Mirror neurons and feedback learning

Steve Womble and Stefan Wermter
Centre of Informatics, University of Sunderland, U.K.

1. Introduction

The neuroscience evidence reviewed in (Rizzolatti & Arbib 1998) suggests that mirror neurons are involved in the comparation of '*goal directed actions*' and the perception of them during competent performance by others. Goal directed actions invariably involve the processing of sequences of more primitive actions. The complex manual tasks such as those discussed in (Rizzolatti & Arbib 1998) share some similarities with simple syntax acquisition. In either case the task is to produce or recognise a useful sequence out of primitive elements. Our model of the mirror system is a synergy of this.

Classical interpretations of language acquisition typically lead to connectionist models of syntax acquisition as the passive acquisition of implicit knowledge concerning a syntax (Reber 1989; Cleeremans 1993). It is also explicit in the axioms on which Gold based his formal learning theorem (Gold 1967). It seems highly unlikely to us that psychologically speaking such an interpretation is correct. After all, what is the point in acquiring knowledge if it is not to *use* it? The model mirror system described in the next section makes *active* use of knowledge already acquired during further learning. In the next section we also discuss a suitable test for the system. We then detail our results which are subsequently discussed. Finally we present our conclusions.

2. Overview of our model system

We propose that as knowledge begins to be acquired through passive adaptation to predominantly correct data, this knowledge is actively used by the learner. In our model this utilisation occurs in two ways. It occurs through the attempted production of syntactically correct sentences. Feedback can then be provided, in the form of recognition of correct constructions, and this additional knowledge can be integrated into the production process. Secondly, we can measure how well the system estimates it has already stored the information contained in the example presented. It can then modify the degree to which it adapts itself to optimise for novel information. Note that this measure does *not* require feedback itself to be calculated. It is simply an estimate generated by the learner on its own use of acquired knowledge. It does not measure if that knowledge is naïve or incorrect only the degree to which it is used. However as the learner's acquired knowledge improves this utilisation representation maps into a confidence measure. This can be *quantified* by feedback.

A high level representation of our model is given in Figure 1. The parts in the square boxes represent the '*mirror neuron system*' that we have developed our abstract model of. In our model the learner examines a newly generated representation and deems the produced sequence to be worthy of production only if the utilisation measure is sufficiently high for all parts of the sequence *and* the learner knows that its knowledge is good. This requires the calculation of a threshold for a filter. This is calculated in a computationally efficient non-neural manner. It represents a minimum firing rate necessary for all neurons to be firing at in order to drive a mirror neuron.

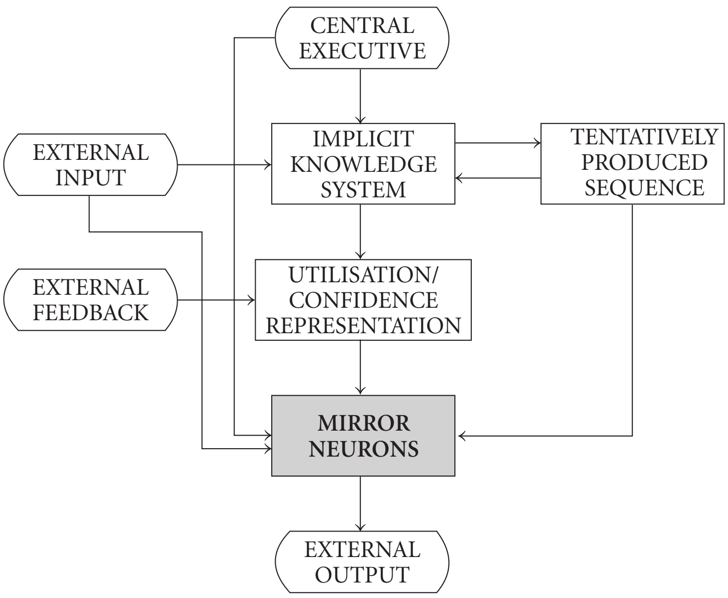

Figure 1. High level representation of model mirror system.

Experimental neuroscience results indicate that mirror cells are highly selective, only firing when their specific associated goal directed action occurs (Rizzolatti & Arbib 1998). In order to facilitate this within our model we require that in addition to input from the utilisation system the mirror field receives high level input from both the sequence production mechanism and from the central executive[1] or external input fields. The former selectively stimulates individual mirror neurons at a sub critical level allowing for full activity to occur when additional stimulation is received from the utilisation system. The latter is necessary to form associations between high level goals and sets of individual sequences.

