Perception-Action Coupling via Imitation and Attention

George Maistros, Yuval Marom and Gillian Hayes
Institute of Perception, Action, and Behaviour,
Division of Informatics, University of Edinburgh,
5 Forrest Hill, Edinburgh, EH1 2QL, UK
{georgem, yuvalm, gmh}@dai.ed.ac.uk

Abstract

When we interact with objects, we give them meaning, *i.e.* we show what they are potentially useful for. We believe that physical entities are anchored to perceptual representations, and *through* them *to* the actions that they 'afford'. This paper brings an imitation mechanism and an attention system together computationally, with the aim of having a system that is capable of creating and maintaining these anchors. The integrated system is implemented on two different platforms: a simulated humanoid robot learning from another how to drink a glass of beer, and a simulated mobile robot learning from another how to follow walls.

Introduction

Gibson (1966) claimed that our perception of the world is expressed in terms of our interactions with it. He introduced the concept of *affordances* of objects that represent the actions that one may apply to those objects. These actions are the ones that the object *affords*. A mug, for example, affords various hand grasps, as a chair affords sitting.

We suggest that perception coupled together with action forms a correspondence between the perceived world and the interactions with it. This correspondence can offer robots the necessary framework with which to ground perceptual information to their own actions, and even use it to 'understand' the observed interactions of others with objects.

Our approach is to anchor sensory data of physical entities *through* perceptual classifications *to* the actions that they 'afford'. This is achieved through a perceptual attention system that matches and classifies input, and a perception-action coupling mechanism that is responsible for the acquisition of motor skills through imitation. In other words, we believe that the integration of these two systems offers appropriate motor skills onto which one can ground the perceptual instances of the world.

There are therefore two main modules that make up the system described in this paper, resulting from work undertaken separately: (1) a *schema network* that handles perception-action coupling, and (2) an *attention system*. The next two sections describe these sub-systems, and also provide their biological and psychological inspirations, respectively.

The paper then proceeds with a section describing how the current work integrates these sub-systems into a single system capable of anchoring affordances to perceptual information. Then, the following sections describe the implementation of the system on two different platforms. Finally, a concluding section provides a discussion of current and future work.

Schema Network

Neurophysiological Model

Neurophysiological experiments in the brain of *macaque* monkeys suggest that neurons in area F5 (the rostral part of inferior area 6), namely *canonical* and *mirror* neurons, have both visual and motor properties (Rizzolatti *et al.* 1988). PET studies illustrate the presence of neurons with similar properties in the human brain as well. Further, it is believed that these neurons may form the fundamental basis for imitation in primates (Rizzolatti, Fogassi, & Gallese 2000).

Single neuron studies by Gallese *et al.* (1996) and Rizzolatti *et al.* (1996) explored further the properties of F5 neurons and exposed a strong relationship between perception and motor control. For instance, mirror neurons fire *both* when the monkey performs an action (motor stimulus) *and* when it observes another monkey or the experimenter perform that same action (perceptual stimulus). Similarly, canonical neurons discharge *both* when the monkey performs an action, say *A*, *and* when the monkey merely observes a 3D object that *affords* action *A*.

The real attractiveness of these neurons reveals itself through their functional role. Most literature agrees on the views summarised by Rizzolatti, Fogassi, & Gallese (2000), where mirror neurons (and F5 neurons in general) perform the necessary visuomotor transformations to offer a 'vocabulary' of potential actions (or 'ideas' on how to (re)act). We also think of these potential actions as affordances.

Implementation

Inspired by these views we have developed an imitation mechanism that employs Arbib's *Schema Theory* (Arbib 1981) in the form of a schema network that expresses this

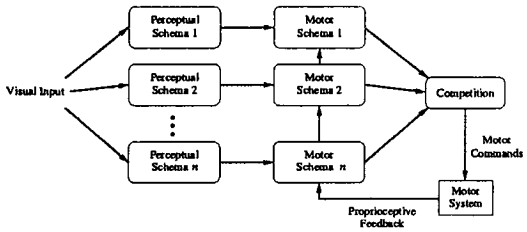

Figure 1: A schema network: several perceptual-motor schema pairs send candidate motor commands to the competition module which send the winner to the motor system for execution.

very role of the mirror neurons. This role together with the notion of affordances has been successfully implemented by Fagg & Arbib (1998), while the mechanisms present in this work are inspired by and resemble the work by Demiris & Hayes (1999) and Matarić *et al.* (1998).

