Towards Biomimetic Neural Learning for Intelligent Robots

Stefan Wermter[1], Günther Palm[2], Cornelius Weber[1], Mark Elshaw[1]

[1] Hybrid Intelligent Systems
University of Sunderland
School of Computing and Technology
St Peter's Way, Sunderland, SR6 0DD, UK
Email: [Stefan.Wermter][Cornelius.Weber][Mark.Elshaw]@sunderland.ac.uk
www.his.sunderland.ac.uk
[2] Neuroinformatics
University of Ulm
Oberer Eselsberg
D-89069 Ulm
Germany
Email: palm@neuro.informatik.uni-ulm.de

Abstract. We present a brief overview of the chapters in this book that relate to the development of intelligent robotic systems that are inspired by neuroscience concepts. Firstly, we concentrate on the research of the MirrorBot project which focuses on biomimetic multimodal learning in a mirror neuron-based robot. This project has made significant developments in biologically inspired neural models using inspiration from the mirror neuron system and modular cerebral cortex organisation of actions for use in an intelligent robot within an extended 'pick and place' type scenario. The hypothesis under investigation in the MirrorBot project is whether a mirror neuron-based cell assembly model can produce a life-like perception system for actions. Various models were developed based on principles such as cell assemblies, associative neural networks, and Hebbian-type learning in order to associate vision, language and motor concepts. Furthermore, we introduce the chapters of this book from other researchers who attended our AI-workshop on NeuroBotics.

1 Introduction

Many classical robot systems ignore biological inspiration and so do not perform in a robust learned manner. This is reflected in most of the conventional approaches to the programming of (semi-) autonomous robots: Many details of the program have to be reprogrammed and fine-tuned by hand even for slight changes in the application, which is time consuming and error-prone. Hence, there is a need for a new type of computation that is able to take its inspiration from neuroscience and perform in an intelligent adaptive manner to create biomimetic robotic systems. Many researchers including the contributors to this book hold that by taking inspiration from biological systems that would allow the development of autonomous robots with more robust functionality than is possible with current robots. An additional benefit of biomimetic robots is that they can provide an indication of how the biological systems actually could work in order to provide feedback to neuroscientists. In order to indicate the progress made towards Biomimetic Neural Learning for Intelligent Robots this book is split into two parts.

In the first part we present some of the research findings from the biomimetic multimodal learning in a mirror neuron-based robot (MirrorBot) project. The aim of this EU FET project was to develop models for biomimetic multimodal learning using a mirror neuron-based robot to investigate an extended 'pick and place' scenario. This task involves the search for objects and integrates multimodal sensory inputs to plan and guide behaviour. These perceptual processes are examined using models of cortical assemblies and mirror neurons to explore the emergence of semantic representations of actions, percepts, language and concepts in a MirrorBot, a biologically-inspired neural robot. The hypothesis under investigation and focus of this book is whether a mirror neuron-based cell assembly model will produce a life-like perception system for actions. The MirrorBot project combines leading researchers in the areas of neuroscience and computational modeling from the University of Sunderland, Parma and Ulm, INRIA Lorraine/LORIA-CNRS and Cognition and Brain Sciences Unit, Cambridge. The findings of the neuroscience partners form the basis of the computational models that are used in the development of the robotic system. The neuroscience partners concentrate on two cerebral cortex systems by examining how humans process and represent different word categories and the mirror neuron system.

The extended 'pick and place' scenario involves the MirrorBot neural robot assistant being positioned between two tables that have multiple objects positioned on them and is required to perform various behaviours on objects based on a human verbal instruction. The robot takes in three or four word instructions that contain an actor, action and object such as 'bot pick plum' or 'bot show brown nut'. The instructional grammar developed for the MirrorBot contains approximately 50 words with the actor being the 'bot'. The actions that are performed are divided into those that are performed by the hand, leg or head. For instance, the action performed by the hand include 'pick', 'put' and 'lift', the leg actions include 'go' and 'move' and the head actions include 'show' and 'turn-head'. The objects include natural objects such as 'orange', 'nut' and artefact objects such as 'ball' and 'cup'. In order to perform the appropriate behaviours the robot assistant using neural learning must perform such diverse activities as language recognition, object localization, object recognition, attention, grasping actions, docking, table localization, navigation, wandering and camera positioning.

In the second part of the book we provide chapters from researchers in the field of biomimetic robotic neural learning systems who attended the AI-Workshop on NeuroBotics. The aim of this workshop and hence of this book is to contribute to robotic systems which use methods of learning or artificial neural networks and/or are inspired by observations and results in neuroscience and animal behaviour. These chapters were selected to give an indication of the diversity of the research that is being performed into biomimetic robotic learning and to provide a broader perspective on neural robotics. For instance, chapters will consider the development of a virtual platform for modeling biomimetic robots, a robotic arm, robot recognition in RoboCup, sensory motor control of robot limbs and navigation. These models are utilised by both robot simulators and actual robots and make use of neural approaches that are both supervised and unsupervised.

