Rob's Robot: Current and Future Challenges for Humanoid Robots

Boris Durán and Serge Thill
University of Skövde
Sweden

1. Introduction

The purpose of this chapter is both to review the current state of the art in robotics and to identify some of the challenges that the field has yet to face. This is done and illustrated by following an imaginary roboticist, Rob, who is currently designing a new humanoid robot from scratch. Rob's ultimate vision is that this robot will be humanoid in the sense that it possesses the abilities of a human being. Therefore, his initial aim is to identify relevant work as well as areas in which innovations are still needed.

Research in robotics has come a long way in the last few decades, whether one considers humanoids specifically or just any type of architecture. The world is increasingly moving away from being confined to industrial robots with precisely programmed movement plans and is beginning to include adaptive robots that can robustly operate in a variety of scenarios as well as robots that show increasing levels of cognitive abilities and intelligence. We have witnessed an increasing collaboration between research areas that, previously, had very little in common. Nowadays, it is not difficult to find research groups uniting neuroscientists, engineers, psychologists, linguists, even philosophers and musicians, all coming together to study the human body as well as the human mind. There is thus plenty of previous work that Rob can build upon when designing his robot. At the same time, it is clear that the field is still a long way from creating truly human-like machines and that there are thus significant challenges to overcome.

In this chapter, we first provide a brief history of humanoid robotics research since it sets the background to Rob's work. We then present the state of the art in the different components of a robot, hardware as well as software, discussed in contrast with Rob's requirements and highlighting particular areas in which further research is needed before Rob's robot can become reality.

2. A brief history of humanoid robotics research

Our fellow roboticist Rob goes back in time in order to trace the roots of his passion. He finds out that attempts to create automated machines can be traced back to the first century of our era with the inventions of Hero of Alexandria, a Greek mathematician and engineer from Roman Egypt. Among the first human-like automata he found one created by another polymath inventor from Mesopotamia in the 12th century known as Al-Jazari. His group of "robots" used water to drive their mechanisms and play four different instruments.

Among others who designed and/or created human-like machines he discovers that:

• Leonardo Da Vinci designed a mechanical knight at the end of the 15th century. This machine was able to sit up, wave its arms and move its head and jaw.

• Hisashige Tanaka, a Japanese craftsman in the 19th century, created several extremely complex toys able to serve tea, fire arrows and paint kanji characters.

• In 1929, Japanese biologist Makoto Nishimura designed and created *Gakutensoku*, a robot driven by air pressure that could move its head and hands.

Rob also comes across some interesting information about the word "robot". It has its origins in the 1921 play "R.U.R." (Rossum's Universal Robots) created by the Czech writer Karel Capek, although he later credited his brother Josef Capek as the actual person who coined the word. In Czech, the word "robota" refers to "forced labor"; in this sense a robot is a machine created to perform repetitive and expensive labor. In 1942 Isaac Asimov introduces the word "robotics" in the short story "Runaround". His robot stories also introduced the idea of a "positronic brain" (used by the character "Data" in Star Trek) and the "three laws of robotics" (later he added the "zeroth" law).

However it wasn't until the second half of the 20th century that fully autonomous robots started being developed and used in greater numbers. Following the rhythm of a demanding industrial revolution the first programmable (autonomous) robot was born. Unimate was used to lift and stack hot pieces of die-cast metal and spot welding at a General Motor plant. From then on, other single arm robots were created to cope with larger production requirements. Nowadays industrial robots are widely used in manufacturing, assembly, packing, transport, surgery, etc.

Humanoid robotics, as a formal line of research was born in the 1970s with the creation of Wabot at Waseda University in Japan. Several projects around the world have been developed since then. The next section will briefly describe the current state-of-the-art in anthropomorphic robots.

2.1 State of the art

This section starts by presenting HONDA's project ASIMO, Fig. 1(a). HONDA's research on humanoid robots began in 1986 with its E-series, a collection of biped structures that developed, without doubt, the most representative example among state of the art humanoid robots (Hirose et al., 2001). Today's ASIMO has 34 degrees-of-freedom (DOF), weights 54kg and has a total height of 130cm. This robot is able to walk and run in straight or circular paths, go up or down stairs and perform an always increasing set of computer vision routines like object tracking, identification, and recognition, localization, obstacle avoidance, etc.

A joint project between the University of Tokyo and Kawada Industries gave birth to the Humanoid Robot "H6" and ended up with HRP-3 in 2007 (Kaneko et al., 2008). The National Institute of Advanced Industrial Science and Technology (AIST) continued this project and developed today's HRP-4, Fig. 1(c), a 34 DOF walking robot weighting 39kg and measuring 151cm in height. This light-weight full size robot was designed with a reduction of production costs in mind, thus allowing access to the average robotics research group around the world to a pretty advanced research tool.

Korea's Advanced Institute of Science and Technology (KAIST) started working on humanoid platforms in 2002 with KHR-1. Today's KAIST humanoid is known as HUBO2 (KHR-4) and has a total of 40 DOF, a height of 130cm and weight of 45kg, Fig. 1(b). HUBO2 is also able to walk, run and perform an increasing set of sensor-motor tasks (Park et al., 2005).

(a) ASIMO   (b) HUBO2   (c) HRP-4

(d) iCub   (e) Robonaut2   (f) Twendy-One
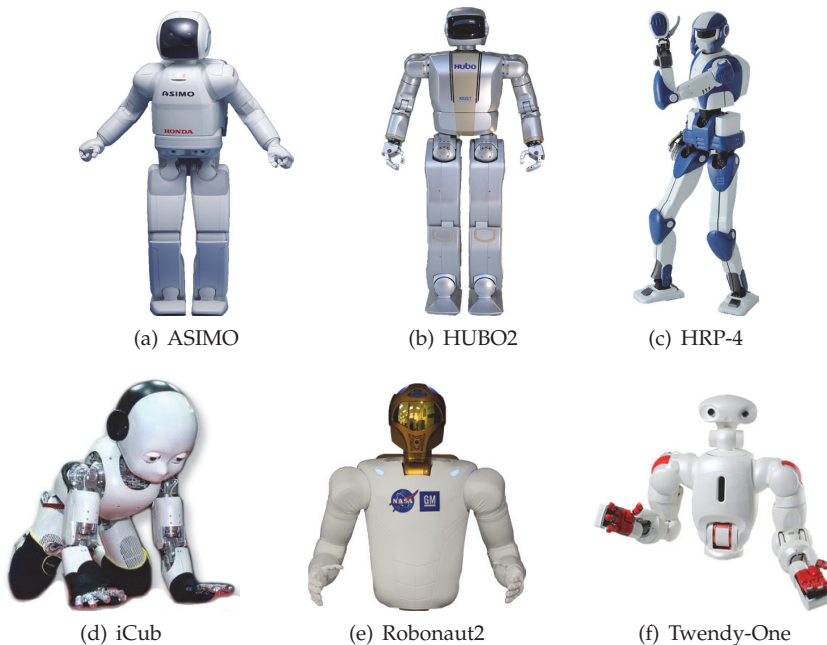
Fig. 1. Some state of the art humanoid platforms.

The RobotCub Consortium formed by 11 European institutions created iCub, a child sized humanoid robot which until 2010 was only able to crawl, Fig. 1(d). This robot has 53 DOF, weights 22kg and has a height of 104cm (Tsakarakis et al., 2007). A remarkable feature of this project is that everything developed within it was made available as open source, including the mechanical design and software products. This move has created a growing community that actively collaborates at all levels of research.

General Motors joined NASA for the second version of NASA's project Robonaut (Ambrose et al., 2000). This upper body humanoid robot (Fig. 1(e)) is a very dexterous robotic platform with 42 DOF, 150kg and 100cm (waist to head). Another upper-body robotic platform worth including as state-of-the-art is Waseda University's TwendyOne, a very dexterous and high-power robot with 47 DOF, 111kg and 147cm height, Fig. 1(f).