To test our model we used the formal deterministic stochastic finite state grammar (DSFSG) displayed in Figure 2. It was developed by Reber (1989) in 1965. It was designed to be just complicated enough to take a little over an afternoon of exposure to learn by competent humans. It has been used in a series of psychology experiments by Reber and his colleagues over a number of years and was used in a sequence prediction task for a connectionist neural network by Cleeremans (1993). This formal language task sits nicely on the bridge between action sequence production such as has been reported on in the papers of Rizzolatti and others (Rizzolatti & Arbib 1998; Gallese, Fadiga, Fogassi, & Rizzolatti 1996), by macaque monkeys, and language processing classically associated with Broca's area in humans.

Syntactically correct 'Reber strings' are generated by walking through the finite state system shown in Figure 2. The grammar possesses six nodes. For a given subsequence generated from Figure 2 it requires both the current and the preceding term in the sequence to accurately identify the current node during a transversal.

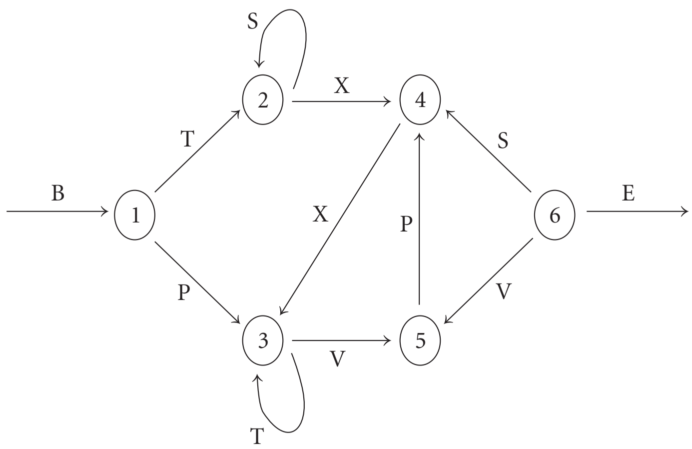

Figure 2. The DSFSG devised by Reber.

The learning task that we set our system was to acquire enough information from the environment concerning the grammar to be able to produce strings that conform to the grammar in a manner that is indistinguishable from the production of strings directly from the DSFSG. This was to be achieved by the coupled use of positive examples and feedback concerning the correctness of tentatively produced strings by the learning agent. Externally supplied strings can also be analysed by the system. Hence in our model there is a link between the evaluation of external input and of internally generated sequences.

We used a simple recurrent neural network (SRN) for the sequence prediction task (Elman 1991). Such networks are known to be well suited to learning DSFSGs since, if they are not over complicated, the SRN's associated error surface is efficiently minimised by learning to represent the nodes of the DSFSG as the internal states of the network. We used three hidden nodes, since three bits span eight binary states and there are six nodes in the Reber grammar. We used standard logistic sigmoid functions as the activation function for the hidden nodes and normalised exponential functions for the outputs. This choice of output generates a controllable 'n-of-c' probability distribution estimate with $n \mapsto 1$ as the temperature parameter of the normalised exponentials $T \mapsto 0$. We used a negative log likelihood function with subtracted entropy as the error function to be minimised. We used backpropagation for training. Training sequences were generated stochastically directly from the DSFSG. We summed errors over a sequence before updating the network and used an explicit momentum term, both to help smooth convergence towards a local optimum.

Elsewhere we have investigated the effects of training this system using a competent teacher which can vary its inputs to improve the acquisition of the network in response to behaviourally realistic output from the learning agent (Womble 2000), where further details concerning the system can also be found. In this work we concentrate on the effects of applying feedback to the system.

To summarise, we wish to compare performance of the basic system conventionally trained using randomly generated sets of strings to both the case where the system analyses its own generated strings to reject those that are likely to be wrong and using self reflexive learning to selectively enhance information stored in the network.

In order to fully analyse these paradigms we investigated the following criteria for successful performance:

– We use a 1-norm measure on the predictive error against the *correct probability distribution* of the next character within a sequence. This measure is summed over a test set of strings and normalised, both with respect to the different lengths of different sequences and over the length of the test set. We generated a test set directly from the DSFSG, selecting the first m most probable strings to be generated.

