introduction of the notion ‘Shape’ in the domain of movement, thus strongly proposing the need to depart from the well known notion of ‘movement trajectory’ to the idea ‘movement shape’. As a simple example, consider actions like turning a steering wheel, uncorking, unwinding, screwing, cycling among others, which result in formation of circular movements in different contexts with different body effectors. Even though the precise spatiotemporal trajectories are different, the shape representation is “invariant”. The notion of ‘circularity’ is a common denominator in all these actions. So can a baby humanoid learning to draw a circle, at the same time also learn something about ‘Circularity’? The general idea was that rather than focusing on movement trajectories, if we teach a humanoid to ‘perceive and synthesize’ shapes instead, we can endow it with the graphical grammar to compose a wide range of movements based on ‘context’, effectively recycling past motor experiences. An example was presented in teaching iCub to use the toy crane (Mohan et al, 2011 d).
However the results are far reaching. A wide range of common human actions, characters’ in scripts of several languages result in trajectories that ultimately can be represented a composition of a few simple ‘shape’ features (see figure 19). Most important among them are line, bump and cusp critical points that can be easily ‘learnt, represented and synthesized’ in formal terms (Mohan et al 2011a). In sum, what we have shown is that by internalizing “shape knowledge” it becomes possible to endow in dexterous robots the powerful capability to ‘compose, recycle’ the previously acquired motor knowledge to swiftly learn wide range of other motor skills. Further, does the capability to perceive the underlying structure in motor actions and ‘spontaneously imitate’ someone performing them with a fair enough ‘first prototype’ becomes possible because the ‘seeds’ already exist in the form of abstract motor knowledge (learnt and stored previously). Do we prefer design tools that ‘conform’ to these specific movement shapes in the extrinsic space hence endowing them with a measure of ‘user friendliness’ in terms minimizing the explorative efforts by recycling past experience? In sum, we suggest that an unified treatment of the dual operations of shape perception/synthesis is critical to better understand the perception-action loop, how we recycle past sensorimotor knowledge, and why we design tools the way we do (taking into account how ‘user friendly’ it is). Future research in this direction will be aimed at answering/validating these questions.

**To learn to push:** In relation to direction and distance, it is well known that both humans and animals represent distance and directional relationships (Doeller & Burgess, 2008; Lever, Wills, Cacucci, Burgess, & O’Keefe, 2002). The existence and properties of these representations are revealed when animals lose their orientation and must draw on memory for the positions of these surfaces to reorient themselves (Cheng & Gallistel, 1984). At the same time knowledge of force and contact is a basic requirement for dexterous manipulation. Most of the time such information
fills in for vision at the edge of physical interactions (consider a ‘screwing’ task where hand occludes the vision of the screw or inserting a key into the key hole to open a door). It is important to realize the need for not just ‘sensing’ but dynamically “filling in missing information” in ways that help in progressing

**To learn to “Push”**

- Contact is necessary to Push
- Object properties influence ‘Pushability’
- Goal directed pushing requires an active sense of directionality
- Pushing gives rise to path of motion from source to goal (source-path-goal schema)
- Has origins and intensities
- There are Counter forces and Diversion
- Pushing supports ‘grasping’

Figure 20. To learn to push and “what” to learn while pushing

incrementally towards the pursued goal. A well investigated scenario from animal and infant cognition that directly relates to knowledge of directionality, contact, force, counterforce etc is the task of learning to ‘push’ (or pull) various objects sometimes using tools in a way to resize otherwise unrealizable goals. As seen in figure there is a lot to learn in such scenarios about the basic spatial and physical causality.