One common characteristic of all the above mentioned platforms is the use of electric torque motors in their joints. There have been, however, new attempts to use hydraulic and pneumatic actuators such as the ones in Anybots' biped Dexter which is also able to jump (Anybots, 2008); or Boston Dynamics' PetMan which introduces a much faster and more human-like gait (Petman, 2011). These two examples make use of pneumatic and hydraulic actuators respectively but there is also a growing interest around the use of pneumatic artificial muscles or air muscles. Shadow Robot's pneumatic hand (Shadow Robot Company, 2003) performing highly dexterous movements and Festo's use of artificial muscles (Festo, 2011) are also worth mentioning as state-of-the-art due to the high accuracy of these otherwise difficult to control kind of actuators.

3. Mechanical requirements and engineering challenges

It is a matter of philosophical discussion how far it is possible to discuss hardware and software separately. Section 4.2 in particular will discuss theories from cognitive science which indicate that in the human, body and mind are intrinsically intertwined and that some of the hardware can in fact be considered to fulfill certain roles of the software too. However, entering into such a debate in detail is beyond the scope of this chapter. Therefore, when considering the different components of Rob's robot, we will largely separate hardware and software in the sense that we will discuss purely mechanical requirements first and computational requirements next. Since hardware can fulfill computational requirements and mechanical requirements can impose software requirements as well, there may be some natural overlap, which in a sense already illustrates the validity of some of the ideas from embodied cognition (to be discussed in section 4.2.1).

First then, Rob needs to identify mechanical and engineering requirements for the creation of his humanoid robot. He is interested in particular in the state of the art of components like sensors and actuators, and must consider how close these elements are to his ambitious requirements and what major challenges still need to be addressed before his vision can become reality. The remainder of this section is dedicated to answering these questions.

3.1 Sensors

Throughout the history of humanoid robotics research there has been an uneven study of the different human senses. Visual, auditory and tactile modalities have received more attention than olfactory and gustatory modalities. It may not be necessary for a robot to eat something but the information contained in taste and odor becomes important when developing higher levels of cognitive processes. Thinking about the future of humanoid robots we should be careful not to leave behind information that may be helpful for these machines to successfully interact with humans.

In this section Rob studies some of the devices that are either currently being used or are under development in different robotic platforms; they are categorized according to the human sense they relate to the most.

3.1.1 Vision

Humanoid robots arguably pose the largest challenge for the field of computer vision due to the unconstrained nature of this application. Rob's plans for his autonomous humanoid involve it coping with an unlimited stream of information changing always in space, time, intensity, color, etc. It makes sense for Rob that among all sensory modalities within robotic applications, vision is (so far) the most computationally expensive. Just to have an idea of how challenging and broad the area of computer vision is, Rob remembers that in 1966 Marvin Minsky, considered by many to be the father of artificial intelligence (AI), thought one master project could "solve the problem of computer vision". More than 40 years have passed and the problem is far from solved.

In robotic applications, a fine balance between hardware and software is always used when working with vision. During the last decade, Rob witnessed an increasing interest on transferring some software tasks to the hardware side of this modality. Inspired by its biological counterpart, several designs of log-polar CCD or CMOS image sensors (Traver & Bernardino, 2010) (Fig. 2(b)) and hemispherical eye-like cameras (Ko et al., 2008) (Fig. 2(c)) have been proposed. However, the always increasing processing power of today's computers and specialized graphical processing units (GPU) have allowed many researchers

to successfully tackle different areas of vision while keeping their low cost Cartesian cameras (Fig. 2(a)) and solve their algorithms with software. Nonetheless, Rob welcomes the attempts to reproduce the high resolution area at the center of the visual field, fovea, with hardware and/or software. He thinks that nature developed this specific structure for our eyes throughout evolution and it may be wise to take advantage of that development.

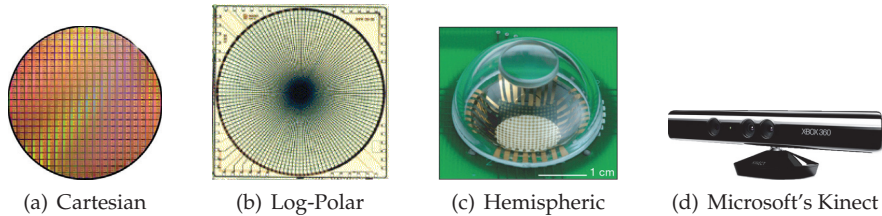| (a) Cartesian | (b) Log-Polar | (c) Hemispheric | (d) Microsoft's Kinect |

Fig. 2. Samples of different configurations for image sensors.

Rob identifies several challenges for the hardware component of vision sensors. Potentially the most important is the need to widen the field of view of current cameras. Normal limits for the human field of view are approximately 160 degrees in the horizontal direction and 135 degrees in the vertical direction. A typical camera used in robotic applications has a field of view around 60 degrees for the horizontal direction and 45 degrees for the vertical direction. Important information is lost in those areas where current cameras can not reach in a single stimulus. Adaptation to changes in light conditions is also a complex and difficult process for artificial vision, especially when working within a dynamic environment. Rob is considering a few options from the video security market based on infra-red and Day/Night technology that could be adapted for use in humanoid robots. Finally, the robotics community has experienced a growing interest on using RGB-D (Red-Green-Blue-Depth) cameras after the release of the low-cost motion capture system Kinect by Microsoft (Shotton & Sharp, 2011), Fig. 2(d). This technology combines visual information and high-resolution depth, opening therefore new possibilities for overcoming the challenges of 3D mapping and localization, object identification and recognition, tracking, manipulation, etc.

3.1.2 Tactile

Even though touch sensors are being used on few specific points of humanoid platforms (e.g. tip of fingers, feet, head), Rob thinks that an efficient solution to acquiring information from large surfaces is needed in order to be able to exploit the richness of textures, shapes, temperatures, firmness, etc. Therefore he argues the importance of developing skin-like sensors as another future challenge for humanoid robotics.

In robotic applications, Rob has used his imagination to create solutions for detecting objects in the path of an arm, finger or leg. One of these solutions has been the detection of peaks of current in their electric motors or using torque and force sensors. This approach could be considered tactile sensing as well but within the area of proprioception since measurements are made as relative positions of neighbouring parts in static and dynamic situations. In humanoid robots, exteroceptive sensors include dedicated pressure sensors placed on points of interest. The technology used in most pressure sensors so far include: resistive, tunnel-effect, capacitive, optical, ultrasonic, magnetic, and piezoelectric. Resistive and capacitive sensors are certainly the most common technology used not only in robotic but in many consumer electronic applications (Dahiya et al., 2010).

Advances in nano-materials are making it possible to integrate a large amount of sensors in flexible and stretchable surfaces (Bao et al., 2009; Peratech, 2011; Takei et al., 2010). Rob finds the approach taken by Prof. Zhenan Bao very interesting since it includes not only a sensitive resolution that surpasses that of human skin, but also the possibility to sense chemicals or biological materials (Bao et al., 2009). In addition, they are also working on embedding solar cells within the same films (Lipomi et al., 2011), which is an outstanding innovation since it would help to solve another major challenge for humanoid robots, i.e. energy.

Rob remembers that human skin is composed of two primary layers: epidermis and dermis. The dermis is located beneath the epidermis and contains all sensors for temperature and touch. He thinks about this because the dermis helps also to cushion the body from stress and strain, a task that is augmented by the morphological interaction of muscles, water, bones, etc. This dynamic plays an important role in tasks such as manipulation and locomotion. Therefore the design of Rob's future humanoid robot's skin should include not only a large amount and various types of sensors, but also an inner mechanism that provides similar dynamics like those found in human bodies.

3.1.3 Sound

Rob knows that microphones are the standard tools for acquiring sound information from the environment in a robotic platform. They can be installed in different configurations and detect/regulate sound in ways that surpass human abilities. But detecting a sound is just the starting point of a more complex challenge. Once a sound has been detected, it is time to localize the origin or source of that sound. Humans' remarkable ability to detect, localize and recognize sound can be exemplified using the *cocktail-party effect*. This term describes the ability of a person to direct his/her attention on the source of a sound amidst other sound sources which makes possible, for example, to talk to another person in a noisy, crowded party (Haykin & Chen, 2005).