– There are three related measures concerning the utilisation and effects of the self reflexive analysis of the learning agent's own production performance. These measures are calculated over a set of strings. This set of strings can be generated by the learning agent in which case they correspond to internal reflection on performance, or the strings could be provided externally. The measures are
  – the largest utilisation measure for an incorrectly selected term in a sequence (denoted $\mathsf{max}\{\mathit{fail}\}$);
  – the smallest utilisation measure for a correctly selected term in a sequence (denoted $\mathsf{min}\{\mathit{pass}\}$);
  – and the relative values of each; in particular we are interested in states for which

$$\min_{S\in T}\{\mathsf{min}\{\mathit{pass}\}\} > \max_{S\in T}\{\mathsf{max}\{\mathit{fail}\}\}, \tag{1}$$

  where S denotes a sequence in a test set T.

– There are three 'behavioural' measures. These examine only what would be externally available to interacting agents (including human observers) and corresponds to the '*t1*' 'toy' Turing test as defined by Harnad (2001). Ultimately these measures are the most important concerning the apparent performance of the system. These three measures are
  – the raw (non-filtered) string production success rate;
  – the filtered string production success rates;
  – the '*t1*' test itself which is a measure on the difference in the distribution of sets of strings generated by the learning agent and those generated by a competent agent. We weight the contribution of each sequence with respect to its asymptotic frequency in the set of strings generated directly from the DSFSG, so that more frequently occurring strings contribute proportionally more to the measure

$$d_{bias} = \sum_{i=1}^{M} f_{RG}^{\infty}(S_i).\left|f_{RG}^{\infty}(S_i) - f_{MS}^{N}(S_i)\right|, \tag{2}$$

  where f denotes a frequency, RG refers to the Reber Grammar, MS refers to the mirror system, S_i is the i-th *different* sequence in a test set of N sequences for which there are M *different* sequences, and $|\ldots|$ denotes the 1-norm distance measure.

3. Results

The 1-norm error was carefully calculated using the 755 most frequent strings generated by the DSFSG, with each string weighted according to the asymptotic frequency distribution. For this value a little over 97% of the asymptotic frequency distribution of the infinite set of Reber strings is spanned. This is a deep search, with the probability of the least frequent of these strings occuring naturally being only $2^{-15} \approx 1$ in 32000. We found it quite easy to reduce the 1-norm error per element in a sequence, per string to below 0.1, and with a little more work to around 0.05. However to get below this value a significant amount of searching is required. Our best results using standard backpropagation generated a 1-norm error of 0.0173. To get this we had to perform many searches, and used adaptation of the learning rate η to facilitate convergence to the best local minimum we could find. The results for the system in this state along with results for our lowest 1-norm error system trained using the full feedback system, are given in the following table:

Training Paradigm	**1-norm Error**	d_{bias} **Mean**	d_{bias} **SD**	**Filtered Success**
BP & Filter	0.0173	22.87e-06	1.65e-06	100.0%
MN System	0.0085	12.32e-06	1.04e-06	100.0%
DSFSG	0.0000	4.904e-06	1.35e-06	100.0%
	Incorrect Rejections	Unfiltered Success	Min Pass	Max Fail
BP & Filter	0.00%	92.05%	0.4905	0.0795
MN System	0.00%	91.63%	0.8646	0.1361
DSFSG	0.00%	100.0%	1.0000	0.0000

The results quoted are based on 5000 self generated test strings for all measures based on system production, and were repeated 10 times for the calculation of d_{bias} means and standard deviations.

Figure 3 show detailed results comparing difference measures for these systems to that of the DSFSG itself. For the best backpropagation trained system we found that the difference measure remains fairly indistinguishable to external analysis up to a test set size of about 100 generated strings, the mean for the learning system lies at a single standard deviation from the DSFSG at about the 150 string size, and the learning system becomes clearly distinguishable (the $\pm$ single standard deviation bands no longer overlap) at about 300 string test sets. While for the lowest 1-norm error feedback trained system, these test set sizes are about 200, 700 and 1000 respectively, and at the 5000 string test set size the full active learning mirror system