A schema network is formed using pairs of perceptual and motor schemas (see Figure 1). Perceptual schemas represent perceptual structures of an action (*e.g.* the joint-angles of an observed demonstrator), and motor schemas represent the motor skills for that action. The motor skills are expressed in terms of internal targets (*e.g* joint-angle targets) that are to be achieved for that particular skill.

All perceptual modules receive perceptual input and compare them with their stored structures to produce a confidence measure. This confidence, if high enough, is used to activate the motor schemas. Then, active motor schemas convert their internal targets into candidate motor commands using proprioception feedback and a forward model. Given a robot's current state and internal targets a forward model suggests the motor commands that best achieve those targets (here we are using a simplified version of the forward model used by Demiris & Hayes 1999). These candidate commands compete in the competition module and the one with the highest confidence (given from the perceptual schemas) wins and is sent to the motor system for execution.

As such, when the schema network receives perceptual input of a familiar behaviour it triggers in sequence those motor skills that imitate that behaviour. Alternatively, an unfamiliar behaviour results in either no or very little motor skill triggering, thus producing poor imitation. For a more detailed description of the schema network and its implementation see Maistros & Hayes (2001).

Figure 2 illustrates a problem of this implementation. Perceptual and motor schemas are given to the network *a priori* so that it can imitate familiar behaviours (*i.e.* the ones represented by its schemas).

For example, the schema network of Figure 2 has been manually built such that it accommodates three behaviours: behaviours 1 and 2, which are quite similar to one another (drinking a glass of beer), and behaviour 3 which is rather different (moving a glass over a table). Each behaviour is considered as a long sequence of joint-angles. These sequences are arbitrarily divided into a number of parts and stored into perceptual and motor schemas. Behaviours 1 and

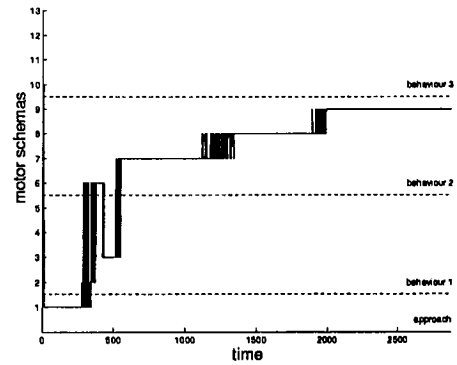

Figure 2: The output of the schema network (see text for explanation).

2 are divided into a common approach part (approaching the glass to be grasped) and another four sequential parts each. Similarly, behaviour 3 is divided into four sequential parts.

Each such part is stored in one perceptual and one motor schema forming a perceptual-motor schema pair. The perceptual schema holds a sequence of joint-angles that are to be recognised, and the motor schema holds the sequence of joint-angle internal targets that are to be achieved. Hence a schema network of a total of 13 pairs is formed where behaviour 1 is perceptually and motorically represented by perceptual and motor schemas 1–2–3–4–5 in sequence, behaviour 2 by 1–6–7–8–9, and behaviour 3 by 10–11–12–13.

In the episode of Figure 2, the perceptual input given is that of behaviour 2, and hence motor schemas 1–6–7–8–9 are expected to be activated in sequence. However, one observes that, at times 300–600, schemas 2 and 6 are alternate winners. As it turns out, schemas 2 and 6 are quite similar to each other (grasping the glass) and hence strongly competitive. There is a redundancy in the representation.

The problem arises from the arbitrary break-down of the behaviour into perceptual and motor schemas, as deemed appropriate by the designer.

The work presented later in this paper is intended to address some of these weaknesses by employing an attention system which replaces perceptual schemas with clusters of perceptual experiences. The agent builds up perceptual structures from experience, rather than receiving them *a priori* from the designer. This reduces the chance of storing duplicate schemas and also ensures that only significant experiences are represented in the network.

Attention System

Psychological Model

In our approach to attention, we are mainly inspired by the *Habituation Hypothesis of Selective Attention* (Cowan 1988), which claims that rather than attention being a simple filtering mechanism that selects certain inputs and disposes of others, it is a system that continually inspects all its input channels, compares them to descriptions stored in

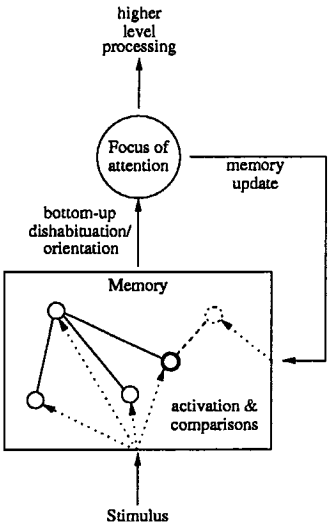

Figure 3: The attention model. Units in memory respond to a stimulus, and this determines whether the stimulus is attended to, due to novelty or forgetting (see text), or ignored. Attending to a stimulus (the orientation response) can result in further, higher-level processing, and modifications in memory.

memory, habituates to familiar stimuli, and has the ability to dishabituate if needed.