It has been shown in humans that a continuous interaction between auditory and visual cues take place in the brain (Cui et al., 2008). In humanoid robots the localization of a sound source is mainly done by computing the difference in times and levels of signals arriving to the different sensors, in this case microphones (Keyrouz, 2008). Attempts to integrate sensor information from visual with auditory modalities has also been done in humanoid platforms (Ruesch et al., 2008).

Rob trusts that at least from the hardware point of view, he will not have many problems. The real challenge for the future is to improve the way different sensor modalities can be merged to replicate or surpass the abilities found in humans. However this point belongs to the *mindware* of his platform, so he will focus on working in optimized configurations of microphones.

3.1.4 Odor and taste

Rob gets a strange feeling at this point and thinks, "Should I even consider spending time thinking about a robot with these two senses? After all, robots do not need to eat anything as humans do, they just need a power cable and will have all the energy they need. And electricity has no smell, right?" But then he starts thinking in more holistic terms– maybe robots will need to somehow recognize smells and flavors; moreover, they will need to associate that information with visual, auditory and tactile cues. Otherwise it is going to be difficult, if not impossible, to interact with humans using more or less the same language. Artificial noses and "tongues" have already been developed and they are also overmatching those of humans (Mahmoudi, 2009). Artificial noses are able to detect and classify thousands

of chemical molecules ranging from biological agents to wine and perfume. As in the case of sound, it does not seem difficult to trust that in the near future Rob will have access to a device that gives his robot information about odor and taste, at the very minimum at human levels.

3.2 Actuators

In the case of actuators, Rob illustrates the current state of the art of active components, although he is lately more biased towards approaches which make use of spring-damper components. He knows that the right choice of these components forms one of the critical challenges for his future humanoid robot. Rob's preference for non-linear components comes from the fact that the human body is made, from an engineering perspective, almost exclusively of spring-damper type elements: cartilages, tendons, muscles, fat, skin, etc. whose interaction results in stability, energy efficiency and adaptability.

3.2.1 Whole-body motion

Humans have a large repertoire of motor actions to move the whole body from one point to another. Walking could be considered the most representative behavior in this repertoire but we learn to adapt and use our bodies depending on the circumstances. Humans are also able to run, crawl, jump, climb or descend stairs, and if using the arms, then we can squeeze our bodies between narrow spaces or if in water we can learn to swim.

Rob realizes that researchers in whole body motion for humanoid robots have focused most, if not all their attention in walking only. The most common approach to control the walking behavior and balance of a humanoid robot is called Zero Moment Point (ZMP) and it has been used since the beginning of humanoid robotics research (Vukobratovic & Borovac, 2004). This approach computes the point where the whole foot needs to be placed in order to have no moment in the horizontal direction. In other words, ZMP is the point where the vertical inertia and gravity force add to zero. Most state of the art humanoid platforms like ASIMO, HRP-4, and HUBO2 make use of this approach. The main drawbacks of ZMP arise from the need to have the whole foot in contact with a flat surface, and it assumes that this surface has enough friction to neglect horizontal forces.

Passive-dynamic walkers are an alternate approach to humanoid locomotion (Collins et al., 2005). These platforms try to exploit not only the non-linear properties of passive spring-damper components but the interaction of the whole body with the environment. The result is, in most cases, a more human-like gait with a heel-toe step in contrast to the flat steps seen in ZMP-based platforms. State of the art passive-dynamic walkers use dynamic balance control (Collette et al., 2007) which provides them with robustness to external disturbances and more human-like whole-body motor reactions. Examples of this approach can be found in platforms such as Dexter (Anybots, 2008), PetMan (Petman, 2011), and Flame (Hobbelen et al., 2008).

Rob also found out that running and jumping have already been implemented in a few of the current humanoid platforms (Anybots, 2008; Niiyama & Kuniyoshi, 2010), although there is still much work to do to reach human-like levels. Once humanoid robotics started to become the meeting point of different scientific groups such as developmental psychologists, neuroscientists and engineers, interesting topics emerged. One of them was the study of infant crawling, its implementation in a humanoid platform was done using the iCub robot (Righetti & Ijspeert, 2006), Fig. 1(d).

For our roboticist Rob, the challenge of making humanoid robots replicate the different types of motor behaviors found in humans is just one part of a larger challenge. An equally

interesting project will be the design of a decision making control that allows the agent to switch between the different motor behaviors. Switching between walking to crawling and vice versa, from walking to trotting or running and back. Those changes will need to be generated autonomously and dynamically as a response to the needs of the environment. The traditional way of programming a robot by following a set of rules in reaction to external stimuli does not work and will not work in dynamic, unconstrained environments. Rob is thinking about Asimo tripping over a step and hitting the floor with his face first. Rob knows that a more dynamic, autonomous and adaptive approach is needed for the future of his humanoid robot.

3.2.2 Manipulation

It is now time for Rob to think about manipulation. He looks at his hands and starts thinking about the quasi-perfection found in this mechanism. He realizes that there is no other being in the animal kingdom that has reached similar levels of dexterity. His hands are the result of an evolutionary process of thousands of years which got them to the sophisticated level of manipulation the enjoy today. This process could have gone "hand-in-hand" with the evolution of higher cognitive abilities (Faisal et al., 2010). Considering human hands only, there is an even larger repertoire of motor actions that range from simple waving and hand signals to complex interactions such as manipulating a pen or spinning a top or shooting a marble (see Bullock & Dollar (2011) for a more detailed classification of human hand behaviors).

Rob is aware of the particular engineering challenge presented by interacting with objects in the real world. Trying to replicate as close as possible the human counterpart, current state of the art humanoid hands (Fig. 3) have four (Iwata et al., 2005; Liu et al., 2007) and five (Davis et al., 2008; Pons et al., 2004; Shadow Robot Company, 2003; Ueno & Oda, 2009) fingers, sometimes under-actuated through a set of tendons/cables as in the case of iCub's hands (Fig. 1(d)). However three-finger hands are also being used for manipulation, the most remarkable case being the work done at the Ishikawa-Oku Lab at the University of Tokyo (Ishikawa & Oku, 2011),Fig. 3(d). Through the use of high-speed cameras and actuators they have showed that it is possible to perform complex actions like throwing and catching or batting a ball, knoting a flexible rope with one manipulator, throwing and catching any type of objects with the same hand, or manipulating tools of different shapes.

The interaction between skin and bones in a human hand creates a dynamic balance between sensing and acting capabilities. By including skin-like materials into the design of robotic end-effectors it will be possible to re-create the optimized functionality of human hands. Skin that includes soft materials in its construction will add spring-damping properties to grasping behaviors. Accurate control of the hardware is transferred to the physical properties of the materials thus saving energy, time and resources. The future humanoid manipulator should be capable of interacting at a minimum with the same objects and displaying similar active and passive properties as human hands.

3.3 Interim summary 1

Rob realizes that the technology necessary to design and build a humanoid platform seems to be reaching a very mature level. In many cases the biggest obstacle is the lack of resources to put together state of the art sensors and actuators in the same project. We have seen remarkable innovations in the area of nano-materials which can help to overcome current obstacles in the acquisition of tactile, visual, and olfactory/taste information. At the same

(a) Shadow Hand

(b) TwendyOne's hand

(c) ELU2's hand
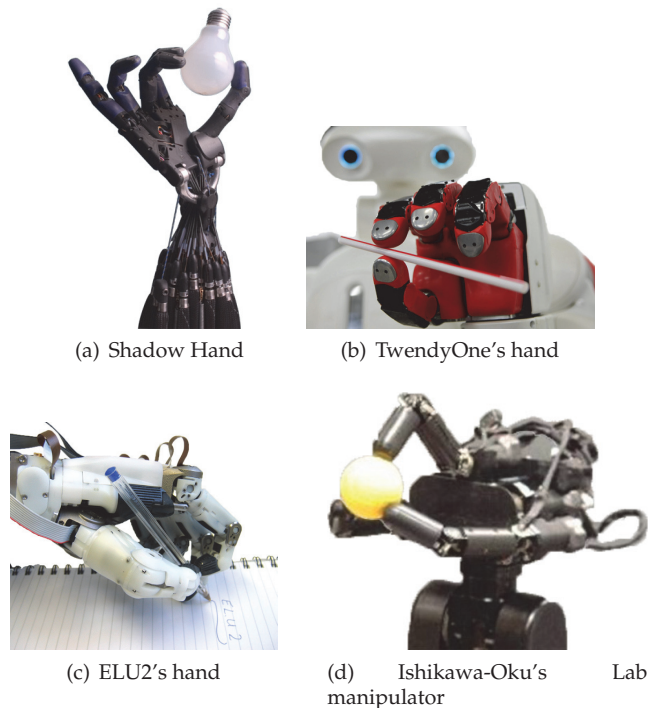
(d)   Ishikawa-Oku's   Lab manipulator

Fig. 3. Samples of different state of the art manipulators.

time, the exponential growth of computational power gives enough freedom to integrate all this information.