2 Modular Cerebral Cortex Organisation of Actions: Neurocognitive Evidence for the MirrorBot Project

Neuroscience evidence reflected in the development of the MirrorBot biomimetic robotic systems comes from research at Cambridge related to how words are processed and represented in the cerebral cortex based on neurocognitive experiments of Pulvermüller. Accordingly, words are represented and processed using Hebbian learning, synfire chains and by use of semantic features. Hebbian learning supports the basis of higher cognitive behaviour through a simple synaptic approach based on cell assemblies for cortical processing [27, 30, 28, 29]. Cell assemblies rely on a connectivity structure between neurons that support one another's firing and hence have a greater probability of being co-activated in a reliable fashion [43, 41, 47, 23]. Synfire chains are formed from the spatiotemporal firing patterns of different associated cell assemblies and rely on the activation of one or more cell assemblies to activate the next assembly in the chain [27, 41, 18]. Hence, neurocognitive evidence on word representation and processing in the cerebral cortex suggests that cognitive representations are distributed among cortical neuronal populations [29, 33, 27]. The word meaning is critical for determining the cortical populations that are activated for the cognitive representation task.

When looking at the web of cell assemblies which process and represent particular word types Pulvermüller [27] notes that activation is found in both hemispheres of the cerebral cortex for content words. Semantic word categories elicit different activity patterns in the fronto-central areas of the cortex, in the areas where body actions are known to be processed [40, 11]. Perception words are represented by assemblies in the perisylvian cortex and posterior cortex [27, 31] and nouns related to animals activate the inferior temporal or occipital cortices [28, 27, 29].

Emotional words are felt to activate the amygdala and cells in the limbic system more than words associated with tools and their manipulation [26]. The link between the assemblies in these two regions is achieved through the amygdala and frontal septum [27]. For action words that involve moving ones own body the perisylvian cell assembly is also associated with assemblies in the motor, premotor and prefrontal cortices [27, 30]. For content words the semantic features that influence the cell assemblies come from various modalities and include the complexity of activity performed, facial expression or sound, the type and number of muscles involved, the colour of the stimulus, the object complexity and movement involved, the tool used, and whether the person can see itself doing this activity. The combination of these characteristics into a single depiction is produced by pathways linking sensory information from diverse modalities to the same neurons. For objects the semantic features represented by cell assemblies typically relate to their colour, smell and shape. If a word is repeatedly presented with a stimulus the depiction of this stimulus is incorporated into the one for the word to produce a new semantic feature. In general, words are depicted via regions historically known as language regions and additional regions connected with the words semantics.

Concerning a division between action related and non-action related words [33], Pulvermüller states that there is a finer-grained grounding of language instruction in actions. This produces a division of the representation in the cerebral cortex based on the part of the body that performs that action between leg, head and hand [11, 29, 30, 27, 12]. It is well known that there is a division in the motor cortex between the regions that perform head/face, hand/arm and leg actions [25]. For instance, the region of the motor cortex that controls face movement is found in the inferior precentral gyrus, hand and arm in the middle region of the precentral gyrus and the leg actions are located in the dorsomedial area [29, 30]. Given the difference in the regions of the cortex that are responsible for performing actions it is also stated by Pulvermüller that a similar difference can be identified when representing action verbs and so grounding language instructions in actions based on the part of the body that performs the action [29].

Pulvermüller and his colleagues have performed various experiments [28, 12, 11, 13, 29, 40] on cerebral cortex processing of action verbs to test their hypothesis on the representation of action verbs based on the body part that performs. These include experiments where (i) different groups of subjects are given leg-, arm- and face-related action verbs and pseudo-words and asked to state whether they are a word; (ii) subjects are asked to use a rating system to answer questions on the cognitive processes a word arouses; (iii) subjects rank words based on whether they are leg-, arm- or head-related; and (iv) there is a comparison between hearing walk- and talk-type verbs. In these experiments EEG electrodes are positioned at various points along the scalp to produce recordings of cerebral cortex activation. From these experiments areas are identified where the activation is the same for all action verbs and more importantly are different depending on the action verbs based on the body parts they relate to.

Differences between the three types of action verbs based on the body parts were observed by Pulvermüller and his colleagues. They found a greater activation for face-words in the frontal-lateral regions of the left hemisphere close to the premotor cortex associated with face and head. For face- and leg-related action verbs there are different regions along the motor strip that are identified to process verbs from these two verb categories. Leg-type words produce greater activation in the cortical region of the cerebral cortex used to produce leg actions and for the face-words there is greater activation in the inferior regions near to the face region of the motor cortex [32]. It is found that hand-related words are located in more lateral regions of the cortex than leg-words. Consistent with the somatotopy of the motor and premotor cortex [25], leg-words elicited greater activation in the central cerebral cortex region around the vertex, with face-words activating the inferior-frontal areas, thereby suggesting that the relevant body part representations are differentially activated when action words are being comprehended.

In addition the average response time for lexical decisions is faster for face-associated words than for arm-associated words, and the arm-associated words are faster than leg ones. There is also greater activation in the right parieto-occipital areas for arm- and leg-words relative to head words. The evidence of these experiments points to the word semantics being represented in different parts of the cerebral cortex in a systematic way. Particularly the representation of the word is related to the actual motor and premotor regions of the cerebral cortex that perform the action.