Some preliminary attempts to learn to “push” have been recently attempted in robotics recently (Dillmann et al, Mohan et al 2011). Even though interesting attempts, there is still a lack of capturing the underlying physical causality of pushing actions in an “abstract” fashion. The former approach employs a standard backpropogation network to the learn poking behavior in order to serve as a support action for grasping. However, the production of novel behaviors is constrained to making generalizations over the learnt input-output data set (using the standard Levenberg Marquart algorithm). The latter approach is based on growing sensorimotor maps (Toussaint et al, 2006) and neural fields (). Sensorimotor exploration, self organization, field computing and value dependent learning assumes significance here. This architecture was further extended in Mohan et al 2011c further with the capability to update the internal model based on ‘new variables’ introduced in the world (like a trap, never
Figure 21. Learning to push: initial trials. Top left panel shows iCub pushing a big green cylinder and perceiving its consequence. Top right panel shows the representation i.e a sensory motor map, the sensory space learning to represent for average displacement of a particular object per unit force applied and motor space learning the distributed coding of direction (bottom left panel). The bidirectional interactions between sensory and motor spaces basically allow iCub to perform goal directed pushing (bottom right panel).

encountered before) based on the contradiction between top down prediction and bottom up sensation (termed as cognitive dissonance) and further learning new value fields by exploration to counter the effect of this new variable introduced in the world. First steps are being taken to further extend the basic model further to teach iCub to push objects in a goal directed fashion, yet retaining a sense of ‘independence’ in terms of spatial representation, objects of action, causes of action (hand, tool etc hence preserving in a motor equivalence). To achieve this iCub is allowed to practice pushing with different objects (cylinders, cubes, balls) while visually observing the consequence. Taking inspiration from the work on ‘shapes’, the trick is to eliminate task specific details (like object locations, end effectors used, amount of force applied) that basically are ‘contextual’ information. The idea is to just code for “invariants” using the same sensorimotor map used previously (Mohan et al, 2011c). Surprisingly, it seems that if the sensory space codes for ‘average displacement per unit force’, the action space has to have a distributed coding of direction (for which there is resounding evidence from neuroscience). Both these information can be easily learnt through experience of pushing. Note that PMP network is deployed to generate the motor commands for pushing, reaching the object to be pushed with the appropriate hand pose. How does this knowledge come to use in the context of the goal: this is the inverse problem of generating goal directed pushing movements (using the task relevant effector/tool). While there are several uses of pushing, the most basic is the problem of visuo-spatial aiming and pushing a target
object to the goal. An example is shown in the bottom panel of figure where the goal of iCub is to push the red object to the goal location indicated by the green cloth. Firstly, based on the spatial position of the target and the object type, the appropriate effector and hand pose is selected (PMP allows to generate motor commands to reach the target object). The neurons in the action space are activated based on the desired direction of pushing (based on the goal and object being pushed). Such activations influence the dynamics in the sensory space through motor modulated lateral connections, generating an incremental prediction of the future position of the target if pushed in this direction. This in turn drives the dynamics of action space based on the predicted location of the target and the goal. The result of this circular dynamics is the synthesis of the “virtual trajectory” (the pull of the puppeteer as in section) that trigger the PMP system to generate necessary motor commands to actually execute the goal directed pushing action, hence solving the inverse problem. The work is preliminary and has to be developed further, this was just the first attempt to liberate the pushing system from task specific details and make it more abstract and task independent. A lot needs to be improved and these are first few explorative attempts. Knowledge related ‘counterforce’ ‘obstacles’ as in the trap tube paradigm, effects of “contact” through touch and force sensing is yet to be incorporated. Future extentions will focus on this “direction”!

5.6 On the “Coevolution” Learning and Reasoning in Darwin, abstraction levels

The ability to reason, to orchestrate thought and action in accordance with internal goals, especially when inhabiting unstructured (and sometimes hostile) environments is a fundamental feature of any kind of cognitive behavior. Though reasoning per se is not a subject of this deliverable, reasoning process is where all learnt experience in terms of perceptual, conceptual and motor knowledge meet in the context of the goal at hand. In addition, reasoning plays a fundamental role in guiding the learning process itself to acquire new knowledge, skills etc. In this sense, it is not reasonable to decouple reasoning and learning, both processes drive each other, in order to enable the robot learn more, reason and better anticipate the functioning of the world during the pursuit of its “goals”. From a global perspective, the reasoning layer under development can be visualized as a loosely coupled network of dynamical systems, with the power of sensorimotor exploration, self organization, field computing and value dependent learning exploited at all levels of hierarchy. Special emphasis is given to acquisition of internal models and coordination of their predictions with the bottom up (real sensorimotor information) in a goal directed fashion. In order handle this effectively, the Darwin architecture is supported with sophisticated memory structures and communication protocols that take care of organizing, retrieving and managing massive amounts of information from different sources, in the context of the goal. Further details on this issue can be found in D6.1. As seen in figure 22 (left panel), internal models exist at different levels of abstraction, a level of circularity between perception-action (real/imagined) always being maintained, inconsistency at some level automatically forming the basis of reasoning/acting/exploring at some other level.

