# A Neural Network Model of Causative Actions

Jeremy Lee-Hand

a thesis submitted for the degree of

Master of Science

at the University of Otago, Dunedin,

New Zealand.

November 13, 2013

**Abstract**

Many of the actions we perform are defined by the effects they bring about, rather than as stereotypical sequences of motor movements. For instance, to open a door we must perform an action which results in the door opening. This thesis is a study of causative actions of this kind. I first introduce the class of causative actions, reviewing evidence for their existence from psychology, neuroscience and linguistics. I then present a computational model of motor control which can learn how to perform causative actions. The model I propose is an extension to an existing model in the literature (Oztop et al., 2004). In Oztop's model simple reach-to-grasp actions are learned through reinforcement using touch sensations which are considered to be intrinsically rewarding. In my extension, I propose that observed external events can also function as rewards if they are observed while the agent is executing a motor action and attending to the object being acted upon. I demonstrate the feasibility of this proposal in an implemented neural network model of causative actions. The model is also novel in that it does not require the trajectory of the hand to be precomputed. I conclude by discussing possible links between my model of causative actions and an account of the syntax of causative constructions in human language.

**Acknowledgements**

I would like to express my gratitude to my supervisor Dr. Alistair Knott. Thank you for all your patience, ideas, discussion, and guidance.

I would also like to thank Tim Neumegen for his early work on this model and for technical help throughout the project.

Lastly, I would like to thank Rebecca for all her support and patience through this period of my life.

# Contents


**1 Introduction** **8**
1.1 Structure of the thesis . . . 9

**2 Effect-based representation of actions: a review of evidence** **10**
2.1 Theories and evidence from psychology . . . 10
2.1.1 Theories of action representation . . . 10
2.1.2 Evidence for effect-based action encoding from psychology . . . 12
2.2 Evidence for effect-based actions in neuroscience . . . 17
2.2.1 Theoretical Background . . . 18
2.2.2 Experimental evidence of effect representation . . . 21
2.2.3 Summary of neuroscience literature . . . 30
2.3 Linguistic Models of Causative Actions . . . 32
2.4 Summary of literature . . . 33

**3 A Review of Simulations of Reaching and Grasping Actions** **35**
3.1 Introduction . . . 35
3.2 Key Concepts in Computational Motor Control . . . 35
3.2.1 Rigid Body Dynamics . . . 35
3.2.2 Motor State . . . 36
3.2.3 Controllers . . . 37
3.3 Other simulations of infant motor action . . . 38
3.4 Grasp Project Model . . . 40
3.4.1 The use of perturbations in a model of motor control . . . 41
3.4.2 Considerations for a model of causative-actions . . . 42

**4 Modeling simple reach-to-grasp actions** **43**
4.1 Introduction . . . 43
4.1.1 Learning . . . 45
4.1.2 A progression of rewards for stages one and two . . . 46
4.1.3 Simplifications of this model . . . 46
4.2 Reach Network: Structure, Training, and Results . . . 47
4.2.1 Introduction . . . 47
4.2.2 Structure of neural network model . . . 48
4.2.3 Training the network . . . 48
4.2.4 Results . . . 49
4.3 A network for learning reach-to-grasp actions: the Perturbation Network 52

4.3.1 Introduction . . . . . 52
4.3.2 Issues with the perturbation network . . . . . 53
4.3.3 Structure of the network and training . . . . . 53
4.3.4 Training the Network . . . . . 56
4.4 Summary of the model . . . . . 59

**5 Implementation of causative actions** **60**
5.1 Model of causitive actions . . . . . 60
5.1.1 Introduction . . . . . 60
5.1.2 A proposal about the reinforcement signal that trains causative actions . . . . . 61
5.1.3 Structure of the model . . . . . 61
5.1.4 Learning . . . . . 62
5.1.5 Creating Objects . . . . . 63
5.1.6 Training the network . . . . . 64
5.1.7 Results . . . . . 65
5.1.8 Summary . . . . . 66

**6 Towards a Syntactic Interpretation of the Model of Causative Actions** **67**
6.1 Introduction to the Minimalist Model of Syntax . . . . . 67
6.1.1 X-bar Schemas . . . . . 67
6.1.2 Phonetic Form and Logical Form . . . . . 69
6.1.3 Movement . . . . . 69
6.1.4 Minimalist Structure of a transitive sentence . . . . . 72
6.1.5 VP-shells . . . . . 73
6.1.6 VP-shells and action-effects . . . . . 74
6.2 Introduction to Knott's Sensorimotor Interpretation of Logical Form . . . . . 75
6.3 Interpretation of My Model . . . . . 77
6.3.1 Interpretation of VP shells . . . . . 77
6.3.2 Higher VP . . . . . 78
6.3.3 Lower VP . . . . . 79
6.4 Clarification of interpretation . . . . . 80