The role of habituation is to inhibit what is known as the *Orienting Response*, which is a combination of neural, physiological, and behavioural changes that an organism undergoes when a novel or significant stimulus is detected (Sokolov 1963; Kahneman 1973). What is interesting is how the orienting response is reinstated (dishabituation), and this can occur due to a number of factors (Balkenius 2000). Our implementation involves two of the main factors: the presentation of a novel stimulus, and the passage of time (forgetting).

According to the model, depicted in Figure 3, the activation and comparisons of structures in memory occur outside of attention, in a passive, automatic manner. When the orienting response is reinstated (due to novelty or forgetting), the information goes through to the focus of attention, which can cause the creation of new structures in memory (see Figure 3). Further, the information from the focus of attention can go through to higher-level processing, such as learning.

Computational System

We have implemented the model discussed above using a variation of the Self Organising Feature Map (SOFM), where structures grow from experience as required, rather than being specified *a priori*. The SOFM has been successfully used in the past in modelling robotic sensory input (Nehmzow 1999). It attempts to cover the sensory input space with a network of nodes, and edges connecting neighbouring nodes are determined by some distance measure, in our case a Euclidean distance. This has the effect of preserving the topology inherent in the space.

We have adopted and suited to our purposes an algorithm developed by Marsland, Nehmzow, & Shapiro (2001), which incorporates notions of habituation, novelty detection, and forgetting. It reflects how perceptual information activates stored structures for comparison; how structures are added, updated, and deleted; and how structures habituate to familiar stimuli, but are able to reinstate the orientation response if required.

The main asset of this algorithm is that it keeps a *habituation* measure for each node in the SOFM. This measure gives an indication of familiarity, *i.e.* the frequency of that node's activation, which provides a useful heuristic. Each time a node is active, its habituation value decreases exponentially.

The nodes and edges of the SOFM form the memory part of Figure 3, where the dotted arrows denote the comparisons between the incoming stimulus and the stored structures, and the node in bold font denotes the 'activated', winning node, *i.e.* the one that best matches the input. One of three possible cases follows this activation:

1. If the match is not good enough novelty is detected. The orientation response is initiated, which corresponds to the creation of a new node (denoted by a 'dashed' node in Figure 3).
2. If the match is good the activated node and its neighbours are modified by a constant fraction to match the input even better, and they are also habituated.
3. If the winning node is fully habituated, it does not respond to any more input, until dishabituation is forced through the passage of a fixed time interval. The information is re-attended to until the node fully habituates again.

To summarise, the system handles attention as follows. Nodes in the network respond and habituate to their respective stimuli. When fully habituated, nodes ignore further stimulation and hence do not get updated. The orienting response is reinstated either due to novelty detection, when a new node is created, or due to forgetting, when a node is dishabituated — in both situations the stimulus is (re-) attended to.

For more information on the general algorithm see Marsland, Nehmzow, & Shapiro (2001), and for our implementation see Marom & Hayes (2001).

The Integrated System

The two systems described in the previous sections, namely the schema network and the attention system, were developed separately in on-going work. The work presented in this paper brings the two sub-systems together computationally, and the aim is to have a system that is capable of creating and maintaining representations of perceptual experiences with respect to the actions that they afford.

In other words, the system grounds perceptual instances to the actions that they afford. The part of the system responsible for grounding (through attention) is the SOFM, and the part responsible for the acquisition of actions, or motor skills (through imitation) is the schema network. A schematic diagram of the system is shown in Figure 4.

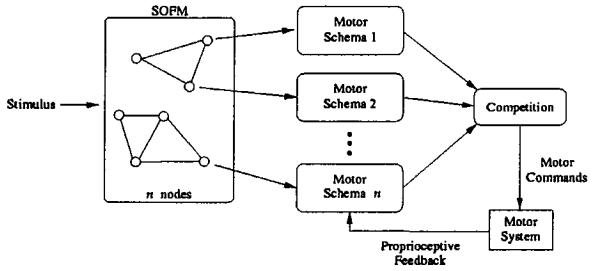

Figure 4: The integrated system: the SOFM handles the stimulus and finds the best matching node at each time; its hard-wired motor schema calculates the motor commands before sending them to the motor system for execution.