a) force-displacement (forward inverse models that deal with mental simulation at the detailed level of forces, positions, motor commands, task constraints, bodily constraints, use of different tools etc, for example the PMP based body model, pushing system, Toy crane skill);

b) object-action (more abstract internal models that encode object- action-consequence loops, goal-initial condition-plan loops, utility values and significance of various action plans, etc., at this
level the internal models doing detailed simulation of actions are abstracted to single neurons that
trigger activation of associated detailed simulation model);
c) goals-plans (motivation-executive control–reasoning loop that deals with most optimal
exploitation of the body resources under presence of multiple competing goals, opportunistic out of
turn executions of actions taking into account a measure of “well being” of the robot).

Each abstraction level itself can have multiple such atomic forward/inverse models, all operating
with similar computational principles underneath (exploration-self organization-field computing-
value dependent learning). While they operate within their local “computational scope” they also
have well defined means to exchange useful information between them through global memory
structures that maintain global harmony.

Figure 22. The ‘continuous’ coupling between reasoning and learning in Darwin and the various levels of
eexistence of internal models all functioning with a level of “circularity” actions of one serving as affordance to
another.

Going deeper inside, the every brick in the cognitive control layer is based on well known
principles like the theory of self organization and their extension to growing networks (Fritzke,
1995), Neural field dynamics (Amari, 1977; Erlhagen and Schoner, 2002) to organize goal directed
planning, prediction, tracking in the various state spaces; the concept of sensorimotor maps to
dynamically couple sensor representations with motor representations during prediction and
planning (Toussaint et al, 2006) extended to switch between exploration and exploitation based on
contradictions between the top down (mentally simulated) and bottom up information (Mohan et al,
2011), temporal Hebbian learning (Dayan, P., & Abbott, L., 2001) to change the internal model
(and lateral connectivity between neurons) in response to dynamic changes in the world detected
by the robot, some intuitive heuristics to distribute rewards/penalties in growing sensorimotor
maps. The local network structure is complemented with powerful systems to organize global
information to retrieve and transport information at the right places based on context. In any case
learning and organization of memory have to go hand in hand. The internal structure of every
internal model is approximately composed of the following four elements:
a) a set of sensorimotor variables (that characterizes the sensory and motoric scope of the model)

b) a set of connectivity structures (that mediates the interactions between the sensorimotor variables: through lateral, intermap, motor modulated, conceptual connections)

c) a set of value fields (that organize goal directed behaviour in the respective sensorimotor spaces taking into account task relevant constraints represented in terms of superimposed fields)

d) a set of trajectories (in the respective sensory and motor spaces, that characterizes the real/virtual behaviour of the system in response to the “goal” issued to it)

Figure 23: Figure 8. Pictorial illustration of the ‘Two sticks paradigm’ (top right panel) applied to the iCub robot in the simplified case that a single stick is suitable and available as a tool. The goal of iCub is the reach the large red cylinder, placed out of reach. As seen, there is a green stick available in the environment. In order to realize its goal in this situation, iCub performs a sequence of overt and covert actions: 1) Mentally estimating weather the goal is directly reachable with either arm using the PMP network for the upper body (figure 12); P1; 2) evaluating the size of the required stick-like tool based on the discrepancy between the goal and the final reachable position predicted by the forward model P1; 3) visually detecting the (green) tool; 4) evaluating whether the long green stick is reachable with an appropriate wrist orientation (using the composite PMP network of figure 14); P2; 5) reaching and grasping the stick using the same PMP network; P2; 6) incorporating the stick in the body schema by updating the Jacobian taking into account the length and orientation of the stick coupled to the end effector; 7) using the stick to reach the target cylinder using the PMP network of figure 23 right panel. Here, since the tool is coupled to the left arm, the right arm network is shown deactivated (goal=initial condition, force field in the right arm network is 0). Since the coordinated tool is the most distal part of the resulting PMP network, goals act on the tool. The field generated by the discrepancy between goal and tool position is mapped into an equivalent torque field by the transpose Jacobian ($J_L^T$). This torque field is mapped into joint rotation patterns for the left arm by the Admittance matrix. The Jacobian $J_L$ now transforms this information into next incremental update in the tool position in the end effector space (tool is the end effector now). This process of incremental updating of every node in
the PMP network continues till the time the force field in the left arm network is also zero (i.e. the tool tip reaches the target).