**7 Discussion** **81**
7.1 Main Points of the Model . . . . . 81
7.2 Structure of the model . . . . . 82
7.3 Problems with this Model and Future Work . . . . . 82
7.3.1 Application of Perturbations . . . . . 82
7.3.2 Circuitry Granularity . . . . . 83
7.3.3 Performance . . . . . 83
7.4 Discussion of Syntactic Interpretation . . . . . 84

**References** **85**

# List of Figures


2.1 Abstract organisation of event codes, adapted from Prinz (1997) . . . . 11
2.2 The Simon task . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13
2.3 Hommel's adapted version of the Simon task. In experiment **A** the buttons were connected to the lights on the same side of the subject. In experiment **B** the buttons were connected to the lights on the opposite side of the subject. . . . . . . . . . . . . . . . . . . . . . . . . . . . 14
2.4 The experiment of Riggio et al. (1986) . . . . . . . . . . . . . . . . . . 15
2.5 Morin and Grant's experiment. . . . . . . . . . . . . . . . . . . . . . . 16
2.6 Diagram of a macaque brain, reproduced with permission from Knott (2012) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
2.7 Diagram of the human brain, reproduced with permission from Knott (2012) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
2.8 Proposed model of mirror neuron circuit involved in action recognition and action execution. The top pathway is the reach pathway through F2 and F4, the bottom pathway through the caudal and anterior intraparietal region is the grasp pathway. Adapted from Knott (2012) . . . 20
2.9 Proposed neural pathway of the action recognition pathway, adapted from Knott (2012) . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
2.10 **A:** Normal pliers require the monkey to close its hand to grip objects. **B:** Reverse pliers require the monkey to open its hand while gripping the pliers to grip an object. Reproduced with permission from Umiltà et al. (2008), Copyright (2008) National Academy of Sciences, USA. . . 22
2.11 Experiment from Umiltà et al. (2001), reproduced with permission from Umiltà et al. (2001). . . . . . . . . . . . . . . . . . . . . . . . . . . 23
2.12 Presented cues in Matsumoto et al. experiment, reproduced with permission from Matsumoto et al. (2003) . . . . . . . . . . . . . . . . . . . 25
2.13 An example of the movies that were shown to subjects. Reproduced with permission from Hamilton and Grafton (2006). . . . . . . . . . . . 28
2.14 In experiment one, subjects were able to vary the size of their grip aperture. In experiment two subjects were able to vary the orientation of their hand. Reproduced with permission from Tunik et al. (2005). . . 29

3.1 A perturbed goal motor state . . . . . . . . . . . . . . . . . . . . . . . 42

4.1 Completed computational model. a) shows the reach network; this network converts an object in retinal-coordinates to a vector of joint-coordinates which will bring the end-effector into contact with the object. This output is given as input to the perturbation network which calculates a perturbation of the goal motor state which will cause the model to grasp the object. Each of the components of the perturbation vector must be calculated individually and are interdependent so are calculated in series here. Since there are many possible perturbations which could lead to a successful grasp a discrete distribution of possible perturbations is calculated in which a winner-take-all (WTA) method is then used to select the perturbation with the highest activation. . . . . 44
4.2 Structure of the reach neural network model. . . . . . . . . . . . . . . . 48
4.3 A typical touch event from two angles. . . . . . . . . . . . . . . . . . . . 50
4.4 Error of testing the neural network. . . . . . . . . . . . . . . . . . . . . . 50
4.5 Error of the training of the neural network. Redo this in matlab for consistency. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 51
4.6 Successful Reach-to-touch locations after training. Successful locations are given by large points, unsuccessful locations by small points. . . . . 52
4.7 Original Structure of the Perturbation Network. . . . . . . . . . . . . . 54
4.8 New Structure of the Perturbation Network . . . . . . . . . . . . . . . . 55
4.9 Population Coding for the perturbation network output . . . . . . . . . . 56
4.10 Vectors of training data in cartesian space. . . . . . . . . . . . . . . . . 57
4.11 Vectors of training data in cartesian space gathered from grid experiments. 58
4.12 Objects that can be successfully reached in Cartesian space. Successful locations are shown as large points, unsuccessful locations as small points. 59