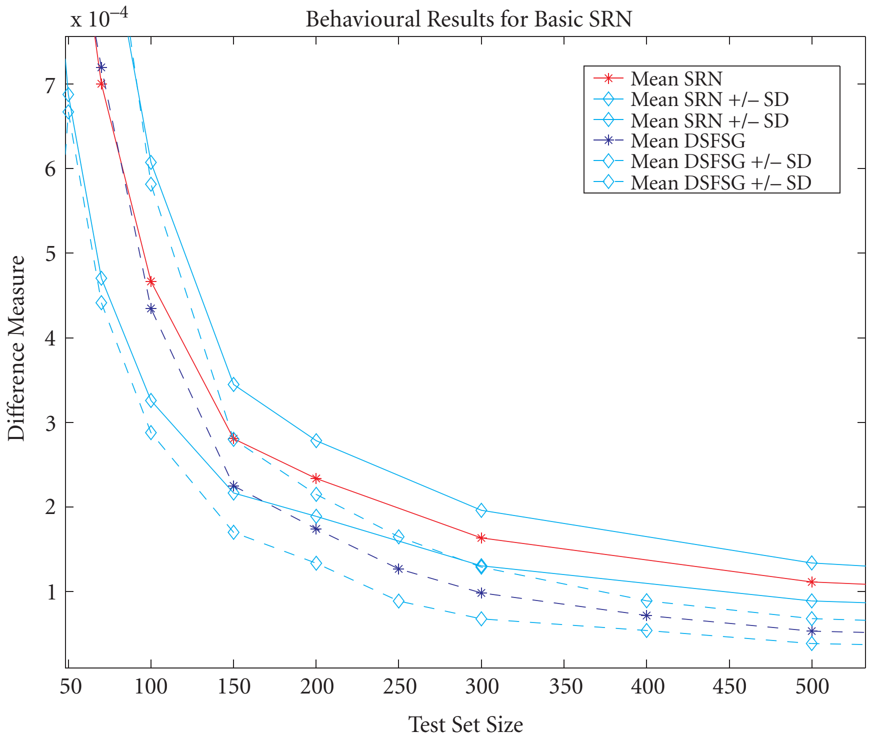

Figure 3. Plot of difference measure for minimum 1-norm error found using the basic neural network. The difference measure for the DSFSG is given for reference. The mean and standard deviations (SD) were calculated from 20 trials for each size of test set.

has a difference measure of only about 54% of the best results using just filtering. Finally we note that without the filtering mechanism the system generates illegal strings about 8% of the time, and is thus clearly distinguishable from the DSFSG.

4. Discussion

The results from the standard backpropagation training show that it is possible to train the SRN to perform the prediction task quite well. The minimum pass results show that our net trained in this way comfortably satisfies the test criteria discussed by Cleeremans (1993).

A useful way to decompose the contributions to the 1-norm error generated by the SRN is to split it into errors in the probability distribution for the next potentially correct characters, and the error caused by non-zero components of

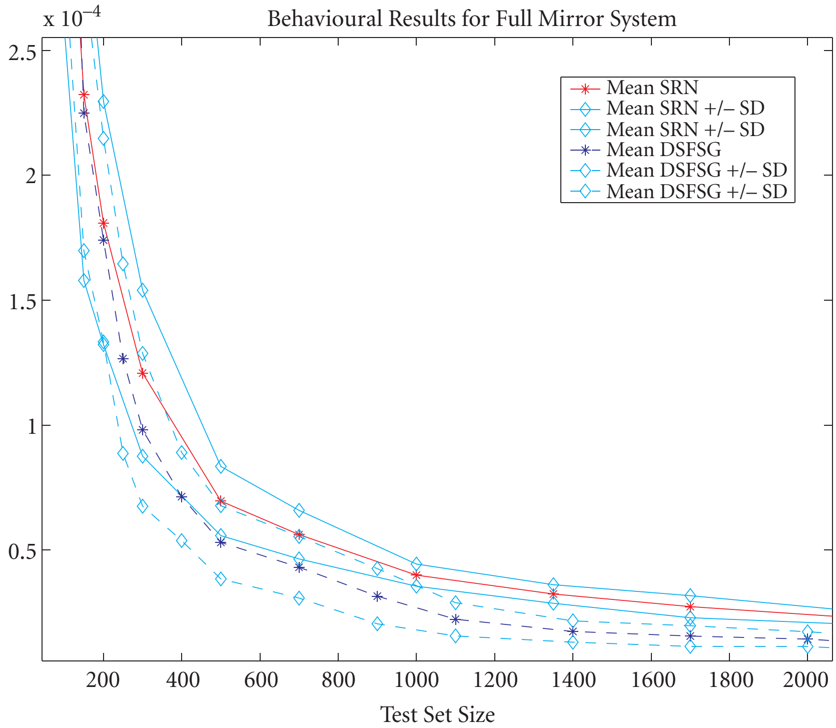

Figure 4. Plot of difference measure for minimum 1-norm error found using the abstracted mirror system. The difference measure for the DSFSG is given for reference. The mean and standard deviations (SD) were calculated from 20 trials for each size of test set.

the distribution associated with incorrect next selections. This splits the error into components associated with metrical and topological error respectively. Under this view a *simple* filter mechanism will produce complete success when the probability for any correct next character is greater than the worst topological error. An approximate condition for this is given by (1).