In terms of actuators, Rob is impressed by the results from the use of spring-damper components. Previous attempts to control this kind of materials gave many researchers (especially to those working with traditional control theory) several headaches. However current implementations for whole-body motion and manipulation have shown the feasibility of this approach, thanks most likely to the use of alternative methodologies, e.g. nonlinear dynamical systems.

4. Computational requirements

Having considered his mechanical requirements, Rob now turns to computational requirements for his robot. In his mind, this means anything that could be needed to make the hardware perform interesting behaviors. This includes, but is not limited to, learning how to solve (not necessarily pre-defined) tasks in uncertain environments under changing or undefined conditions and interacting with humans in a non-trivial manner.

Within robotics, Rob has distinguished two main approaches of interest to him: one which we shall call the traditional symbolic approach (also known as *Cognitivism*) and one that we will name embodied approach (also known as *Emergent Systems*). Although both approaches can sometimes use the same technologies (neural networks are a prime example of that), they differ philosophically in the sense that symbolic is first and foremost an AI approach

with its roots in computer science, whereas the embodied approach has its roots in the cognitive sciences. An important consequence is that the traditional symbolic approach sees computations as fundamentally disassociated from any specific hardware and therefore operating in an abstract, amodel level while this is not true for the embodied approach. This section therefore begins by briefly discussing some examples of the symbolic approach and its limitations before moving on to embodiment and its relevance to robotics in general and Rob's robot in particular.

4.1 Traditional symbolic approaches

The symbolic paradigm sees cognition as a type of computation between any information that can be represented in code and the outcome of the operations carried out on those codes (Vernon, 2008). In other words, symbols are manipulated through a set of rules. This determines that the way an artificial system sees the world around it is only through the eyes of the designer/programmer of that system. All its freedom and limitations will be dependent on the amount of symbols and rules that the person who created this model wrote in code.

Symbolic approaches make use of mathematical tools that help them select the best response within a specific goal. Among the most important are probabilistic modeling and machine learning techniques. Both of these have been relatively successful in solving task specific problems. Hundreds if not thousands of commercially available products for character, voice, and face recognition are on the market right now. Classification and regression problems are easily solved by this approach. Nonetheless, the dependency of the system to the programmer's view of the world challenges symbolic approaches in solving problems such as the symbol grounding problem, the frame problem, and the combinatorial problem (Vernon, 2008).

Rob has seen Asimo tripping and falling a couple of times without any reaction to protect itself from the fall. This gives Rob a clue about the limitations of the symbolic approach for controlling unexpected situations. At some point a group of engineers devised the most likely situations that Asimo would encounter when climbing or descending a number of stairs given that they would be flat, non slippery and nobody else would be around. So they knew about the symbols present in an specific sequence of events, they knew the constraints needed and imposed and finally, they wrote enough rules to fulfill the task.

Rob starts thinking about those events that cannot be foreseen and therefore cannot be calculated. What if a pebble is in the wrong place or at the wrong time? What if a cat crosses my robot's path? What if a person pushes my robot? After all, my robot will be moving around in places where any other human moves around, and we encounter these types of situations all the time. Rob has already considered the importance of a good design for both actuators and their wrapping materials. This would help by transferring some of the problems from the digital control part into the physical interactions with the environment. Yet a symbolic approach does not seem to be as open as needed in the unconstrained environments that humans work and live.

4.2 Embodied and cognitively inspired approaches

Rob thus comes to the realization that the traditional symbolic approaches do not fulfill his requirements. In this part, he therefore discusses embodied approaches. In a nutshell, this class of approaches sees the particular body of an agent as intertwined with its mind: the cognitive abilities of an agent is a consequence of both and cannot be understood by studying one in the absence of the other.

In the simplest form, the embodiment can provide a solution to the *symbol grounding problem* (Harnad, 1990), namely the problem of attaching a real meaning to symbols. For example, the sequence of symbols *to grasp* has no meaning by itself; it only becomes meaningful when an agent can associate it with the corresponding sensory perceptions and motor actions. In a more advanced form, the embodiment of a particular agent can actually reduce the computations necessary by a controller in the traditional sense by what is called morphological computing (Pfeifer et al., 2006).

The remainder of this section thus first briefly introduces the core concept of embodied cognition as relevant to roboticists such as Rob. We then consider examples of research which uses embodiment in the sensory-motor grounding sense as well as examples of morphological computing as such. Finally, we briefly discuss *dynamic field theory*, a particular modeling approach explicitly built on ideas from embodied cognition.

4.2.1 Embodied cognition

Thill (2011) provides a brief introduction of embodied cognition as relevant to the design of artificial cognitive systems in general (rather than the specific case of robots). The brief introduction here follows this discussion, albeit adapted to suit the needs of roboticists in particular.

The basic claim within embodied cognition (Anderson, 2003; Chrisley & Ziemke, 2003; Gallagher, 2005), as mentioned before, is that the body intrinsically shapes the mind. A simple illustration of this influence comes from an early study by Strack et al. (1988), who showed that people rate cartoons as more funny when holding a pen between their teeth (activating smiling muscles) than when holding a pen between their lips (activating frowning muscles). Another example is the SNARC (Spatial-Numeric Association of Response Codes) effect (Fischer, 2008, see Pezzulo et al. 2011). In essence, people respond to smaller numbers faster with the left hand than with the right hand and vice versa for large numbers. Similarly, when asked to produce random numbers while simultaneously shaking their heads left to right, people are biased towards smaller numbers during left turns than during right ones. A further illustration of the body's influence on mental processes can be seen in language processing, particularly when involving action verbs. Such processing (for instance while reading a sentence) can be shown to activate motor regions within the brain and lead to either facilitation of or interference with executing an action (involving the same end-effector as the sentence, see Chersi et al., 2010, for a review and a more thorough discussion).

Although examples such as those above and several more not discussed in detail here (but see Pezzulo et al., 2011, for additional ones) clearly show that body and mind are intertwined, it is still an open debate how intricate the relationship is. While researchers like Pfeifer & Bongard (2007) or Pezzulo et al. (2011) argue strongly in favor of such an intertwined relationship, Mahon & Caramazza (2008), for instance, are amongst those who favour a view that sees mental processes operating at an abstract symbolic representation, with concepts that are merely grounded by sensory-motor information. In other words, in this view (also related to Harnad, 1990), cognition does not require a body as such, although the latter may be necessary to ground the symbols used.

The relevance of embodied cognition to robotics in general is thus clear: when designing the controller for a robot, one faces decisions as to how much the higher-level cognitive abilities of the machine need to involve the particular embodiment and sensory features. Thus the relationship between robotics and the study of embodied cognition is mutually informative: on one hand, a robot provides cognitive scientists with a real body in which to test their

theories and computational models thereof (Pezzulo et al., 2011) while on the other hand, insights from embodied cognition can allow the better design of robots. For Rob's robot, the latter is clearly the most interesting aspect and hence the focus of the remainder of this section.

4.2.2 Mirror neurons and the social robot

One of the key requirements for Rob's robot is the ability to be social in the sense that it can interact with human beings in a sensible, autonomous manner. Such an interaction can take many forms. One example is *learning by imitation*, which would allow the robot to learn novel tasks by observing a human demonstrating the desired behavior. It could also take the form of *cooperation*, in which Rob's robot has to solve a task together with a human. Or it could be *verbal interaction*, which may be required in all social scenarios involving robots.