5.1 Structure of the perturbation network. . . . . . . . . . . . . . . . . . . 63
5.2 Squashable Object . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 64
5.3 Bendable Object . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 64

6.1 An X-bar schema . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 68
6.2 The phrases present at the specifier and complement positions (*the man* and *the cup*) of this XP are arguments of the head (*grasped*). . . . . . . 68
6.3 The logical form of a simple sentence. IP is the inflection x-bar, AgrP is the agreement x-bar, VP is the verb x-bar, and DP is the determiner x-bar. Evidence for these different projections can be found in Carnie (2002). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 69
6.4 DP movement . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 70
6.5 A verb head raises to its inflection in the IP projection . . . . . . . . . 71
6.6 Left: shows the overt movements in this sentence. Right: covert movements that take place after pronouncement. . . . . . . . . . . . . . . . 72
6.7 LF structure of a transitive sentence . . . . . . . . . . . . . . . . . . . 72
6.8 LF of a double-object sentence using VP shells . . . . . . . . . . . . . . 73
6.9 LF Syntax tree of *The man squashed the ball* . . . . . . . . . . . . . . . 74
6.10 Syntax tree interpreted as a description of a sequence sensorimotor operations. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 76

6.11 Sensorimotor characterisation of X-bar schema . . . . . . . . . . . . . . . 76
6.12 VP shells of figure 6.9: *the man squashed the ball* . . . . . . . . . . . . . 78

# Chapter 1

# Introduction

Human actions are performed in order to change the world around us. The effects that an action bring about are almost always more important than the action itself; all effects can be caused with an almost infinite number of different actions. Modern models of human motor acquisition and performance often focus on simple reach-to-grasp learning. In this thesis I propose a model of the acquisition and execution of **causative actions**: an action which is defined by the effects that it brings about in the world.

As part of this research I have developed a computational model using neural networks, motivated from the literature in infant development and neurobiology. The focus of this model is to develop several ideas that have mostly been overlooked in the literature.

The first idea that I have attempted to develop relates to the evolution of rewards involved in motor learning. Most computational models of motor development are based on reinforcement learning: the idea that the agent learns to perform actions which lead to rewards. In many of these models, the reward signal is unchanging throughout development, or if it changes, does so in a relatively simple way - for instance it may decay in relation to the number of times that the action has been performed or to the degree to which the reward is expected (using temporal difference methods (Sutton, 1988)). My proposal is that for the causative actions which I am interested in there is a qualitative change in the stimuli which generate rewards: simple behaviours are trained by simple haptic rewards, but for causative actions, the perceptual mechanism which recognises episodes occurring in the external world is able to deliver reward signals. This means that the effects of action become intrinsically rewarding themselves. I use this mechanism to produce my model of causative action acquisition and execution.

The second idea that I want to develop involves the pre-computed trajectories that are present in most computational models of motor action. Although these seem intuitive there is mounting evidence that this pre-computation does not occur; instead it is likely that this complex trajectory is gradually developed over the course of a movement through feedback mechanisms. The model that I produce to replicate this uses the idea of perturbations of the goal hand state.

The third idea I have attempted to develop is motivated by an interpretation of sensorimotor cognition and language due to Knott (2012). Knott proposes an interpretation of modern linguistic theory that shows a close relationship with the sequence of sensorimotor events which make up human action. I will make use of this theory to try to explain the structure of the model of causative actions that I have produced.

## 1.1 Structure of the thesis

In chapter 2 I begin with a review of literature in order to motivate the conclusion that many human actions are defined by the effects they cause. This is not an abstract idea, many studies support the idea that the effects of action are important. Loosing the feedback of the effect of action can drastically affect the performance of an agent. In chapter 3 I examine other computational models of human motor action and leverage some of these ideas in the construction of my own model. Chapters 4 and 5 are devoted to the two parts of my motor model. The first deals with learning how to represent action trajectories in general, given that full trajectories cannot be precomputed. This part of the model only learns simple reach-to-grasp actions and has nothing specifically to do with causative actions. The second part of my model deals specifically with an extension to the simple reach-to-grasp model which is capable of learning to perform actions which bring about certain effects. In chapter 6 I develop a link between my model, linguistics, and sensorimotor cognition using a modern linguistic theory. This chapter is a distinct part of this thesis: it looks for links between the structure of causative actions in a motor model and in a model of natural language syntax. By doing this I hope to provide computational support for an embodied model of language. In chapter 7 I resolve this thesis with a discussion of the ideas contained, problems that still remain, and potential future solutions to these problems.