3 Mirror Neuron System Inspiration for MirrorBot Project

Research at Parma has provided a great deal of evidence on the mirror neuron system that inspired the robotic research for the MirrorBot project. Rizzolatti and co-workers [35, 8] found that neurons located in the rostral region of a primate's inferior frontal cortex area, the F5 area, are activated by the movement of the hand, mouth or both. These neurons fire as a result of the action, but not of the isolated movements that make up this action. The recognition of motor actions comes from the presence of a goal and so the motor system does not solely control movements [9, 37]. Hence, what turns a set of movement into an action is the goal and holding the belief that performing the movements will achieve a specific goal [1]. The F5 neurons are organised into diverse categories based on the actions that cause them to fire, which are 'grasping', 'holding' and 'tearing' [34, 9].

Certain grasping-related neurons fire when grasping an object whether it be performed by the hand, mouth or both [7]. This supports both the view that these neurons do not represent the motor action but the actual goal of performing the grasping task. Within area F5 there are two types of neuron: the first known as canonical neurons only respond to the performing of the action and the second mirror neurons that respond not only when performing an action but also when seeing or hearing the action performed [17, 36, 34]. Hence, the mirror neuron system produces a neural representation that is identical for the performance and recognition of the action [1].

These mirror neurons are typically found in area F5c and do not fire in response to the presence of the object or mimicking of the action. Mirror neurons required the action to interact with the actual object. They respond not only to the aim of the action but also how the action is carried out [44]. However, as shown by Umilta et al. 2001 [44] an understanding of an invisible present object causes the activation of the mirror neurons if the hand reaches for the object in the appropriate manner. This is achieved when they are first shown the action being performed completely visible and then with the hand-object interaction hidden. As the performance and recognition of an action causes activation in the premotor areas which is responsible for the hand movements when observing the action there is a set of mechanisms that prevent the behaviour being mimicked. The mirror neuron system indicates that the motor cortex is not only involved in the production of actions but in the action understanding from perceptual information [36] and so the observer has the same internal representation of action as the actor [44].

In this book Gallese [6] considers an exciting extension to findings related to the mirror neuron system based on the system providing an understanding of the emotional state of the performer for the observer. We do not exist independent of the actions, emotions and sensations of others as we understand the intentions of others. This indicates that as well as recognising an action the mirror neuron system has a role in predicting the consequences of what is being performed. Furthermore by allocating intentions to the actions monkeys and humans are able to use the mirror neurons to aid social interactions. This is achieved through the mirror neuron system providing intention to the motor sequence to identify further goals from this sequence. Hence, Gallese [6] in this chapter notes that we do not just see and recognise an action using the mirror neuron system but by using this system we also associate emotions and sensations to this observed behaviour. This occurs as if the observer is performing a similar action and feeling the same feelings and sensations. This offers a form of mind reading by the observed by attaching intentions to the behaviour. Hence, this points to the ability through embodied simulation to gain insight into the minds of others. Although this does not account for all social cognition.

It is observed that mirror neurons in humans are also excited by both the performance and observation of an action [9]. The F5 area in primates corresponds to various cortical areas in humans including the left superior temporal sulcus, the left inferior parietal lobule and the anterior region of Broca's area. The association of mirror neurons with Broca's area in humans and F5 in primates provides an indication that mirror neurons might have evolved in humans into the language system [34]. The role of the mirror neuron system in language can be seen from the findings of Pulvermüller [30, 11] in that processing and representation of words includes the activation of some of the same regions as those that are found to perform the action. The ability in the first instance to recognise an action is required for the development of a communication system between members of a group and finally for an elaborate language system [17, 34]. The concept of the mirror neuron system being the foundation of the language system directs the multimodal models developed as part of the MirrorBot project.

4 Computational Models in the MirrorBot Project

The neuroscience findings related to word processing and representation based on the findings of research at Cambridge and the mirror neuron system from research at Parma formed the basis for various computational robotic models in the MirrorBot project. In this book we incorporate chapters that provide some of these models that are able to perform language sequence detection [16, 18], spatial visual attention [46, 4], auditory processing [22], navigation [3] and language based multimodal input integration for robot control and multimodal information fusion [18, 4, 20, 48, 24]. These models are a significant contribution to the field of biomimetic neural learning for intelligent robots as they offer brain-inspired robotic performance that is able to produce behaviour based on the fusion of sensory data from multiple sources.

The research of Knoblauch and Pulvermüller [16] described in this book relates to the use of a computational system to consider if word sequences are grammatically correct and so perform sequence detection. A particularly important feature of language is its syntactic structure. For a robot to be able to perform language processing in a biomimetic manner it should be able to distinguish between grammatically correct and incorrect word sequences, categorise words into syntactic classes and produce rules. The model of Knoblauch and Pulvermüller [16] incorporates a biologically realistic element in that it uses numerous sequence detectors to show that associative Hebb-like learning is able to identify word sequences, produces neural representations of grammatical structures, and links sequence detectors into neural assemblies that provides a biological basis of syntactic rule knowledge. The approach consists of two populations of neuron, the WW population for word webs and population SD for sequence detectors. Each neuron is based on leaky-integrate units. The model was found to create auto-associative substitution learning and generalized sequential order to new examples and achieves the learning of putative neural correlate of syntactical rules.