What is common to all these scenarios is first and foremost the requirement that Rob's robot be able to *understand* the actions of a human and related concepts, whether shown to the robot by demonstration or given via verbal labels. In other words, these scenarios can benefit from the embodied theories on sensory-motor grounding in order to facilitate this understanding.

Within this context, the discovery of mirror neurons –neurons that are active both when an agent executes a (goal-directed) action and when he observes the action being executed by another (Gallese et al., 1996) –is attracting Rob's attention. Indeed, one of the most prominent hypotheses on the functional role of mirror neurons is that they form the link between action perception and action understanding Bonini et al. (2010); Cattaneo et al. (2007); Fogassi et al. (2005); Rizzolatti & Sinigaglia (2010); Umiltà et al. (2008), which appears highly relevant for Rob's robot. Even so, it is always worthwhile to remember that this claim remains a *hypothesis*, not a proven fact, and does not come without criticism (Hickok, 2008).

It is further thought that mirror neurons may play a role in learning by imitation (Oztop et al., 2006) as well as in the sensory-motor grounding of language (see Chersi et al. (2010) for a discussion). Although these are again hypotheses rather than facts (, for instance rightly points out that most of the neurophysiologic data supporting the theories come from macaque monkeys, which neither imitate nor use language), they have inspired some robotics research (see for instance Yamashita & Tani (2008) for an example of learning by imitation and Wermter et al. (2005) for an example of language use). Overall, it thus comes as no surprise that mirror neurons are also of high interest to the field of humanoid robotics, since they may provide the key to grounding cognition in sensory-motor experiences. For this reason, Rob is interested in some of the work within the field of robotics that is based on insights learned from mirror neurons.

4.2.3 Mirror neuron based robots

Oztop et al. (2006) provides a general review of computational mirror neuron models (including models that can be seen to have components which function similarly to mirror neurons rather than being explicitly inspired by mirror neuron research), together with a proposed taxonomy. Not all of these are relevant to robotics, but we will briefly mention a few controllers that are examples of early mirror-neuron related work in robotics.

One of the first such examples is the recurrent neural network with parametric bias (RNNPB) by Tani et al. (2004), in which "parametric bias" units of a recurrent neural network are associated with actions, encoding each action with a specific vector. Then the system may then either be given a PB vector to generate the associated action, or it may be used to predict the PB vector associated with a given sensory-motor input. As such, these PB units can be understood as replicating some of the mirror functionality. Tani et al. (2004) demonstrate the

utility of the overall architecture in three tasks. First, they show that the architecture enables a robot to learn hand movements by imitation. Second, they demonstrate that it can learn both end-point and cyclic movements. Finally, they illustrate the ability of the architecture to associate word sequences with the corresponding sensory-motor behaviors.

Yamashita & Tani (2008) similarly use a recurrent neural network at the heart of their robot controller but endow it with neurons that have two different timescales. The robot then learns repetitive movements and it is shown that the neurons with the faster timescale encode so-called movement primitives while the neurons with the slower timescale encode the sequencing of these primitives. This enables the robot to create novel behavior sequences by merely adapting the slower neurons. The encoding of different movement primitives within the neural structure also replicates the organization of parietal mirror neurons (Fogassi et al., 2005), which is at the core of other computational models of the mirror system (Chersi et al., 2010; Thill et al., In Press; Thill & Ziemke, 2010).

While the RNNPB architecture encodes behavior as different parametric bias vectors, Demiris & Hayes (2002); Demiris & Johnson (2003) propose an architecture in which every behavior is encoded by a separate module. This architecture combines inverse and forward models, leading to the ability to both recognize and execute actions with the same architecture. Learning is done by imitation, where the current state of the demonstrator is received and fed to all forward modules. These forward modules each predict the next state of the demonstrator based on the behavior they encode. The predicted states are compared with the actual states, resulting in confidence values that a certain behavior is being executed. If the behavior is known (a module produces a high confidence value), the motors are then actuated accordingly. If not, a new behavioral module is created to learn the novel behavior being demonstrated. A somewhat similar model of human motor control, also using multiple forward and inverse models has been proposed by Wolpert & Kawato (2001), with the main difference being that in this work, all models (rather than simply the one with the highest confidence value) contribute to the final motor command (albeit in different amounts).

Finally, Wermter et al. (2003; 2005; 2004) developed a self-organizing architecture which "takes as inputs language, vision and actions . . . [and] . . . is able to associate these so that it can produce or recognize the appropriate action. The architecture either takes a language instruction and produces the behavior or receives the visual input and action at the particular time-step and produces the language representation" (Wermter et al., 2005, cf. Wermter et al., 2004). This architecture was implemented in a wheeled (non-humanoid) robot based on the PeopleBot platform. This robot can thus be seen to "understand" actions by either observing them or from its stored representation related to observing the action. This is therefore an example of a robot control architecture that makes use of embodied representations of actions. In related work on understanding of concepts/language in mirror-neuron-like neural robotic controllers (Wermter et al., 2005) researchers use the insight that language can be grounded in semantic representations derived from sensory-motor input to construct multimodal neural network controllers for the PeopleBot platform that are capable of learning. The robot in this scenario is capable of locating a certain object, navigating towards it and picking it up. A modular associator-based architecture is used to perform these tasks. One module is used for vision, another one for the execution of the motor actions. A third module is used to process linguistic input while an overall associator network combines the inputs from each module.

What all these examples illustrate is that insights from mirror neuron studies (in particular their potential role in grounding higher-level cognition in an agent's sensory-motor experiences) can be useful in robotics. In terms of using insights from embodied cognition,

these are relatively simple examples since the main role of the body in these cases is to enable the grounding of concepts. For instance, a robot would "know" what a grasp is because it can relate it to its own grasping movements via the mirror neurons.

However, from Rob's perspective, there is still substantial work that needs to be done in this direction. In essence, what the field is currently missing are robots that can display higher-level cognition which goes beyond simple toy problems. For example, most of the examples above dealing with learning by imitation understand imitation as reproducing the trajectory of, for instance, the end-effector. However, imitation learning is more complex than that and involves decisions on whether it is really the trajectory that should be copied or the overall goal of the action (see *e.g.* Breazeal & Scassellati, 2002, for a discussion of such issues). Similarly, while there is work on endowing robots with embodied language skills (*e.g.* Wermter et al., 2005), it rarely goes beyond associating verbal labels to sensory-motor experiences (Although see Cangelosi & Riga (2006) for an attempt to build more abstract concepts using linguistic labels for such experiences). Again, while this is a worthwhile exercise, it is not really an example of a robot with true linguistic abilities.

4.2.4 Morphological computation

Closely related to embodiment is the concept of *morphological computation* (Pfeifer & Iida, 2005; Pfeifer et al., 2006). This explicitly takes into account that the body can do more than just provide a grounding for concepts; it can actually, through its very morphology, perform computations which no longer need to be taken care of by the controller itself.

The classic example of this is the passive dynamic walker (McGeer, 1990), a pair of legs that can walk down a slope in a biologically realistic manner with no active controllers (or indeed any form of computation) involved at all. Pfeifer et al. (2006) offer additional examples: the eye of a house-fly is constructed to intrinsically compensate for motion parallax (Franceschini et al., 1992) and a similar design can facilitate the vision of a moving robot. Another example is the "Yokoi hand" ((Yokoi et al., 2004)), whose flexible and elastic material allows it to naturally adapt to an object it is grasping without any need for an external controller to evaluate the shape of the object beforehand.

A completely different type of gripper that fulfils a similar role as the Yokoi hand (namely the ability to grasp an object without the need to evaluate its shape beforehand) is presented by Brown et al. (2010). This gripper is in the shape of a ball filled with a granular material that contracts and hardens when a vacuum is applied. When in contact with an object, it reshapes to fit the object and can lift it. Here, the morphology of the hand goes as far as removing the need for any joints in the gripper, thus significantly reducing the computational requirements associated with it.