# Chapter 2

# Effect-based representation of actions: a review of evidence

In this chapter I will review evidence for a class of actions which are defined by the effects that they bring about in the world. Sections 2.1 and 2.2 of this chapter cover studies from psychology and neuroscience which indicate that some motor actions are represented in the brain by the effects they are cause in the world. The studies I review in this chapter focus on reach and grasp actions which are the predominant motor actions used in psychological and neurophysiology studies. Although actions must also encode some information about motor movements many actions are performed only to bring about some desired effect, thus it is not unexpected that the cognitive representation of actions includes details of the actions effects. Section 2.3 briefly covers some linguistic arguments which also suggest the importance of action's effects.

## 2.1 Theories and evidence from psychology

### 2.1.1 Theories of action representation

Two influential theories which explore the relationship between perception and the motor system are the theory of common coding (Prinz, 1997) and the theory of event coding (Hommel et al., 2001a).

In the theory of common coding Prinz proposes **common coding** and the **action effect** principle. The common coding principle states that perceived events and planned actions share a common representational level. The action effect principle follows from this and states that actions are most likely planned and controlled around the effects that they are expected to bring about in the world.

The theory of event coding (TEC) is a successor of the theory of common coding and has been developed into a theoretical framework of action planning and perception. It shares the common coding and action effect principles of Prinz's common coding and develops them to include some structural details of the common code representation.

This representation is known as an **event code** and is a temporary integration of other existing neural **feature codes**. A feature code represents some particular property of an external event which can be perceptual, tactile, kinesthetic etc. Event codes can be planned actions or perceived events, TEC assumes that all that is being represented internally is the features of an external (planned or perceived) event.

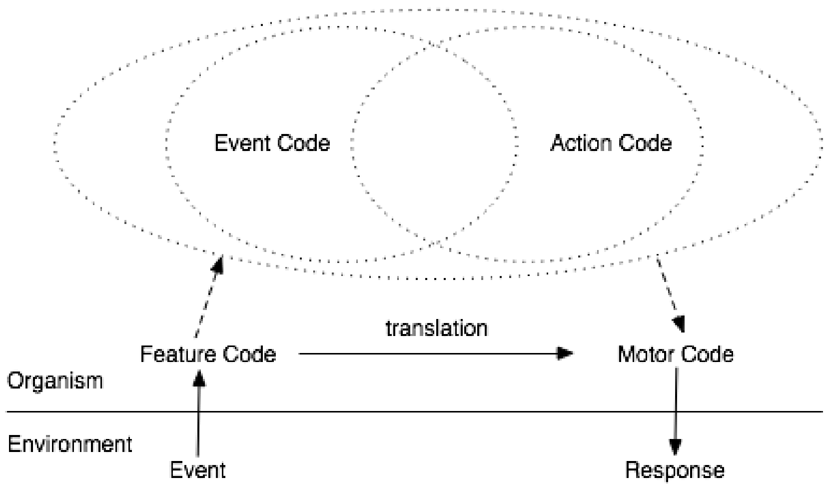



Figure 2.1: Abstract organisation of event codes, adapted from Prinz (1997)

Figure 2.1 shows the abstract representation of an event code. Feature codes are temporarily bound to each other into an event code. If one feature code is activated it will partially activate the other feature codes that it has been bound to. This also allows intentions to be represented as priming of a particular feature code in order to activate all plans involving that feature.

An important aspect of this model is the **bidirectional links** between percieved actions and their sensory effects. This bidirectionality is the only plausible way for this model to deal with humans ability to infer the effects of novel actions and to learn associations between an observed effect and the action that caused it; single linkages in either direction between codes would only allow one of these. The **action codes** shown in figure 2.1 will create associations with event codes and create action-effect associations.

The TEC has been criticised for its vagueness and underspecification of key ideas in the theory. Feature codes are explained in examples of color and size of an object. The specifics of how you would represent an action effect as a feature code are never discussed. The TEC also fails to explain anything about representing multiple competing action plans or whether they expect this to occur at a different level of representation. The authors do state that the TEC is not a complete theory and does not deal with "early perception and late action" (Hommel et al., 2001b), however the lack of coverage of how a **distal** event is represented is surprising considering how important it is to the TEC.