Our results show that for standard backpropagation it is hard to find a solution for which criterium (1) holds. However for our best network it did. An examination of the utilisation measure shows that topological error has been minimised very well for this network, but that there still exists a behaviourally significant metrical error, something that the filter mechanism set up utilising basic feedback cannot improve upon. However our results clearly show that using full feedback learning as described by our model mirror system the metrical information can be significantly improved. This is due to the modification in the batch learning technique we have introduced in the feedback mechanism. Using the utilisation measure, contributions to the direction in weight space through which the network adapts is biased in favour of contributions for which the utilisation was low i.e. for which the network has a relatively poor representation. The effect of this is to improve the performance of the SRN *selectively* around the regions where the metrical information is poor. However this is at the cost of some interference. The reason for the significant improvement in the behavioural error measures lies in the fact that providing the condition (1) holds then any topological error can be filtered out from the perceived external output of the learning agent.

It is noteworthy that a system which learns how to process sequences in this manner will, when producing strings, suffer from errors reminiscent of characteristic errors in classical Broca's aphasics, when the filtering system is disabled. This provides further circumstantial support for the argument that while biological mechanisms may be significantly different at a low level, at least at a modular level our connectionist network contains some of the characteristics of the biological system from which it was inspired.

The results reported here provide empirical evidence supporting the claim that the apparent problems of learning from only positive examples (Gold 1967) can be neatly circumvented using feedback learning, and that this approach is a plausible mechanism for L1 acquisition in humans. While our system is not a detailed model of a biological mirror system at a neural level, we do claim that the *high level* mechanisms used in our model for learning (positive examples, feedback, and optionally intelligent teaching by a competent teaching agent) are plausible mechanisms during infant language learning, and our results show that they have the potential to be successful. As a minimum our results indicate that at least for context free grammars up to the complexity of Reber grammars, positive examples and feedback are *sufficient* for the acquisition process to succeed.

5. Conclusion

The results clearly show that contrary to the claims made by Cleeremans (1993) the use of positive only data for the implicit acquisition of the DSFSG of Reber, when applied to an SRN hand crafted to use the ideal network topology for the acquisition of the grammar, is a difficult task. Given that it is usually thought that the gradient decent backpropagation techniques used during the adaptation of the artificial system are more powerful than those available to the biological system, and that the Reber Grammar is obviously significantly simpler than any natural language we argue that the positive only learning mechanism is likely to be insufficient for language acquisition, in line with Gold's formal analysis (Gold 1967). However our results show that when feedback is available and is used by our abstracted mirror neuron system to both analyse tentative production and to modify the learning process the production performance of a learning agent on the syntax acquisition task presented by the Reber Grammar can become behaviourally indistinguishable from that of a competent agent. Finally we note that since the system developed here is equally applicable to goal directed actions as it is to syntax acquisition the mirror neuron production/perception comparation system on which it is based could quite plausibly provide an explanation for the emergence of modern natural language processing from a mechanism previously adapted for complex goal directed actions synthesized from a *vocabulary* of more basic actions.

Note

1. For the purposes of our model we mean by central executive only that an instruction to spontaneously generate a particular sequence is initiated externally to the mirror system.

References

Cleeremans, A. (1993). *Mechanisms of Implicit Learning. Connectionist Models of Sequence Processing*, 35–74. Cambridge, Massachusetts: MIT Press.

Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning, 7*, 195–224.

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

Gold, E. M. (1967). Language identification in the limit. *Information and Control, 16*, 447–474.

Harnad, S. (2001). Minds, machines, and Turing: The indistinguishability of indistinguishables. *Journal of Logic, Language, and Information*, Special Issue on 'Alan Turing and Artificial Intelligence': (in press).

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*(3), 219–235.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in the Neurosciences, 21*, 188–194.

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

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

Womble, S. P. (2000). Connectionist model of a mirror neuron system http://www.his.sunderland.ac.uk/womble/tech2000a. Technical report, University of Sunderland, Hybrid Intelligent Systems Research Group, Informatics Centre, School of Computing and Technology, Sunderland, United Kingdom.