Vitay et al. [46] have developed a distributed model that allows to sequentially focus salient targets on a real-world image. This computational model relies on dynamical lateral-inhibition interactions within different neural maps organized according to a biologically-inspired architecture. Despite the localized computations, the global emergent behaviour mimics the serial mechanism of attention-switching in the visual domain. Attention is understood here as the ability to focus a given stimulus despite noise and distractors, what is represented here by a localized group of activated neurons. The task is to sequentially move this bubble of activity on the different salient targets in the image. They use a preprocessed representation of the image according to task requirements: objects potentially interesting are made artificially salient by enhancing their visual representation. Three mechanisms are involved to achieve that task: first a mechanism allowing to focus a given stimulus despite noise and competing stimuli; second a switching mechanism that can inhibit at demand the currently focused target to let the first mechanism focus another location; third a working memory system that remembers previously focused locations to avoid coming back to a previously inspected object. The cooperation between these different mechanisms is not sequential but totally dynamic and distributed: no need for a central executive that would control the timing between the functionally different systems, the inner dynamics of the neurons make the work. This model has been successfully implemented on a robotic platform.

The chapter by Murray et al. [22] offers a biomimetically inspired hybrid architecture that uses cross-correlation and recurrent neural networks for acoustic tracking in robots. This research is motivated by gaining an understanding of how the mammalian auditory system tracks sound sources and how to model the mechanisms of the auditory cortex to enable acoustic localisation. The system is based on certain concepts from the auditory system and by using a recurrent neural network can dynamically track a sound as it changes azimuthally in the environment. The first stage in the model determines the azimuth of the sound source from within the environment by using Cross-Correlation and provides this angle to the neural predictor to predict the next angle in the sequence with the use of the recurrent neural network. To train the recurrent neural network to recognize various speeds, a separate training sub-group was created for each individual speed. This is to ensure that the network learns the correct temporal sequence it needs to recognize and provide prediction for the speeds. It has been shown that within the brain there is short term memory to perform such prediction tasks and in order to forecast the trajectory of an object it is required that previous positions are remembered to establish predicitons.

The navigation approach of Chokshi et al. [3] is based on modeling the place cells by using self-organising maps. The overall aim of this approach is to localise a robot using two locations based on visual stimulus. An internal representation of the world is produced by the robot in an unsupervised manner. The self-organising maps receive visual images that are used to produce internal representation of the environment that act like place codes using landmarks. Localisation by the robot is achieved by a particular position being associated with a specific active neuron. An overall architecture has been developed that uses modules to do diverse operations such as visual information derivation and motor control. The localisation was performed using a Khepera robot in an environment divided into 4 sections. These sections were also divided into squares that were used to determine the place cell error. Landmarks of different coloured cubes and pyramids were positioned at the edge of the environment. Each square represented a place cell, with the training and testing involving the images associated with each cell. An interesting finding of this study was that as the robot approached a specific landmark it was found that the appropriate place cell in the self-organising map output layer had great activation and once this robot leaves the landmark the activation reduces until it reaches 0. Clustering was also seen for landmarks that were close together and distinct landmarks were positioned further apart on the self-organising map.

The first language model based on multimodal inputs from the MirrorBot project considered in this book, is that of Markert et al. [18] who have developed an approach that through associative memories and sparse distributed representations can associate words with objects and characteristics of the object and actions. The approach enables a robot to process language instructions in a manner that is neurobiologically inspired using cell assemblies. The fundamental concept behind the cortical model developed is the use of cortical regions for different properties of an entity. Hence, a feature of the cortical model is the combining of visual, tactile, auditory language, goal and motor regions. Each of the regions is implemented using a spike counter architecture. Looking at regional activations it is possible to describe how inputs from multiple modalities are combined to create an action.

This language model using multimodal inputs is also used by Fay et al [4] who developed a neurobiologically plausible robotic system that combines visual attention, object recognition, language and action processing using neural associative memory. This involves finding and pointing with the camera at a specific fruit or object in a complex scene based on a spoken instruction. This requires the understanding of the instruction, relating the noun with a specific object that is recognised using the camera and coordinating motor output with planning and sensory information processing. The kinds of spoken commands that the robot is able to parse include 'bot show plum and 'bot put apple to yellow cup'. In the architecture preprocessing involves extracting features from the auditory and the visual input chosen by attention control. The cortical model used to perform speech recognition, language processing, action planning and object recognition consists of various neural networks such as radial basis networks and associator networks. A speech recogniser is used to receive the language instruction which is checked for syntactic consistency. This work is closely related to the work by Knoblauch and Pulvermüller described above. If the word sequence is syntactically correct the global goal is divided into a sequence of subgoals whose solution fulfills the overall goal. Object recognition is performed by a hierarchy of radial-basis-function networks which divide a complex recognition task into various less complex tasks. The model is able to associate a restricted set of objects with sentence like language instructions by associating the noun with properties of the object such as colour and actions. The model is of particular significance to this book as it shows how data from diverse modalities of language and vision can be brought together to perform actions.

Wermter et al. [48] produce two architectures that are able to successfully combine the three input modalities of high-level vision, language instructions and motor directions to produce simulated robot behaviour. The flat multimodal architecture uses a Helmholtz machine and receives the three modalities at the same time to learn to perform and recognise the three behaviours of 'go, 'pick' and 'lift' . The hierarchical architecture at the lower-level uses a Helmholtz machine and at the upper-level a SOM to perform feature binding. These multimodal architectures are neuroscience-inspired by using concepts from action verb processing based on neurocognitive evidence of Pulvermüller and specifically features of the mirror neuron system. The architectures are able to display certain features of the mirror neuron system which is a valuable development as the activation patterns for both the performance of the action and its recognition are close. Furthermore, we are able to indicate the role played by the mirror neuron system in language processing. A particular interesting finding obtained with the hierarchical architecture is that certain neurons in the Helmholtz machine layer of the hierarchical model respond only to the motor input and so act like the canonical neurons in area F5 and others to the visual stimulus and so are analogous to the mirror neurons. Furthermore, the lower-level Helmholtz machine is analogous to area F5 of the primate cortex and the SOM area to area F6. The F5 area contains neurons that can produce a number of different grasping activities. F6 performs as a switch, facilitating or suppressing the effects of F5 unit activations so that only the required units in the F5 region are activated to perform or recognise the required action.