From the perspective of Rob's robot, the key insight from the examples given here is that a suitably designed robot can reduce the computations required within the controller itself by offloading them onto the embodiment. At the same time, he also identifies a shortcoming, namely that there are no examples of humanoids that take this into account to a significant degree. Rather, the demonstrations mostly focus on small robots with a very limited behavioral repertoire designed solely to illustrate the utility of the particular embodiment in that specific case. Although Rob finds the approach itself promising, he feels that it still needs to overcome the restrictions in form of the focus on limited behavioral repertoires.

4.2.5 Dynamic Field Theory

Dynamic Field Theory (DFT) is a mathematical framework based on the concepts of Dynamical Systems and the guidelines from Neurophysiology. A field represents a population of neurons and their activations follow continuous responses to external stimuli. Amari (1977) studied the properties of these networks as a model of the activation observed in cortical tissue. Fields have the same structure of a recurrent neural network since their connections can have, depending on the relative location within the network, a local excitation or a global inhibition, Fig. 4.
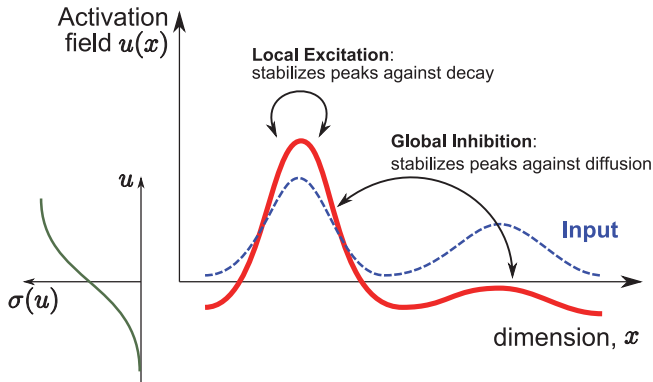

Fig. 4. Typical activations in a dynamics field, from Schöner (2008).

Fields are used to represent perceptual features, motion or cognitive decisions, e.g. position, orientation, color, speed. The dynamics of these fields allow the creation of peaks which are the units of representation in DFT (Schöner, 2008). Different configurations of one or more fields are possible, being the designer responsible for creating a proper connectivity and tuning of parameters. The result of activating this type of network is a continuously adaptive system that responds dynamically to any change coming from external stimuli.

Rob has learned about the different properties and potentials of dynamic fields for using it as part of a robust cognitive architecture. Some of the most attractive features of this approach include the possibility of having a hebbian-type of learning by exploiting the short-term memory features implicit in the dynamics of this algorithm. Long-term memory, decision making mechanism and noise robustness (also implicit in the dynamics of fields), and single-shot learning are all important tools that can and must be included in any cognitive architecture. Several applications modeling experiments on human behavior (Dineva (2005); Johnson et al. (2008); Lowe et al. (2010)) and robotic implementations (Bicho et al. (2000); Erlhagen et al. (2006); Zibner et al. (2011)) have demonstrated DFT's potential.

Nonetheless, from Rob's perspective, the current work with dynamic fields still needs to overcome a number of challenges. Dynamic field controllers are currently designed by hand, through elaborate parameter space explorations, to solve a very specific problem. Although these models nicely illustrate that decision-making can be grounded directly in sensory-motor experiences, their learning ability is limited and any particular model does not generalize well to other tasks, even though modular combinations of different models each solving a particular task seems possible (Johnson et al. (2008); Simmering et al. (2008); Simmering & Spencer (2008)).

4.3 Interim summary 2

To summarize, it seems likely that embodied approaches are the future, both from a theoretical perspective (given insights in the cognitive sciences about the functioning of the human mind) and practically speaking (given the numerous examples of the associated benefits in robotic applications) and thus, this is the approach that Rob would prefer when designing his robot. However, he thinks about the successful applications of symbolic approaches and leaves open the door to possible interactions between both approaches. It could be possible to use dynamic tools to face unexpected circumstances and once that information looks stable let a symbolic algorithm work with it and give some feedback to the body.

There are significant challenges still to overcome for embodied approaches as well. Most of the existing work (while demonstrating that an embodied approach is indeed viable) currently focuses on proof-of-concept solutions to specific problems. These limitations go hand-in-hand with the development of physical platforms to test new models.

5. Discussion

Rob has completed his brief study of the current state of the art in humanoid robotics. Along the way he has found very interesting developments for both hardware and software. In almost all fields he found technology that oftentimes exceeds the abilities found in humans. However, there are several questions that come to his mind: why is it so difficult to create human behavior? Or is it intelligence? Or maybe cognition? From his analysis it seems that the most important parts are there, but still something is missing in the overall picture.

The field of humanoid robotics was formally born when a group of engineers decided to replicate human locomotion. This work is more than 40 years old and we are still unable to replicate the simplest walking gait found in humans. In section 3 Rob found several state of the art platforms that are showing signs of improvement for locomotion. He is confident that in the near future a more robust and stable human-like walking gait will be available for biped platforms. However, he does not feel so positive about other forms of motion and the transitions between them. Rob picks the embodied approach, and more specifically the dynamical systems approach, as the best option for handling instabilities and transitions between stable states such as walking, running, crawling, jumping, etc.

In Rob's opinion, grasping and manipulation do seem to be close to be maturing as well. In hardware terms, several platforms are progressing thanks to the integration of state of the art technology in skin-like materials. The addition of this component to the traditional "naked" robotic hand has improved substantially the results of grasping and manipulation. Rob is convinced that the same strategy should be applied in the lower limbs of his humanoid platform. A large part of controlling balance and motion can be transferred to the interactions between materials and environment as it has been showed by morphological computation in section 4.2.4.

There are however several challenges that worry Rob. He is still not sure about the best way to go regarding his cognitive architecture. He is not worried about the quantity or quality of the information collected through state of the art sensors; his study showed that sensors are not the problem. The problem is how to merge all that information and make some sense out of it. The missing piece seems to be finding the right way of creating and then recovering associations. Finally, a proper way of interacting with humans is needed, e.g. language.

The biggest challenge for the area of humanoid robotics right now is the design and implementation of a large project with the collaboration of an equally large number of researchers. If we continue doing research in pieces, then the development of cognitive

abilities in humanoid robots will be very slow and the use of humanoids in the future very limited. The above mentioned mega-project should have as common goal the construction of a humanoid robot with state of the art sensors and actuators. If anything can be learned from the embodiment approach 4.2, it is that a complete and sophisticated body like this is needed mainly because of the biggest challenge for the future of humanoid robotics, i.e. the development of cognitive abilities through dynamic interactions with unconstrained environments.

6. References

Amari, S.-I. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields, *Biological Cybernetics* 27: 77–87.

Ambrose, R. O., Aldridge, H., Askew, R. S., Burridge, R. R., Bluethmann, W., Diftler, M., Lovchik, C., Magruder, D. & Rehnmark, F. (2000). Robonaut: Nasa's space humanoid, *IEEE Intelligent Systems* 15: 57–63.
    URL: *http://dx.doi.org/10.1109/5254.867913*

Anderson, M. L. (2003). Embodied cognition: a field guide., *Artificial Intelligence* 149: 91–130.

Anybots (2008). Dexter.
    URL: *http://www.anybots.com/*

Bao, Z., McCulloch, I., (Society), S., Staff, S. S., Company, A. C., Staff, A. C. C., Incorporated, C., Staff, C. I., (Firm), S. S. A. & Staff, S. S. A. F. (2009). *Organic Field-Effect Transistors VIII: 3-5 August 2009, San Diego, California, United States*, Proceedings of SPIE–the International Society for Optical Engineering, S P I E-International Society for Optical Engineering.
    URL: *http://books.google.com/books?id=IjGvSgAACAAJ*

Bicho, E., Mallet, P. & Schoner, G. (2000). Target representation on an autonomous vehicle with low-level sensors., *The International Journal of Robotics Research* 19(5): 424–447.
    URL: *http://dx.doi.org/10.1177/02783640022066950*

Bonini, L., Rozzi, S., Serventi, F. U., Simone, L., Ferrari, P. F. & Fogassi, L. (2010). Ventral premotor and inferior parietal cortices make distinct contributions to action organization and intention understanding, *Cerebral Cortex* 20: 1372–1385.