All atomic models are endowed with the capability continuously learning/ incorporating newer subtleties discovered in the environment concerned within the context of the sensorimotor scope they represent and goal they seek to solve. By loosely demarcating the internal dynamics, connectivity (associated neural networks), associated reward structures of every computational model local to itself, it becomes possible to “abstract” the complexity at the local level, when it comes to dealing with complexity at a higher, more global level. For example, a reasoning process that wants to evaluate the possibility of using a stick as a tool or the consequence of pushing a ball with the grasped stick, need not be concerned about the complexity at the level of end effector trajectories, necessary wrist orientation to grasp the stick, the joint angles needed to coordinate the body or the direction of pushing based on the location of traps. All it needs to learn is to find who is the right person (or the action variable) to do the job! See figure 5 and () for initial results in this direction. There are few preliminary results and we agree this may even be trivial considering the overall objective of the project. There are the first few steps and the efforts at this phase are focused on laying strong theoretical foundations to allow Darwin realize its full potential in the years ahead.

To sum up the basic reasoning module is being designed with a dual objective of both supporting/enhancing “learning” as well as enabling the robot to perform goal directed action. Outcome of the performance (real/virtual) of one atomic action automatically creates an affordance for some the other action to take place. Actions executed by one arm, becomes an affordance for the other arm, actions executed by one robot becomes an affordance for some other robot. At the level of the “world”, bodies compete/cooperate; at the level of the “body”, goals compete/cooperate; at the level of “goals”, reasoning processes compete/cooperate; at the level of “reasoning”, actions compete/cooperate; at the level of “actions”, force fields compete/cooperate, hence taking us back to ground zero!

Co c u io A ro or u ture ictor i u r

Darwin “learning” and “reasoning” architecture is being developed with a goal to slowly and gradually enable cognitive robots to face the real world. We realize that the vision is ambitious and most often we anticipate to find ourselves in uncharted areas. This gives us both a sense of challenge and a sense of fulfillment while we cruise ahead. In this deliverable, we summarized the core functional requirements in terms of learning in Darwin, a roadmap of scenarios organized in the form of four core loops to aid cumulative learning in Darwin, practical assumptions and guiding principles that are being taken into account while designing the first prototype of the cognitive control layer, some results of ongoing work, key challenges and how they will be developed further. The research issues addressed constitute some of the basic challenges for research in cognitive robotics in general in terms of creating architectures that enable robots to gradually expand their perceptual, motor and cognitive horizons through interaction and learning. Other core issues like development of linguistic/communicative capabilities (new developments in phonetic and articulatory systems, influence of language on concept acquisition and other processes), research on emotional interfaces, social intelligence and human robot interaction nevertheless crucial are out of scope in terms of the core objectives of Darwin. Hence, they have not been addressed here. If we consider future advancements in cognitive robotics and the parallel progresses in the various cognitive and behavioural capabilities required in the cognitive robot itself, we can identify a
potential sequence of milestones that provide a possible set of goals and test-scenarios. We do not propose a fully defined discrete, sequential list of milestone, especially as there will be overlap of cognitive capabilities development in the transition between milestones/stages. This milestone list, together with other roadmap proposals (Baldassare et al, 2012), can contribute to the evaluation of advances for future developmental cognitive robotics research.