The attention system (SOFM) replaces the perceptual schemas of Figure 1. Thus, instead of arbitrarily categorising the input space, we now use attention for this purpose. The SOFM takes input directly from the sensors and processes them as described in the previous section. For each node in the SOFM there is a single motor schema hard-wired to it,[1] where internal targets are to be stored.

At the learning phase both the SOFM and the motor schemas are built in response to the input. Since during training the nodes in the SOFM move in the input space, we create and modify motor schemas only when a node is fully habituated, and therefore settled. These modifications depend on the habituation value of the winning node:

1. If the winning node is not fully habituated, the corresponding motor schema is left intact.
2. If the winning node is fully habituated and the perceptual input is part of the same learning episode, the current target is appended to the sequence of targets (if any) of the corresponding motor schema.
3. If the winning node is fully habituated and the perceptual input is part of a new learning episode, all targets of the corresponding motor schema are deleted and overwritten by the current target.

Due to forgetting (dishabituation), mentioned in the "Attention System" section, a node can re-train itself and its motor schema, to adapt to changes in the environment, or to unsuccessful previous trainings. We have not yet developed sophisticated methods for modifying motor schemas (for re-training); currently we are using the simple and crude method described above.

To summarise, an episode is recorded in a motor schema that corresponds to a fully habituated SOFM node, and this episode is overwritten by future episodes.

The emergent level of granularity in the motor schemas is governed by the sensitivity of the SOFM nodes, which in turn is controlled by the threshold used for detecting novelty. This is the classical over-generalisation (too few nodes) versus over-fitting (too many nodes) issue.

[1]The possibility of hard-wiring more motor schemas to a perceptual categorisation is being considered.

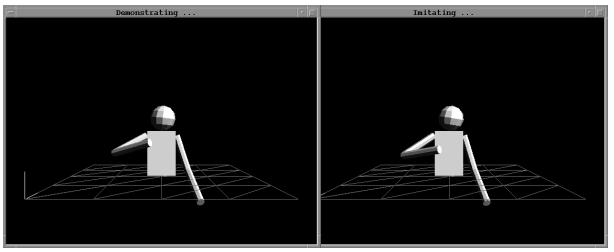

Figure 5: The simulation platform: the imitator (right) is imitating the demonstrator (left).

At the recall phase we use the SOFM and motor schemas created in the learning phase to recall the appropriate motor skills required at particular perceptual states. At each step the input activates the SOFM node that best matches it, and the motor schema associated with that node is then responsible for producing motor commands. We are interested in testing how well the system can reproduce the learned behaviour, and more specifically in the dynamics of perceptual-motor activations.

We have tested the integrated system on two different platforms, both in simulation, and they are presented in the following two sections. We are also currently testing the system on a physical robot (a Real World Interface B21 robot).

Experiment 1: Postural Imitation

Setup

The experiments presented in this section simulate the dynamics of two eleven degrees of freedom simulated robots (waist upwards, see Figure 5): a demonstrator and an imitator. Each robot has three degrees of freedom at the neck, three at each shoulder, and one at each elbow.

The imitator has explicit access to the joint-angles and instantaneous velocities of the demonstrator along with object information (visual perception), and to its own angles and velocities (proprioception), after noise is added onto them.

At the learning phase the imitator observes the demonstrator perform an action, and tries to learn it. At the recall phase, the imitator observes the demonstrator perform that action again, and tries to reproduce it, *i.e.* imitate it. The action involves 'picking up' a glass off a table, 'drinking' its contents, and 'putting' it back on the table.

Note that, due to software limitations, the robots do *not* interact with objects, although objects *are* represented, recognised, and virtually — yet effortlessly — moved (hence the visual absence of the glass in Figure 5).

Imitation, here, is used in the recall phase to explicitly expose the imitator to perceptual cues that may trigger appropriate motor schemas to reproduce the learned action. In this experiment, the motor schemas store joint-angles as internal targets which are converted into motor commands via a Proportional-Integral-Derivative (PID) controller (a forward model). Notice that these motor skills correspond to affordances onto which perceptual information is to be grounded.