An additional multimodal model from the MirrorBot project described in this book is that of Ménard et al. [20] which is a self-organising approach inspired by cortical maps. When an input is given to the map, a distributed activation pattern appears. From this a small group of units is selected by mutual inhibition that contains the most active units. Unlike Kohonens self-organising maps where this decision is made based on a global winner-takes-all approach, the decision is based on a numerical distribution process. The fundamental processing element of the Biologically-Inspired Joint Associative MAps (BIJAMA) model is a disk-shaped map consisting of identical processing units. The global behaviour of the map incorporates adaptive matching procedures and competitive learning. Ménard et al. [20] use several self-organising maps that are linked to one another to achieve cooperate processing by achieving word-action association. This multi-associative model is used to associate multimodalities of language and action and is able to deal with an association between modalities that is not one-one. The main issue associated with this approach is that learning in a map is dependent on the other maps, so that the inter-map connectivity biases the convergence to a certain state. The model is able to organise word representation in such a way that for instance a body word is associated with a body action.

Panchev [24] also developed a multimodal language processing model related to the MirrorBot project. This model uses a spiking neural model that is able to recognise a human instruction and then produce robot actions. Learning is achieved through leaky Integrate-And-Fire neurons that have active dendrites and dynamic synapses. Using spiking neurons the overall aim of this research is to model the primary sensory regions, higher cognitive functional regions and motor control regions. Hence, in this architecture there are modules that are able to recognise single word instructions, recognise objects based on colour and shape and a control system for navigation. The model uses a working memory that is based on oscillatory activity of neural assemblies from the diverse modalities. As this is one of the first robot control models that is based on spiking neurons it offers the opportunity to consider new behaviours and computational experiments that could be compared with the activity identified in the brain. This research is a significant contribution to the MirrorBot project as it shows the use of spiking neurons for spatiotemporal data processing. It is felt that as this model is able to approximate current neuroscience evidence it could directly address future neuroscience studies in the area of multimodal language processing. A future direction of this work is to consider the incorporation of goal behaviours by having certain objects more attractive than others.

5 Biomimetic Cognitive Behaviour in Neural Robots

We now turn to the chapters in the book that relate to research into biomimetic robots outside the MirrorBot project. The chapters summarised in this section show the diversity necessary to build biomimetic intelligent systems.

Reinforcement and reward based learning has proved a successful technique in biomimetic robots. In this book there are four chapter related to this technique from Jasso and Triesch [15], Hafner and Kaplan [10], Sung et al. [42] and Sheynikhovich et al. [39]. Jasso and Triesch [15] who consider the development of a virtual reality platform that is useful for biomimetic robots as it can be used to model cognitive behaviour. This environment allows the examination how cognitive skills are developed as it is now understood that these skills can be learned through the interaction with the parent and the environment in changeable social settings. This learning relates to visual activities such as gaze and point following and shared attention skills. Usually the models incorporate a single child and a parent. The environment is a room that contains objects and furniture and virtual agents that receive images from their camera. These images are processed and used to influence the behaviour of the agent. The chapter shows that the platform is suitable for modeling how gaze following emerges through infant-parent interactions. Gaze following is the capacity to alter ones own attention to an object that is the attention of another person. The environment consists of a living room containing toys and furniture and contains a parent agent and child agents. The child agents use reinforcement learning to alter its gaze to that of the parent based on a reward.

Hafner and Kaplan [10] in this book present research on biomimetic robots that are able to learn and understand pointing gestures from each other. Using a simple feature-based neural approach it is possible to achieve discrimination between left and right pointing gestures. The model is based on reward mechanisms and is implemented on two AIBO robot dogs. The adult robot is positioned pointing to an object using its left or right front leg and the child robot is positioned watching it. From the pointing gesture, the child robot learns to guess the direction of the object the adult robot is attending to and starts searching for it. The experiment used different viewing angles and distances between the two robots as well as different lighting conditions. This is a first step in order to bootstrap shared communication systems between robots by attention detection and manipualtion.

A further learning approach for biomimetic robots is that of Sung et al. [42] who use grid based function approximators for reinforcement learning. The approach uses techniques gained from active learning to achieve active data acquisition and make use of Q-learning methods that incorporate piecewise linear grid-based approximators. A feature of active learning is active data acquisition with algorithms being developed to reduce the effort required to produce training data. The learning algorithm that relates piecewise linear based approximators to reinforcement learning consists of two components. The first component is used for data acquisition for learning and the second carries out the learning. The suitability of the approach is tested on the 'Mountain-Car' problem which is typically used to evaluate reinforcement learning approaches. When doing so it was found that the number of required state transitions during learning was reduced. It is anticipated that this approach will be applicable to reinforcement learning for a real-world problem.