Breazeal, C. & Scassellati, B. (2002). Robots that imitate humans, *Trends in Cognitive Sciences* 6(11): 481–487.

Brown, E., Rodenberg, N., Amend, J., Mozeika, A., Steltz, E., Zakin, M. R., Lipson, H. & Jaeger, H. M. (2010). Universal robotic gripper based on the jamming of granular material, *Proceedings of the National Academy of Sciences USA* 107(44): 18809–18814.

Bullock, I. & Dollar, A. (2011). Classifying human manipulation behavior, *IEEE International Conference on Rehabilitation Robotics, Proceedings of the.*

Cangelosi, A. & Riga, T. (2006). An embodied model for sensorimotor grounding and grounding transfer: experiments with epigenetic robots, *Cognitive Science* 30: 673–689.

Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G. & Rizzolatti, G. (2007). Impairment of actions chains in autism and its possible role in intention understanding, *Proceedings of the National Academy of Sciences USA* 104(45): 17825–17830.

Chersi, F., Thill, S., Ziemke, T. & Borghi, A. M. (2010). Sentence processing: linking language to motor chains, *Frontiers in Neurorobotics* 4(4): DOI:10.3389/fnbot.2010.00004.

Chrisley, R. & Ziemke, T. (2003). *Encyclopedia of Cognitive Science*, Macmillan Publishers, chapter Embodiment, pp. 1102–1108.

Collette, C., Micaelli, A., Andriot, C. & Lemerle, P. (2007). Dynamic balance control of humanoids for multiple grasps and non coplanar frictional contacts, *Humanoid Robots, 2007 7th IEEE-RAS International Conference on*, pp. 81 –88.

Collins, S., Ruina, A., Tedrake, R. & Wisse, M. (2005). Efficient bipedal robots based on passive-dynamic walkers, *Science* 307(5712): 1082–1085.
   URL: *http://www.sciencemag.org/content/307/5712/1082.abstract*

Cui, Q. N., Bachus, L., Knoth, E., O'Neill, W. E. & Paige, G. D. (2008). Eye position and cross-sensory learning both contribute to prism adaptation of auditory space, *in* C. Kennard & R. J. Leigh (eds), *Using Eye Movements as an Experimental Probe of Brain function - A Symposium in Honor of Jean BÃijttner-Ennever*, Vol. 171 of *Progress in Brain Research*, Elsevier, pp. 265 – 270.
   URL: *http://www.sciencedirect.com/science/article/pii/S0079612308006377*

Dahiya, R., Metta, G., Valle, M. & Sandini, G. (2010). Tactile sensing - from humans to humanoids, *Robotics, IEEE Transactions on* 26(1): 1 –20.

Davis, S., Tsagarakis, N. & Caldwell, D. (2008). The initial design and manufacturing process of a low cost hand for the robot icub, *IEEE International Conference on Humanoid Robots*.

Demiris, Y. & Hayes, G. (2002). *Imitation in animals and artefacts*, MIT Press, Cambridge, MA, chapter Imitation as a doual-route process featuring predictive and learning components: A biologically-plausible computational model.

Demiris, Y. & Johnson, M. (2003). Distributed, predictive perception of actions: A biologically inspired robotics architecture for imitation and learning., *Connection Science* 15(4): 231–243.

Dineva, E. (2005). *Dynamical Field Theory of Infant Reaching and its Dependence on Behavioral History and Context.*, PhD thesis, Institut für Neuroinformatik & Intrnational Graduate School for Neuroscience, Ruhr-Universiät-Bochum, Germany.

Erlhagen, W., Mukovski, A. & Bicho, E. (2006). A dynamic model for action understanding and goal-directed imitation, *Brain Research* 1083: 174–188.

Faisal, A., Stout, D., Apel, J. & Bradley, B. (2010). The manipulative complexity of lower paleolithic stone toolmaking, *PLoS ONE* 5(11): e13718.
   URL: *http://dx.doi.org/10.1371/journal.pone.0013718*

Festo (2011).
   URL: *http://www.festo.com/*

Fischer, M. H. (2008). Finger counting habits modulate spatial-numerical associations, *Cortex* 44: 386–392.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F. & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding, *Science* 308: 662–667.

Franceschini, N., Pichon, J. M. & Blanes, C. (1992). From insect vision to robot vision, *Philosophical Transactions of The Royal Society Of London. Series B: Biological Sciences* 337: 283–294.

Gallagher, S. (2005). *How the body shapes the mind*, Oxford University Press, New York.

Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. (1996). Action recognition in the premotor cortex, *Brain Research* 119: 593–609.

Harnad, S. (1990). The symbol grounding problem, *Physica D: Nonlinear Phenomena* 42(1-3): 335–346.

Haykin, S. & Chen, Z. (2005). The cocktail party problem, *Neural Computation* 17(9): 1875–1902.
   URL: *http://www.mitpressjournals.org/doi/abs/10.1162/0899766054322964*

Hickok, G. (2008). Eight problems for the mirror neuron theory of action understanding in monkeys and humans, *Journal of Cognitive Neuroscience* 21(7): 1229–1243.

Hirose, M., Haikawa, Y., Takenaka, T. & Hirai, K. (2001). Development of humanoid robot asimo, *International Conference on Intelligent Robots and Systems*.

Hobbelen, D., de Boer, T. & Wisse, M. (2008). System overview of bipedal robots flame and tulip: Tailor-made for limit cycle walking, *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pp. 2486 –2491.

Ishikawa, M. & Oku, H. (2011). Ishikawa oku laboratory, university of tokyo.
URL: *http://www.k2.t.u-tokyo.ac.jp/*

Iwata, H., Kobashi, S., Aono, T. & Sugano, S. (2005). Design of anthropomorphic 4-dof tactile interaction manipulator with passive joints, *2005 IEEERSJ International Conference on Intelligent Robots and Systems* pp. 1785–1790.
URL: *http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1545261*

Johnson, J. S., Spencer, J. P. & Schï£¡ner, G. (2008). Moving to higher ground: The dynamic field theory and the dynamics of visual cognition, *New Ideas in Psychology* 26(2): 227 – 251. Dynamics and Psychology.
URL: *http://www.sciencedirect.com/science/article/pii/S0732118X07000505*

Kaneko, K., Harada, K., Kanehiro, F., Miyamori, G. & Akachi, K. (2008). Humanoid robot hrp-3, *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pp. 2471 –2478.

Keyrouz, F. (2008). Efficient binaural sound localization for humanoid robots and telepresence applications.

Ko, H. C., Stoykovich, M. P., Song, J., Malyarchuk, V., Choi, W. M., Yu, C.-J., III, J. B. G., Xiao, J., Wang, S., Huang, Y. & Rogers, J. A. (2008). A hemispherical electronic eye camera based on compressible silicon optoelectronics, *Nature* 454: 748–753.

Lipomi, D., Tee, B., Vosgueritchian, M. & Bao, Z. (2011). Stretchable organic solar cells., *Adv Mater* 23(15): 1771–5.

Liu, H., Meusel, P., Seitz, N., Willberg, B., Hirzinger, G., Jin, M., Liu, Y., Wei, R. & Xie, Z. (2007). The modular multisensory dlr-hit-hand, *Mechanism and Machine Theory* 42(5): 612 – 625.
URL: *http://www.sciencedirect.com/science/article/pii/S0094114X0600098X*

Lowe, R., Duran, B. & Ziemke, T. (2010). A dynamic field theoretic model of iowa gambling task performance, *Development and Learning (ICDL), 2010 IEEE 9th International Conference on*, pp. 297 –304.

Mahmoudi, E. (2009). Electronic nose technology and its applications, *Sensors & Transducers* 107: 17–25.
URL: *http://www.sensorsportal.com/HTML/DIGEST/august_09/P_470.pdf*

Mahon, B. Z. & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content, *Journal of Physiology Paris* 102: 59–70.