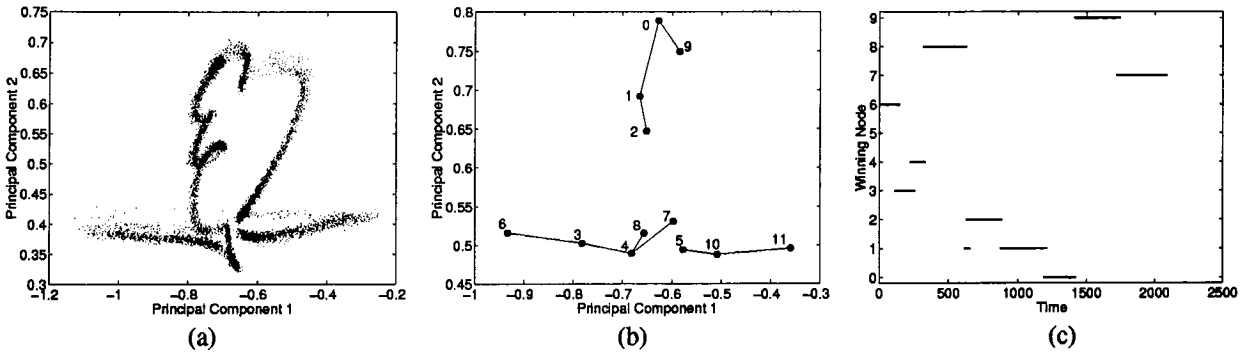

Figure 6: Experiment 1: (a) the perceptual data exposed to at the learning phase, projected onto the first 2 principal components; (b) the SOFM produced by the attention system for these data; (c) the SOFM node activation at the recall phase.

Learning Phase

Figure 6(a) shows the perceptual data that the imitator observes at the learning phase, while Figure 6(b) shows the resulting SOFM network. Since the dimensionality of the input space is quite high (28), we have used a dimensionality reduction technique called Principal Component Analysis (PCA), for display and analysis purposes. PCA finds the most statistically significant dimensions, called Principal Components, in a multivariate dataset (see Afifi & Clark (1996) for more information on PCA).

Figure 6(a) is the projection of the perceptual data onto the first 2 principal components found by PCA.[2] Similarly, Figure 6(b) is the projection of the SOFM network onto the first 2 principal components. Interpreting the PCA plots is not straightforward because it is hard to inspect the principal components and trace them back to the original dimensions, especially when the original dimensionality is quite high. However, from our knowledge of what the learner *should* be perceiving, we can conjecture what the different parts of the plots mean.

It seems that the first Principal Component expresses a generalised velocity ('velocity') of the right hand joint-angles, while the second expresses a generalised height ('height') of the right hand end point or wrist. Notice that the first PC is 'zero' around -0.7, while the second is at 'table height' around 0.4.

We can distinguish three shapes in Figure 6(a): one fairly straight curve at the bottom left of the figure (going left to centre), a central loop in the middle (going clockwise), and another fairly straight curve at the bottom right (going centre to right). In fact, this *does* correspond to the three major parts of the "drinking a glass of beer" action: (i) approach the glass; (ii) pick up, drink, and put down; (iii) move away from the glass.

[2]It's important to check how much of the variance is actually accounted for by the first 2 principal components, and this is given by the sum of their eigenvalues. This gives us an indication of how representative a 2-dimensional plot is of the true complexity in the data. In this case, 2 dimensions account for approximately 78% of the total variance.

Further, in part (i), as the hand approaches the glass, the 'velocity' tends to 'zero' (it decelerates), whereas in part (iii), as the hand moves away from the glass, the 'velocity' increases from 'zero' (it accelerates).

Note however that a single episode ends when the demonstrator puts the glass back on the table, *i.e.* parts (i) and (ii) have been completed. Repeating the demonstration episode first takes the hand off and away from the glass, back towards the starting location (*i.e.* part (iii)) and then through parts (i) and (ii). Nevertheless, part (iii) is clustered and associated with appropriate motor skills.

Indeed, in Figure 6(b) we observe the attention system clusters these three parts with enough nodes, with the only exception of the right hand part of the central loop. This is due to the small number of data points in that area. Further work will consider the issue of whether it is desirable to sample at a faster rate when fast movements (as the one mentioned above) are to be imitated.

Recall Phase

In the recall phase, while the demonstrator performs the action, the SOFM receives continuous perceptual input, finds the node that best matches the input, and activates the corresponding motor schema. That schema recalls the current internal target state and calculates the motor commands to achieve that state using its PID controller.