As navigation is such an important activity for biomimetic robots this book includes a second chapter on this by Sheynikhovich et al. [39]. Their biologically inspired model is based on rodent navigation. The model is multimodal in that it combines visual inputs from images and odometer readings to produce firing in artificial place cells using Hebbian synapses. By using reinforcement type learning the model is able to recreate behaviours and actual neurophysiological readings from the rodent. Although the model starts with no knowledge of the environment learning occurs through populations of place cells as the artificial rodent interacts with the environment. In the model visual information is correlated with odometer data related to the rotation and displacement using Hebbian learning in order that ambiguity in the visual data is resolved using the odometer readings. The model was implemented on a simulator and a mobile Khephera robot and is able to achieve similar performance to animals. This was seen when the robot learned the navigational task of reaching a hidden platform from random positions in the environment. In training the robot was given reward each time it found the platform and was able overtime to reduce the number of steps required to find the platform.

With regards to biomimetic learning the approach of Bach [2] looks at using distributed and localised representations to achieve learning and planning. To perform plan-based control there is a need to have a localist representation of the objects and events in the model of the world. In this approach compositional hierarchies implemented using MicroPsi node nets are used as a form of executable semantic networks that are seen as knowledge-based artificial neural networks. By using MicroPsi node nets it is possible to achieve backpropagation learning and symbolic plan representations. MicroPsi agents have a group of motivational variables that determine demands that direct how the agent performs. In the first instance, the agent does not know what actions will fulfill their desires and so performs trial-and-error actions. When the action is felt to have a positive impact on the demand a link is established between them. Experiments are performed by using a simulated environment that provides the agents with resources and dangers.

Folgheraiter and Gini [5] have developed an artificial robotic arm that replicates the functionality and structure of the human arm. In order to test the control system architecture the arm was developed with a spherical joint with 3 degrees of freedom, and an elbow with 1 degree of freedom. The control is arranged in a modular-hierarchical manner that has three levels: the lower level replicates the spinal reflexes that is used to control artificial muscle activities; the middle level produces the required arm movement trajectories; and at the higher level the circuits in the cerebral cortex and the cerebellum are found to control the path generator operation. The model uses a multilayer perceptron in order to solve the inverse kinematics problem as it is possible to train the network from actual readings from the human arm. In order to model the reflex behaviours a simplified model of the human spinal cord was used that concentrates on modeling membrane potential instead of spiking behaviour of the neurons. By using light materials it is possible to include this arm into a humanoid robot.

The next chapter in this book includes a model of human-like controlling of a biped robot [38]. Humanoid robots require complex design and complicated mathematical models. Scarfogliero et al. [38] demonstrate that a Light Adaptive-Reactive biped is simple, cheap and effective and is able to model the human lower limb. By use of this model it is possible to understand how humans walk and how this might be incorporated into a robot. The model is able to alter joint stiffness for position-control by use of servo motors. Scarfogliero et al. [38] devised a motor device based on torsional spring and damper to recreate the elastic characteristics of muscles and tendons. By using this approach it is possible to have good shock resistance and to determine the external load. As the robot uses location and velocity feedback it is able to perform a fine position operation even though it has no a-priori knowledge of external load.

The chapter in this book by Meng and Lee [21] considers the production of a biologically plausible novelty habituation model based on topological mapping for sensory-motor learning. Meng and Lee [21] examine embedded developmental learning algorithms by using a robotic system made up of two arms that are fitted with two-fingered gripper and a pan/tilt head that includes a colour camera. This robot is built in such away that it recreates the positioning of the head and arms of a child. Novelty and habituation are fundamental for early learning by assisting a system to examine new events/places while still monitoring the current state to gain experience for the full environment. The chapter considers the problem of sensory-motor control of limbs based on a childs arm movements during the first three months of life. Learning is achieved through a hierarchical mapping arrangement which consists of fields of diverse sizes and overlap at diverse mapping levels. The fields include local information such as sensory data and data on movement from the motor and stimulus data. The maps contain units known as fields that are receptive regions. In the experiments various parameters are considered such as the condition of the environment, field extent and habituation variables. By using the concepts from novelty and habitation as the basis of early robot learning it is possible to learn sensory-motor coordination skills in the critical areas first, before going onto the less critical areas. Hence, the robot uses novelty to learn to coordinate motor behaviour with sensory feedback.

Robot control using visual information was performed by Hermann et al. [14] to examine modular learning using neural networks for biomimetic robots. This chapter describes a modular architecture that is used to control the position/orientation of a robot manipulator by feedback from the visual system.

The outlined modular approach is felt to overcome some of the main limitations associated with neural networks. Using modular learning is a useful approach for robots as there is limited data for training, robots must function in real-time in a real-time environment. A single neural network may not be sufficient to perform a complex task, however by using a modular sequential and bidirectional arranges of neural modules solutions can be found. Motivated by biological modularity Hermann et al. [14] use extended Kohonen maps that combine a self-organising map with ADA-Line networks (SOM-LLM) in the modular model. This neural network approach was selected as the SOM-LLM is simple and offers a topological representation. To test the modular architecture it was used to control a robot arm using two cameras that are positioned on a robot head. This chapter is able to offer an approach that can combine a set of neural modules that can converge and so learn complex systems.