McGeer, T. (1990). Passive dynamic walking, *The International Journal of Robotics Research* 9(2): 62–82.

Niiyama, R. & Kuniyoshi, Y. (2010). Design principle based on maximum output force profile for a musculoskeletal robot, *Industrial Robot: An International Journal* 37(3).

Oztop, E., Kawato, M. & Arbib, M. A. (2006). Mirror neurons and imitation: A computationally guided review, *Neural Networks* 19: 254–271.

Park, I.-W., Kim, J.-Y., Lee, J. & Oh, J.-H. (2005). Mechanical design of humanoid robot platform khr-3 (kaist humanoid robot 3: Hubo), pp. 321 –326.

Peratech (2011).
    URL: *http://www.peratech.com/*

Petman (2011). Boston dynamics.
    URL: *http://www.bostondynamics.com/*

Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fischer, M. H., McRae, K. & Spivey, M. J. (2011). The mechanics of embodiment: a dialog on embodiment and computational modeling, *Frontiers in Psychology* 2(5).

Pfeifer, R. & Bongard, J. (2007). *How the body shapes the way we think*, MIT Press, Cambridge.

Pfeifer, R. & Iida, F. (2005). Morphological computation: conecting brain, body and environment, *Japanese Scientific Monthly* 58(2): 48–54.

Pfeifer, R., Iida, F. & Gómez, G. (2006). Morphological computing for adaptiver behavior and cognition, *International Congress Series* 1291: 22–29.

Pons, J., Rocon, E., Ceres, R., Reynaerts, D., Saro, B., Levin, S. & Van Moorleghem, W. (2004). The manus-hand dextrous robotics upper limb prosthesis: Mechanical and manipulation aspects, *Autonomous Robots* 16: 143–163. 10.1023/B:AURO.0000016862.38337.f1.
    URL: *http://dx.doi.org/10.1023/B:AURO.0000016862.38337.f1*

Righetti, L. & Ijspeert, A. J. (2006). Design methodologies for central pattern generators: an application to crawling humanoids, *Proceedings of Robotics: Science and Systems*, pp. 191–198.

Rizzolatti, G. & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations, *Nature Reviews Neuroscience* 11(4): 264–274.

Ruesch, J., Lopes, M., Bernardino, A., Hornstein, J., Santos-Victor, J. & Pfeifer, R. (2008). Multimodal saliency-based bottom-up attention a framework for the humanoid robot icub, *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, pp. 962 –967.

Schöner, G. (2008). *Toward a New Grand Theory of Development? Connectionism and Dynamic Systems Theory Re-Considered.*, Oxford University Press, New York, chapter Development as Change of System Dynamics: Stability, Instability, and Emergence., p. In press.

Shadow Robot Company (2003). Design of a dextrous hand for advanced CLAWAR applications, *Climbing and Walking Robots and the Supporting Technologies for Mobile Machines*, pp. 691–698.
    URL: *http://www.shadowrobot.com/*

Shotton, J. & Sharp, T. (2011). Real-time human pose recognition in parts from single depth images, *statosuedu* 2: (to appear).
    URL: *http://www.stat.osu.edu/ dmsl/BodyPartRecognition.pdf*

Simmering, V. R., Schutte, A. R. & Spencer, J. P. (2008). Generalizing the dynamic field theory of spatial cognition across real and developmental time scales, *Brain Research* 1202: 68 – 86. Computational Cognitive Neuroscience.
    URL: *http://www.sciencedirect.com/science/article/pii/S0006899307013807*

Simmering, V. R. & Spencer, J. P. (2008). Generality with specificity: the dynamic field theory generalizes across tasks and time scales, *Developmental Science* 11(4): 541–555.
    URL: *http://dx.doi.org/10.1111/j.1467-7687.2008.00700.x*

Strack, F., Marin, L. & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis, *Journal of Personality and Social Psychology* 54: 768–777.

Takei, K., Takahashi, T., Ho, J. C., Ko, H., Gillies, A. G., Leu, P. W., Fearing, R. S. & Javey, A. (2010). Nanowire active-matrix circuitry for low-voltage macroscale artificial skin, *Nature Materials* 9(10): 821âĂŞ826.

Tani, J., Ito, M. & Sugita, Y. (2004). Self-organization of distributedly represented multiple behavior schemata in a mirror system: reviews of robot experiments using RNNPB, *Neural Networks* 17: 1273–1289.

Thill, S. (2011). Considerations for a neuroscience-inspired approach to the design of artificial intelligent systems, *in* J. Schmidhuber, K. R. Thórisson & M. Looks (eds), *Proceedings of the Fourth Conference on Artificial General Intelligence, LNAI 6830*, Springer, Heidelberg, pp. 247–254.

Thill, S., Svensson, H. & Ziemke, T. (In Press). Modeling the development of goal-specificity in mirror neurons, *Cognitive Computation* .

Thill, S. & Ziemke, T. (2010). Learning new motion primitives in the mirror neuron system: A self-organising computational model., *in* S Doncieux et al (ed.), *SAB 2010, LNAI 6226*, Springer, Heidelberg, pp. 413–423.

Traver, J. & Bernardino, A. (2010). A review of log-polar imaging for visual perception in robotics., *Robotics and Autonomous Systems* pp. 378–398.

Tsakarakis, N., Metta, G., Sandini, G., Vernon, D., Beira, R., Becchi, F., Righetti, L., Santos-Victor, J., Ijspeert, A., Carrozza, M. & Caldwell, D. (2007). icub - the design and realization of an open humanoid platform for cognitive and neuroscience research, *Journal of Advanced Robotics, Special Issue on Robotic platforms for Research in Neuroscience* 21(10): 1151–1175.

Ueno, T. & Oda, M. (2009). Robotic hand developed for both space missions on the international space station and commercial applications on the ground, *Proceedings of the 2009 IEEE/RSJ international conference on Intelligent robots and systems*, IROS'09, IEEE Press, Piscataway, NJ, USA, pp. 1791–1796.
URL: *http://portal.acm.org/citation.cfm?id=1733023.1733039*

Umiltà, M. A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., Jezzini, A., Gallese, V. & Rizzolatti, G. (2008). When pliers become fingers in the monkey motor system, *Proceedings of the National Academy of Sciences USA* 105(6): 2209–2213.

Vernon, D. (2008). Cognitive vision: The case for embodied perception, *Image Vision Comput.* 26(1): 127–140.

Vukobratovic, M. & Borovac, B. (2004). Zero-moment point - thirty five years of its life, *International Journal of Humanoid Robotics* 1: 157–173.

Wermter, S., Elshaw, M. & Farrand, S. (2003). A modular approach to self-organization of robot control based on language instruction, *Connection Science* 15(2-3): 73–94.

Wermter, S., Weber, C. & Elshaw, M. (2005). Associative neural models for biomimetic multimodal learning in a mirror-neuron based robot, *Progess in Neural Processing* 16: 31–46.

Wermter, S., Weber, C., Elshaw, M., Panchev, C., Erwin, H. & Pulvermüller, F. (2004). Towards multimodal neural robot learning, *Robotics and Autonomous Systems* 47(2-3): 171–175.

Wolpert, D. & Kawato, M. (2001). Multiple paired forward and inverse models for motor control, *Neural Networks* 11: 1317–1329.

Yamashita, Y. & Tani, J. (2008). Emergence of functional hierarchy in a multiple timescale
        neural network model: a humanoid robot experiment, *PLoS Computational Biology*
        4(11): 1–18.
Yokoi, H., Arieta, A. H., Katoh, R., Yu, W., Watanabe, I. & Maruishi, M. (2004). Mutual
        adaption in a prosthetics application, *in* F. Iida, R. Pfeifer, L. Steels & Y. Kuniyoshi
        (eds), *Embodied Artificial Intelligence, LNAI 3139*, Springer, Heidelberg, pp. 146–159.
Zibner, S., Faubel, C., Iossifidis, I. & Schoner, G. (2011). Dynamic neural fields as building
        blocks of a cortex-inspired architecture for robotic scene representation, *Autonomous
        Mental Development, IEEE Transactions on* 3(1): 74 –91.