The authors' response (Hommel et al., 2001b) to these criticisms emphasised the TEC as a theoretical framework (rather than a complete theory) and claimed the TEC was deliberately vague so that different models of action planning and perception could be integrated into it. The authors answered the criticisms of a lack of specificity of a feature code with the response that to identify specific feature codes may not be possible. More recent work by one of the authors (Hommel, 2004, 2009) have added some aspects of motor control into the TEC but little seems to have been done to demystify how feature codes are represented and the sort of information that they can encode.

### 2.1.2 Evidence for effect-based action encoding from psychology

**The Simon Task** In order to understand some of the experiments that are adduced in support of the TEC and effect based representation of actions, the reader will need to understand the stimulus-response compatibility effect known as the **Simon effect**. The Simon effect was originally thought to be a reduction in response performance due to an incompatible spatial relationship between the stimulus in a task and the location of the response action. It was first discovered by J.R. Simon (Simon and Rudell, 1967) using the following experiment which has become known as the Simon task.

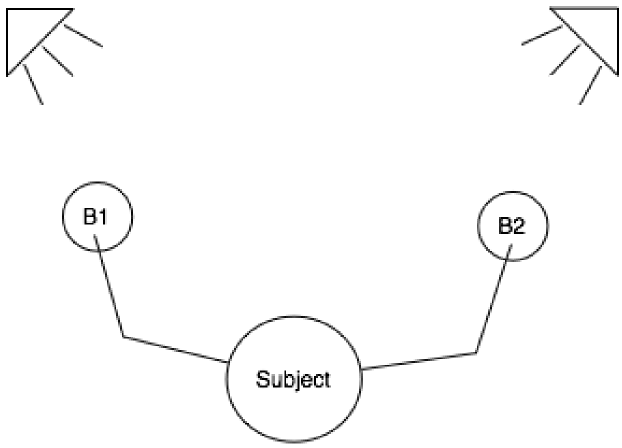



Figure 2.2: The Simon task

Figure 2.2 shows the Simon task. Subjects had a choice of two buttons to press, one for the left hand, and one for the right. The stimulus for the task is an auditory signal which can either be a high or low pitched tone issued from a speaker to the left or right of the subject. Subjects were told to press a particular button in response to an auditory stimulus but to ignore the direction the sound came from. Simon found that the subject's response times were usually much faster if the stimulus shared the same spatial location as the response. This was thought to be due to **locational correspondence**, i.e. the correspondence between the locations the stimulus were presented and the location the response actions were performed (both the stimulus and the button were on the same side of the subject).

**Hommel, 1993** One problem with the Simon task is that it does not completely dissociate the location of the subject's actions from the location of the effect that they cause. One adapted version of the Simon task (Hommel, 1993) found that the cause of the Simon effect seems to be the relative locations of the stimulus and the effect of its response.

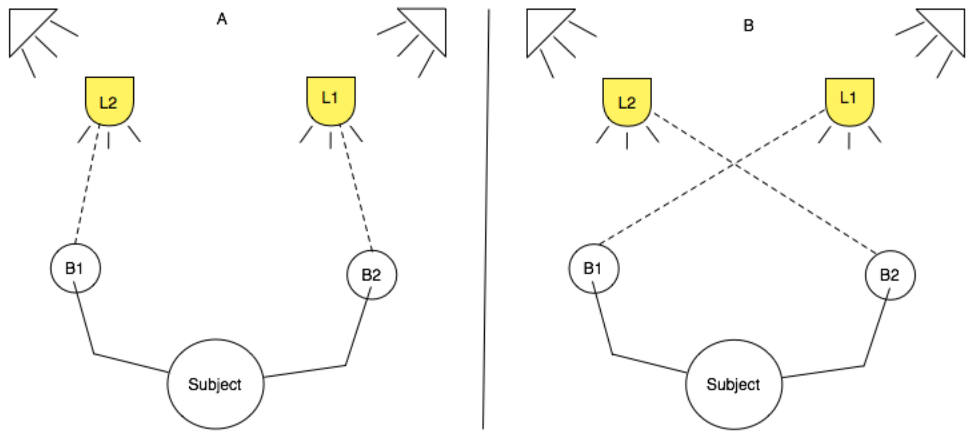



Figure 2.3: Hommel's adapted version of the Simon task. In experiment **A** the buttons were connected to the lights on the same side of the subject. In experiment **B** the buttons were connected to the lights on the opposite side of the subject.