Figure 6(c) shows the SOFM activation of a single episode at the recall phase, *i.e.* the sequence of nodes that are activated in response to the input. We observe that the nodes in the clusters, created at the learning phase (Figure 6(b)), are activated in sequence at the recall phase. The sequence 6–3–4–8 controls the imitator from the starting posture, to 'grasp' the glass. Similarly, motor schemas 2–1–0 control the imitator to 'lift' the glass and 'move' it towards the mouth, and schemas 0–9–7 from the mouth back towards the table. As mentioned above, a single episode ends when the glass is put back on the table; which explains the absence of winning nodes 5–10–11.

Figure 6(c) merely demonstrates the correct activation dynamics of the SOFM nodes (the winning node at each time step). However, we expect that motor schemas hard-wired

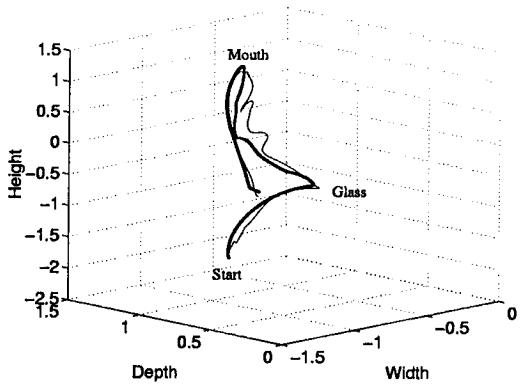

Figure 7: The trajectories of the right hand wrists of the demonstrator (bold font) and the imitator (normal font) in a single episode in the recall phase.

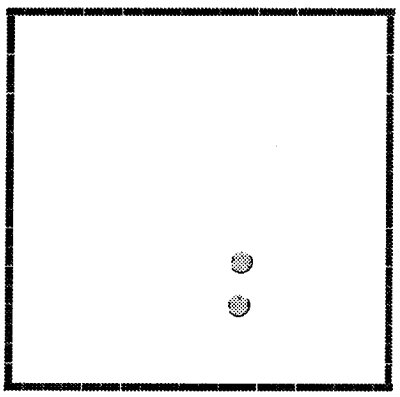

Figure 8: The Khepera mobile robot simulator.

to those nodes will generate appropriate motor commands to efficiently imitate the demonstrated behaviour.

On the other hand, since there is no easy or standardised way to evaluate imitation, 'efficient imitation' is hard to define (Matarić 2000). For the purposes of this task, though, we consider the position of the hand(s) over time relative to the position of the demonstrator's hand(s) sufficient as an indication of 'efficient imitation'.

Figure 7 shows the trajectories of the right-hand wrists of both the demonstrator (in bold font) and the imitator (normal font) in a single episode in the recall phase. Notice that the trajectories are close to each other, save for one part of the action. This reflects the actual behaviour of the imitator where all sub-goals of the task are achieved with a reasonable degree of accuracy. One possible way to numerically evaluate this degree of accuracy is to calculate the area between the curves; this is part of ongoing work.

Experiment 2: Imitation by Following

Setup

The experiments presented in this section were performed using a Khepera mobile robot simulator with a learner agent following behind a teacher agent (see Figure 8). The learner robot perceives information through 6 infra-red sensors around its front. It uses its built-in following behaviour to keep behind the demonstrator, as the demonstrator executes the task, which in this case is wall-following.

We regard this as 'imitation' in the sense that information is implicitly shared between the demonstrator and learner about a specific task to be learned. The object-interactions in this case correspond to how the robot responds to being near a wall: the affordances of the *wall* and of the *no-wall*.

In this experiment the forward model is not as straight forward as in Experiment 1 where the PID provides an intuitive forward model. As mentioned earlier, the motor schemas provide motor skills expressed as a sequence of internal *targets*. The forward model should suggest to the robot how to attain a particular *target* state, given the *current* state, and the possible actions available to the robot.

We let the robot build up this forward model from experience before we start the experiment. The robot wanders around the environment on its own, with an added obstacle-avoidance skill, and collects information about state-transitions. The transition matrix is then stored and used later as a forward model.

Learning phase

Figure 9(a) shows the perceptual data that the learner is exposed to as it follows the demonstrator around the environment, and Figure 9(b) shows the corresponding network produced by the attention system. As before, since the dimensionality of the sensor space is too high to visualise, we have used PCA to reduce the number of dimensions to two (from six).[3]

For a simple task such as wall-following, we do not expect many perceptual clusters, and in fact we see 3 main clusters: one cluster is the intersection of the 2 apparent lines in Figure 9(a), which is the area of low (weak) wall-detection, *i.e.* 'no wall', and as we move away from this intersection in either direction, we see data corresponding to 'right wall' and 'left wall', at different distances from the wall, clustered in distinct areas. The attention system captures these clusters by placing nodes in appropriate places, as shown in Figure 9(b). We therefore expect these clusters to afford the appropriate motor skills.