<table>
<thead>
<tr>
<th>Action Development</th>
<th>Development of Basic primitives like reaching, grasping taking into account motor redundancy, structural and basic task related constraints</th>
<th>Advanced grasp functionality, Bimanual actions. Development of hierarchical and compositional actions, Preliminary action perception and imitation (ex. Movement trajectories)..building up on basic primitives</th>
<th>Synchronized use of two hands to manipulate objects. Use of imitative and explorative learning to actively learn motor skills: like tool use (building up on previous motor repertoires), highly accurate reach and grasp functionality in unstructured scenarios.</th>
<th>Ability to actively recycle past Motor experience while learning new skills, development of basic action understanding of others, preliminary interface of action with language</th>
<th>Ability of correlate composite actions with composite linguistic descriptions. Ability to learn rich action repertoires based effortlessly building up on the existing motor vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceptual development</td>
<td>Preliminary object identification (based on color, and possibly surface markings), 3D localization and extraction of basic geometric attributes to afford action. Preliminary integration with touch focused on binary information about contact.</td>
<td>Ability to discriminate, more complex object properties beyond local patches e.g. (using shape). Dense multi view stereo vision. Active learning to extract object affordances to support basic actions like reach / grasp (push with a demand for accuracy (i.e. manipulating small objects, industrial parts, elastic objects)).</td>
<td>Categorization of novel objects though a combination of color, shape, texture information. Integration of depth images and subsequent merging into a complete 3D model. Basic use vision with touch information to aid manipulation (like inserting a screw or cable into a Meccano part)</td>
<td>Active use of action and attention to aid perception. Categorization and localization of a large range of objects required to execute specific industrial assembly asks (under constraints of industrial requirements in terms of cycle time, accuracy etc).</td>
<td>Multimodal sensory abstraction capabilities. Classification of novie objects through visuotactile properties. Full integration of tactile, proprioceptive and force related information with vision to fill in perceptual gaps at the edge of physical interactions with objects.</td>
</tr>
<tr>
<td>Cognitive development</td>
<td>Basic reactive mechanisms to detect failures in achievements of goal and being able to replan the behavior from the new initial condition.</td>
<td>First prototype motivation, concept and reasoning system deployed to organize global behavior mainly learning of affordances, actions and their consequences and use of such knowledge to plan simple goal directed action sequences. Fully functional internal body model for the coordination of both covert and overt action using the body.</td>
<td>High level of knowledge related to physical causality, affordances of various tools and objects, predictive powers to determine how they behave, what they are good for and what new can be learnt. Identify potential problems for carrying out an assembly task and reason what new information is needed in order to complete it</td>
<td>Full maturation of the Darwin cognitive architecture with powerful perceptual, motor and reasoning capabilities to allow the robot to play, interact and learn autonomously in semi structured, noisy and novel environments, handling multiple goals in an opportunistic fashion.</td>
<td>Tighter integration of physical intelligence with social intelligence enhanced linguistic and communicative capabilities. Cooperative and competitive coevolution of multiple Darwin’s learning from each other, sharing rewards during their pursuit to realize goals</td>
</tr>
<tr>
<td>Integrated Cognitive</td>
<td>(End Year 1 Darwin)</td>
<td>(End Year 2 Darwin)</td>
<td>(End Year 3 Darwin)</td>
<td>(End Year 4 Darwin)</td>
<td>Post Darwin (3-4 years’ goal)</td>
</tr>
</tbody>
</table>

To conclude, we remark that this effort may just be equivalent to playing on the beach and vast range of problems still remain to be untangled in the quest to better understand the ‘forces and the causes’ that should play a role in shaping ‘reasons and actions’ of our cognitive robots to make them as explorative, intuitive, expressive, irrational, unpredictable as we really are. Coming to the life time of Darwin project itself, we conclude with a “picture” that summarizes the how we envision the gradual expansion of the cognitive horizons of the Darwin cognitive architecture as we move ahead.
Figure 24. A hypothetical timeline of the Darwin robot attempting a simple task of “grasping a red ball” under various environmental conditions coupled with gradual increments in its cognitive capabilities (in order to scale up to the equivalent complexity in the environment in which it has to achieve this simple goal). As time progresses, we also see the link between self and other being ‘strengthened’ in order to allow stronger coupling between multiple Darwins and humans, to support coevolution and co development of learning to serve joint “goals” of mutual interest.

### References


64