Figure 2.3 shows Hommel's adapted version of the Simon task. This is almost the same as the original Simon task with the exception that pushing a button causes a light to be illuminated. There are two lights, one for each button, placed to the left and right of the subject. In the first experiment the lights were attached to the buttons on the same side of the subject as shown in 2.3:A. In the second experiment the lights were attached to the buttons on the opposite side of subject as shown in 2.3:B. Hommel found that the subject's reaction speeds were reduced when the stimulus (the auditory tone) shared the same spatial location as the effect produced by the subject's response (the light coming on). This was true even when, in experiment two, the subject's action did not have the same spatial location as the stimulus.

**Riggio et al., 1986** Hommel's experiment is similar to another by Riggio et al. (1986) which investigated the effect of crossing subject's hands in response time tasks. Previous experiments had found that crossing subject's hands while performing a task had increased response time and reversed the direction of spatial compatibility. This means that subjects who crossed their hands responded more quickly to a stimulus from their left with their right hand than they would if the they responded to the same stimulus with their left hand. Although this reversal occurs, the overall response times of participants while crossing their hands was slower.

Subjects were placed in front of a screen on which a visual stimulus was presented. They were told to respond to this stimulus as quickly as possible by pressing one of two buttons using one of two plastic sticks which they held in each hand. During trials the sticks were either crossed or uncrossed as shown in figure 2.4.

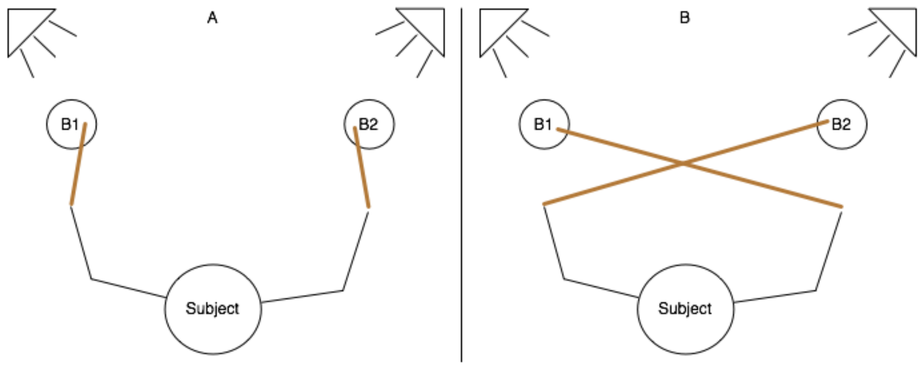



Figure 2.4: The experiment of Riggio et al. (1986)

The results of the experiment showed an overall increase in response time and a reversal in the direction of spatial compatibility when the sticks were crossed. This reproduction of the effect of crossing hands without actually crossing them showed that the location of the hand in relation to the stimulus has little to do with an increase in response times. Riggio et al. postulate that "what matters for spatial compatibility is the position of the response goal not the effector" (Riggio et al., 1986). More recently experiments in the field of neuroscience have shown that tools seem to be integrated into the motor system of monkeys, if this is true then Riggio et al. may be showing that the end-effector becomes the stick rather than the hand.

**Morin and Grant, 1955** Another experiment that provides evidence for the representation of actions by their effects is a stimulus-response experiment by Morin and Grant (1955). The experiment was designed to test the effects on learning of different levels of spatial correspondence of stimulus and response. The experimental setup is shown in figure 2.5. Eight red stimulus lights were placed above eight green response lights. Below this there were eight keys which were connected to the green response lamps in an arbitrary manner not visible to the subjects. When subjects pressed a key the green bulb it was connected to was illuminated.