In the area of RoboCup Mayer et al. [19] have developed a neural detection system that achieves colour-based attention control and neural object recognition to determine whether another robot is observed. The ability to recognise teammate robots and opposition robots is fundamental for robot soccer games. Problems associated with self-localisation and communication between robot teammates has lead to the need for an approach that is able to detect teammates and the opposition in a reliable manner based on vision. The approach identified in Mayer et al. [19] in this book uses the following steps (i) identify region of interest; (ii) gain features from this region; (iii) classify the features using neural networks; and (iv) arbitrate the classification outcome to establish if a robot is recognised and whether it is part of the own or opposition team. Regions of interest are typically determined using blob search using a segmented and colour-indexed picture. The types of features that are used in the robot are width, how much black is in the image and an orientation histogram. Once the features are established they are passed into two multilayer preceptron neural networks that are used to classify the features. One network was used to process the simple features and the other for the orientated histogram levels. These two networks produce a probability value stating if a robot is present. A decision as to whether a robot is present is based on whether the joint probablity from these two neural networks is greater than a threshold. The team the robot belongs to depends on the recognition of a colour marker. This approach has proved to give very good performance when classifing the presence of a robot and whether it belongs to the opposition or the own team.

Two approaches for visual homing using a descriptor that characterises local image sections in a scale invariant fashion are considered by Vardy and Oppacher [45] in this book. Visual homing is returning to a location by contrasting the current image with the one at the goal. This approach is based on the behaviour of insects like bees or ants. The two homing approaches rely on edges being extracted from the input images using a Sobel filter. The first approach uses the common technique of corresponding descriptors among images and the second approach establishes a home vector by determining the local image regions which are most similar between the two images, and assuming that these correspond to the foci of expansion and contraction. This second approach makes use of the structure of the motion field for pure translation. The second method found a home vector more directly using the stationary local image region closest from the two images. The first approach was able to out-perform the warping method, while the second performs equivalently to the warping method.

6 Conclusion

As can be seen from this book there is much research being carried out towards biomimetic robots. In particular, the MirrorBot project has contributed to the development of biomimetic robots by taking neuroscience evidence and producing neural models, but also several other joint European projects have worked in the same direction (SpikeFORCE, BIOLOCH, CIRCE, CYBREHAND, MIRROR). The breadth of the international research community sharing the research goal of biomimetic robotics can be seen not only from the contributions to this workshop, but also from many closely related conferences that have been organised in recent years. The chapters included in this book show that the MirrorBot project has successfully developed models that are able to check the syntactic consistency of word sequences, visually explore scenes and integrate multiple inputs to produce sophisticated robotics systems. This shows that we can overcome the present limitations of robotics and improve on some of the progress made by basing robots on biological inspiration such as the mirror neuron concept and modular cerebral cortex organisation of actions. The second part of the book shows the diversity of the research in the field of biomimetic neural robot learning. Although this research produces different approaches to diverse sets of robot function they are all connected by performance, flexibility and reliability that can be achieved by those based on biological systems. The biological systems thereby act as a common guideline for these diverse, cooperative, cooperating and competing approaches. Hence, there is a need to base robotic systems on biological concepts to achieve robust intelligent systems. As shown in this book the current progress in biomimetic robotics is significant, however more time is needed before we see it in full operation showing fully autonomous biomimetic robots.

References

1. M. Arbib. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Science*, pages 1–9, 2004.
2. J. Bach. Representations for a complex world: Combining distributed and localist representations for learning and planning. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
3. K. Chokshi, S. Wermter, P. Panchev, and K. Burn. Image invariant robot navigation based on self organising neural place codes. In S. Wermter, G. Palm, and

M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
4. R. Fay, U. Kaufmann, A. Knoblauch, H. Markert, and G. Palm. Combining visual attention, object recognition and associative information processing in a neurobotic system. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
5. M. Folgheraiter and G. Gini. Maximumone: an anthropomorphic arm with bio-inspired control system. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
6. V. Gallese. The intentional attunement hypothesis. the mirror neuron system and its role in interpersonal relations. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
7. V. Gallese, L. Escola, I. Intskiveli, M. Umilta, M. Rochat, and G. Rizzolatti. Goal-relatedness in area F5 of the macaque monkey during tool use. Technical Report 17, MirrorBot, 2003.
8. V. Gallese, L. Fadiga, L. Fogassi, and G. Rizzolatti. Action recognition in the premotor cortex. *Current Opinion in Neurobiology*, 119:593–609, 1996.
9. V. Gallese and A. Goldman. Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Science*, 2(12):493–501, 1998.
10. V. Hafner and F. Kaplan. Learning to interpret pointing gestures: Experiments with four-legged autonomous robots. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
11. O. Hauk, I. Johnsrude, and F. Pulvermüller. Somatotopic representation of action of action words in human motor and premotor cortex. *Neuron*, 41:301–307, 2004.
12. O. Hauk and F. Pulvermüller. Neurophysiological distinction of action words in the frontal-central cortex. Technical Report 7, MirrorBot, 2003.
13. O. Hauk and F. Pulvermüller. Neurophysiological distinction of action words in the frontal lobe: An ERP study using minimum current estimates. *European Journal of Neuroscience*, 21:1–10, 2004.
14. G. Hermann, P. Wira, and J-P Urban. Modular learning schemes for visual robot control. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
15. H. Jasso and J. Triesch. A virtual reality platform for modeling cognitive development. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
16. A. Knoblauch and F. Pulvermüller. Sequence detector networks and associative learning of grammatical categories. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
17. E. Kohler, C. Keysers, M. Umilta, L. Fogassi, V. Gallese, and G. Rizzolatti. Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, 297:846–848, 2002.
18. H. Markert, A. Knoblauch, and G. Palm. Detecting sequences and understanding language with neural associative memories and cell assemblies. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.