Recall phase

After the robot has clustered its perceptual space and created corresponding motor schemas at the learning phase, it is placed in the same environment on its own for the recall phase. The default behaviour is 'wandering', and at each step it tries to activate one of the nodes in its SOFM. The winning node recalls the targets stored in the corresponding

[3]The 2 dimensions used in Figure 9(a) account for approximately 98% of the total variance.

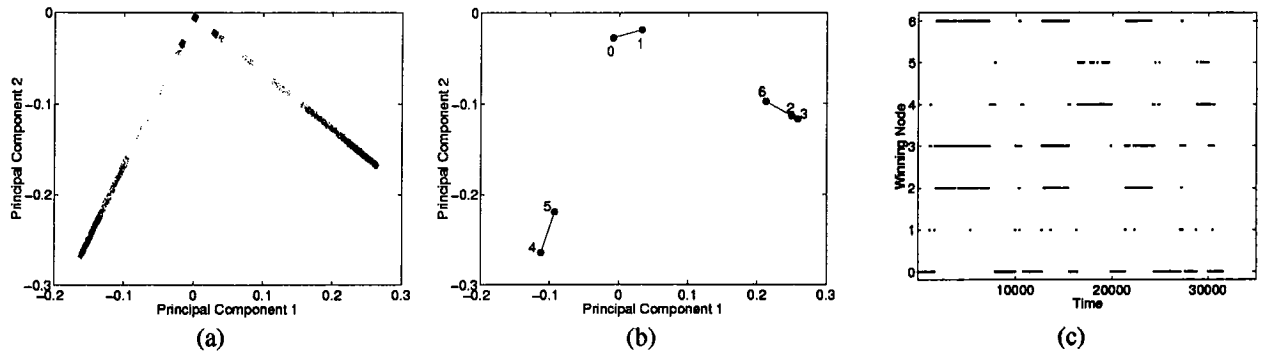

Figure 9: Experiment 2: (a) the perceptual data exposed to at the learning phase, projected onto the first 2 principal components; (b) the SOFM produced by the attention system for these data; (c) the SOFM node activation at the recall phase.

motor schema, and these are passed to the forward model which selects the best action likely to achieve them.

Figure 9(c) shows the SOFM activation at the recall phase. Firstly, we see that the nodes, created at the learning phase, that form clusters (Figure 9(b)), are also activated together and intermittently at the recall phase. Secondly, we see an emergent sequence of activations as the robot moves around trying to recall the information from the motor schemas: the activations of nodes for wall perception on either side are separated by activations of nodes for no-wall perception. This reflects the actual behaviour of the learner: following a right wall, then no wall, then following a right wall again, etc.

Note that the activation dynamics are different than in Experiment 1 (Figure 6(c)), which is inherent in what is being modelled. In Experiment 1 the behaviour being modelled by the motor schemas has a true sense of sequence in it (moving the hand towards a glass, 'picking' it up, etc.); each motor schema stores a part of that sequence.

In this experiment, each motor schema, or rather cluster of schemas, corresponds to being in a particular perceptual 'state', and the motor skills are responsible for *maintaining* that state (for example fine-tuning to stay next to a wall on the left). For this reason we do not expect to have a clear winning node since the robot will keep adjusting itself by activating alternate nodes within the same cluster (where each one affords different fine-tuning).

Visual inspection of the behaviour confirms that the activation of the schemas does correspond to a behaviour closely resembling the one exhibited by the demonstrator. We are also developing a method to numerically evaluate the recalled behaviour in terms of the actual task; we do this by calculating the 'energy' that the robot acquires from the wall at particular configurations from it. At the end of the recall episode we can look at the accumulated energy as a measure of how well the robot performs the wall-following task.

Discussion & Future Work

We have presented on-going work on the integration of an attention system and an imitation mechanism. The aim of this integration is threefold: first, to classify our perceptual experiences such that we distinguish familiar from novel experiences; second, to be able to recall the right motor skills at the right time to achieve feasible (natural) motor control for robots; and third, to ground our perceptual experiences to these motor skills.

To summarise, the overall system needs to be shown the perceptual input together with the action that it affords such that it creates appropriate perceptual and motor structures. Further, a forward model needs to be available such that the internal targets in the motor schemas can be converted into actions.