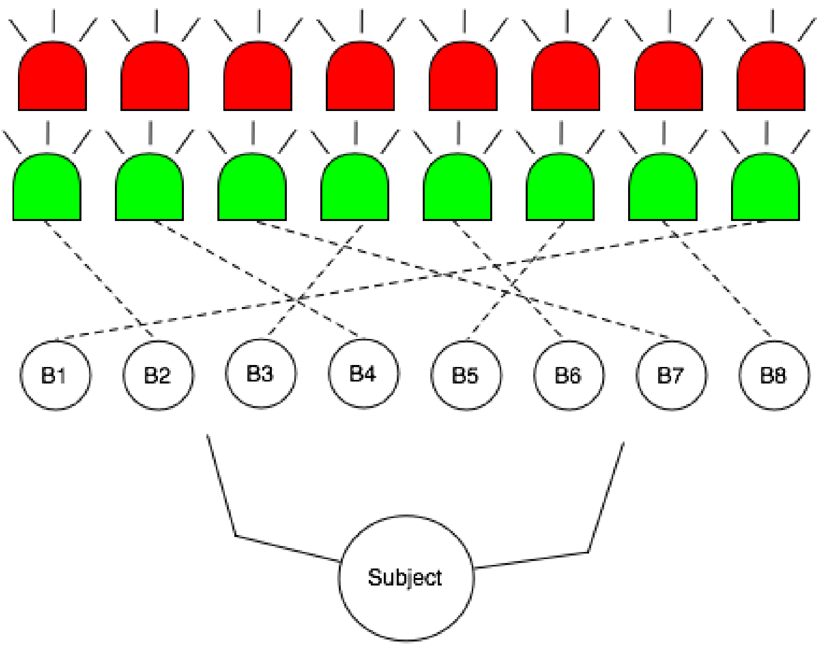



Figure 2.5: Morin and Grant's experiment.

Subjects first had to learn the mappings between keys and response lights, a difficult task given the number of possible combinations. After subjects had successfully learnt the mappings they underwent eleven testing phases each with 25 patterns of two red stimulus lights. In the first nine phases, when subjects were presented with a red stimulus light they simply had to press the key which would make the green response light below it illuminate. As the trials continued all subjects' response times decreased. In the tenth phase the green response lights were disconnected from the keys removing the visual feedback from the task. Subjects then had to press the key that they had learned would light the correct response light but would not get visual feedback. In this phase the subjects' performance immediately dropped to pre-training performance.

All subjects were accurately able to produce a picture of the mappings between keys and response lights but for some reason their response time was hugely increased without the visual feedback. At the time it was suggested that "they could not decode rapidly enough for fast performance"(Morin and Grant, 1955). One possible explanation of this is that what was encoded was the effect of action (illuminating the correct green light) rather than the action. This would mean that when the effect was removed the action would have to be relearnt with or without a new effect.

**Verschoor et al. 2010** Work by Verschoor et al. (2010) attempted to find evidence for action-effect associations in infants. Three groups of infants were involved in the test, twenty-two 9-month olds, twenty-one 12-month olds, and twenty-two 18-month olds. Infants were placed in front of a large screen which was used to produce audio and visual stimuli. A large touch sensitive key was placed in front of the infant.

Infants were taught to respond to two types of stimuli, action-independent and self-produced. When action-independent stimuli were presented the infants caregiver held their hands so they were unable to press the button. Each stimulus was presented to the infant until it had attended to the screen which was monitored by the experimenters. In self-produced phases infants were free to press the button if they desired which would cause a short audiovisual stimulus to start. The infants caregivers were told to encourage the child to press the key during these learning phases.

Testing followed immediately after learning (30 seconds gap) and consisted of the presentation of a learned action-independent stimulus (no action-effect mapping), self-produced stimulus (action-effect mapping learned) or no stimulus. Babies were free to touch the key during all testing phases and were expected to touch the key for self-produced stimuli and ignore it for action-independent stimuli and no stimuli.

Experimenters recorded the latency of each infant's actions, the number of motor responses and whether they needed help from their caregiver. As well as this the number of undetected (attempted movements that were undetected by the button) actions were also recorded.

The results of the experiment show clearly that 18-month olds were able to learn an action-effect association, they were able to correctly perform actions 50% of the time and ignored the key 70% of the time during action-independant stimuli. In the two younger age groups the performance was statistically similar to chance.

Observed learning of action-effect associations in young infants adds support to the idea that some actions are defined by their effects as they are able to associate an arbitrary stimulus with the motor action producing it suggesting at least a temporary storage of an action-effect.

## 2.2 Evidence for effect-based actions in neuroscience

Studies in neuroscience can find evidence for the neural correlates of cognitive tasks using several different techniques. These can involve single-neuron recordings in monkeys; in humans they more often involve brain imaging techniques, which are less invasive.

In this section I will present studies which attempt to identify the neural regions that represent the effects of simple motor actions, and that suggest that these effects are an integral part of the neural representation of actions.