19. G. Mayer, U. Kaufmann, G. Kraetzschmar, and Palm G. Neural robot detection in robocup. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
20. O. Menard, F. Alexandre, and H. Frezza-Buet. Towards word semantics from multi-modal acoustico-motor integration: Application of the bijama model to the setting of action-dependant phonetic representations. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
21. Q. Meng and M. Lee. Novelty and habituation: The driving force in early stage learning for developmental robotics. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
22. J. Murray, H. Erwin, and S. Wermter. A hybrid architecture using cross-correlation and recurrent neural networks for acoustic tracking in robots. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
23. G. Palm. *Neural Assemblies. An Alternative Approach to Artificial Intelligen e.* Springer-Verlag, 1982.
24. C. Panchev. A spiking neural network model of multi-modal language processing of robot instructions. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
25. W. Penfield and T. Rasmussen. *The cerebral cortex of man*. Macmillan, Cambridge, MA, 1950.
26. D. Perani, S. Cappa, T. Schnur, M. Tettamanti, S. Collina, M. Rosa, and F. Fazio. The neural correlates of verbs and noun processing a PET study. *Brain*, 122:2337–2344, 1999.
27. F. Pulvermüller. Words in the brain's language. *Behavioral and Brain Sciences*, 22(2):253–336, 1999.
28. F. Pulvermüller. Brain reflections of words and their meaning. *Trends in Cognitive Neuroscience*, 5(12):517–524, 2001.
29. F. Pulvermüller. A brain perspective on language mechanisms: from discrete neuronal ensembles to serial order. *Progress in Neurobiology*, 67:85–111, 2002.
30. F. Pulvermüller. *The Neuroscience of Language: On Bain Circuits of Words and.* Cambridge Press, Cambridge, UK, 2003.
31. F. Pulvermüller, R. Assadollahi, and T. Elbert. Neuromagnetic evidence for early semantic access in word recognition. *European Journal of Neuroscience*, 13:201–205, 2001.
32. F. Pulvermüller, M. Häre, and F. Hummel. Neurophysiological distinction of verb categories. *Cognitive Neuroscience*, 11(12):2789–2793, 2000a.
33. F. Pulvermüller, B. Mohr, and H. Schleichert. Semantic or lexico-syntactic factors: What determines word class specific activity in the human brain? *Neuroscience Letters*, 275(81-84):2789–2793, 1999.
34. G. Rizzolatti and M. Arbib. Language within our grasp. *Trends in Neuroscience*, 21(5):188–194, 1998.
35. G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi. The mirror system, imitation, and the evolution of language. *Cognitive Brain Research*, 3:131–141, 1996.
36. G. Rizzolatti, L. Fogassi, and V. Gallese. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Review*, 2:661–670, 2001.
37. G. Rizzolatti, L. Fogassi, and V. Gallese. Motor and cognitive functions of the ventral premotor cortex. *Current Opinion in Neurobiology*, 12:149–154, 2002.

38. U. Scarfogliero, M. Folgheraiter, and G. Gini. Larp, biped robotics conceived as human modelling. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
39. D. Sheynikhovich, R. Chavarriaga, T. Strösslin, and W. Gerstner. Spatial representation and navigation in a bio-inspired robot. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
40. O. Shtyrov, Y. Hauk and F. Pulvermüller. Distributed neuronal networks for encoding category-specific semantic information: the mismatch negativity to action words. *European Journal of Neuroscience*, 19:1–10, 2004.
41. M. Spitzer. *The Mind Within the Net: Models of Learning, Thinking and Acting*. MIT Press, Cambridge, MA, 1999.
42. A. Sung, A. Merke, and M. Riedmiller. Reinforcement learning using a grid based function approximator. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
43. A. Treves and E. Rolls. Computational analysis of the role of the hippocampus in memory. *Hippocampus*, 4(3):374–391, 1994.
44. M. Umilta, E. Kohler, V. Gallese, L. Fogassi, L. Fadiga, and G. Keysers, C.and Rizzolatti. I know what you are doing: A neurophysical study. *Neuron*, 31:155–165, 2001.
45. A. Vardy and F. Oppacher. A scale invariant local image descriptor for visual homing. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
46. J. Vitay, N. Rougier, and F. Alexandre. A distributed model of spatial visual attention. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.
47. S. Wermter, J. Austin, D. Willshaw, and M. Elshaw. Towards novel neuroscience-inspired computing. In S. Wermter, J. Austin, and D. Willshaw, editors, *Emergent Neural Computational Architectures based on Neuroscience*, pages 1–19. Springer-Verlag, Heidelberg, Germany, 2001.
48. S. Wermter, C. Weber, V. Gallese, and F. Pulvermüller. Neural grounding robot language in action. In S. Wermter, G. Palm, and M. Elshaw, editors, *Biomimetic Neural Learning for Intelligent Robots*. Springer-Verlag, Heidelberg, Germany, 2005.