Currently, perceptual and motor structures are implicitly connected by a simple one-to-one relationship. We are, however, interested in ways to afford more than one action per perceptual classification, as we are also considering more sophisticated ways to update and maintain the motor schemas.

This overall system has been implemented on two different platforms where we observe that the attention component correctly classifies the input space and creates perceptual nodes and motor schemas, and that the imitation component takes full advantage of these affordances (to generate behaviour).

As mentioned earlier, evaluating 'efficient imitation' is as difficult a task as is challenging and it is part of on-going work in the field. Evaluating this system is therefore equally difficult. However, as mentioned throughout the paper, we are currently working on ways to measure numerically the overall performance of the system on a task-specific basis.

When a conspecific performs an action that we are familiar with, not only are we able to imitate that action, but we are also able to 'understand' it, *i.e.* we 'know' the perceptual context that triggered that action together with its immediate consequences. This can be regarded as 'understanding' in the sense that we can internally imitate an action, and a higher-level system that overlooks its consequences may have the capacity to infer its goal(s) and/or purpose. Our work is an attempt to encompass these ideas in a computational framework.

References

Afifi, A. A., and Clark, V. A. 1996. *Computer-Aided Multivariate Analysis*. London: Chapman & Hall. ISBN 041273060x.

Arbib, M. A. 1981. Perceptual structures and distributed motor control. In Brooks, V. B., ed., *Handbook of Physiology - The Nervous System II. Motor Control*. Amer. Physiol. Soc. 1449–1480.

Balkenius, C. 2000. Attention, habituation and conditioning: Toward a computational model. *Cognitive Science Quarterly* 1(2).

Cowan, N. 1988. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin* 104:163–191.

Demiris, J., and Hayes, G. M. 1999. Active and passive routes to imitation. In *Proceedings of the AISB symposium on imitation in animals and artifacts*. Edinburgh, Scotland, UK.

Fagg, A. H., and Arbib, M. A. 1998. Modeling parietal-premotor interactions in primate control of grasping. *Neural Networks* 11(7/8):1277–1303.

Gallese, V.; Fadiga, L.; Fogassi, L.; and Rizzolatti, G. 1996. Action recognition in the premotor cortex. *Brain* 119:593–609.

Gibson, J. J. 1966. *The senses considered as perceptual systems*. Houghton Mifflin Company.

Kahneman, D. 1973. *Attention and Effort*. Englewood-Cliffs, New Jersey: Prentice Hall.

Maistros, G., and Hayes, G. M. 2001. An imitation mechanism for goal-directed actions. In *Proceedings of TIMR 2001 – Towards Intelligent Mobile Robots*. Manchester, UK.

Marom, Y., and Hayes, G. M. 2001. Attention and social situatedness for skill acquisition. In *Proceedings of the First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*.

Marsland, S.; Nehmzow, U.; and Shapiro, J. 2001. Novelty detection in large environments. In Nehmzow, U., and Melhuish, C., eds., *Proceedings of Towards Intelligent Mobile Robots (TIMR) 2001*.

Matarić, M. J.; Williamson, M.; Demiris, J.; and Mohan, A. 1998. Behavior-based primitives for articulated control. In Pfeifer, R.; Blumberg, B.; Meyer, J.-A.; and Wilson, S. W., eds., *Proceedings: From Animals to Animats 5, Fifth International Conference on Simulation of Adaptive Behavior (SAB-98)*, 165–170. Zürich, Switzerland.

Matarić, M. J. 2000. Getting humanoids to move and imitate. *IEEE Intelligent Systems* 14:18–24.

Nehmzow, U. 1999. *Mobile Robotics: A Practical Introduction*. Springer Verlag. ISBN 1-85233-173-9.

Rizzolatti, G.; Carmada, R.; Fogassi, L.; Gentilucci, M.; Luppino, G.; and Matelli, M. 1988. Functional organization of inferior area 6 in macaque monkey. 2. Area F5 and the control of distal movements. *Experimental Brain Research* 71(3):491–507.

Rizzolatti, G.; Fadiga, L.; Gallese, V.; and Fogassi, L. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3(2):131–141.

Rizzolatti, G.; Fogassi, L.; and Gallese, V. 2000. Cortical mechanisms subserving object grasping and action recognition: A new view on the cortical motor functions. In Gazzaniga, M., ed., *The New Cognitive Neurosciences*. Cambridge, MA: MIT Press. 539–552.

Sokolov, E. 1963. A neuronal model of the stimulus and the orienting reflex. In Sokolov, E., ed., *Perception and the Conditioned Reflex*. Pergamon Press. 282–294.