### 2.2.1 Theoretical Background

In this section I will reproduce basic models of the neural circuits involved in action execution and recognition. These models will help to situate the neural correlates of intended action-effects or goals as part of the overall neural circuits involved in action. I will focus on a very simple, and well-studied action: reach-to-grasp and orientation of the hand. Most of the evidence presented in this summary relates to data from a macaque monkey. For a more thorough review of the models presented and the evidence that underlies them see Knott (2012). This section discusses human and macaque neural regions which can be seen in figures 2.6 and 2.7.

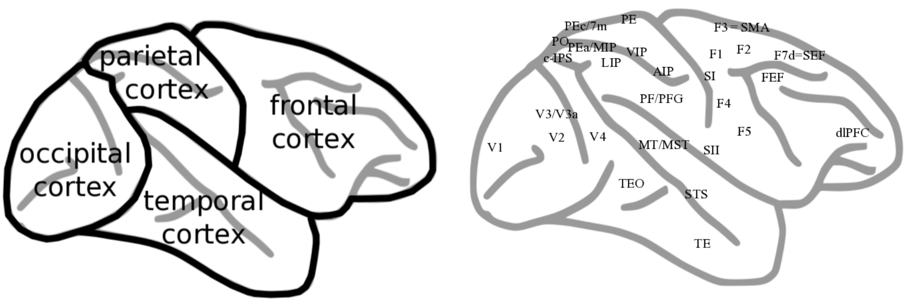



Figure 2.6: Diagram of a macaque brain, reproduced with permission from Knott (2012)

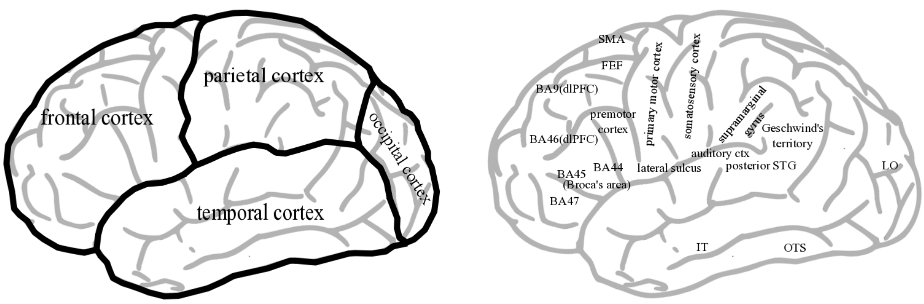



Figure 2.7: Diagram of the human brain, reproduced with permission from Knott (2012)

**The action execution pathway** The basic idea of the action execution pathway is to deliver motor impulses which cause an **effector** — in this case, the agent's hand — to travel through space to reach a target — in our case, the object to be grasped. The model of the neural circuit for execution of a reach-to-grasp action that is explained assumes that there are two relatively separate circuits: the **reach pathway** and **grasp pathway** (Jeannerod, 1996). The reach pathway directs movements' of the shoulder, elbow and wrist joints whereas the grasp pathway guides movements' of the fingers and thumb. These two neural circuits converge to the primary motor cortex (F1) which controls low-level motor action.

The reach pathway is a neural circuit in the superior parietal lobule, the premotor cortex, and the motor cortex. From an abstract perspective the reach pathway is involved in the creation of something similar to a **movement vector** which specifies the direction and distance the hand must travel to reach the target. When a movement vector has been created it will then be used dynamically as the agent tracks the progress of the effector through space. Various kinds of feedback and feed forward control are likely to be used to control action such as visual feedback and a forward model of arm position.

Two key areas involved in the reach pathway are the premotor regions F2 and F4. These regions seem to code a mixture of visual and tactile information in an effector-centered coordinate system (distance to target from current position); they compute something similar to a movement vector. Area F4 encodes information about the head and receives input from ventral intraparietal sulcus (VIP) (Schlack et al., 2005). F2 on the other hand seems to encode information about the body and reach actions within the animal's perispace (Fogassi et al., 1999). Area F2 and F4 both project to area F1, the primary motor cortex.

Another important property of the reach pathway is that it seems to perform some action selection. The areas leading to and including area F2 contain **matching cells** and **condition cells**. Matching cells are responsive to a movement in a particular direction but also when an observed object affords such an action. This suggests that matching cells encode possible motor commands. Condition cells also fire in response to an action in a particular direction but also in any context were such a movement is desirable. These cells may encode prepared or desired actions. If this is true then area F2 may select one of these actions and project it to area F1.

The grasp pathway begins at the caudal intraparietal sulcus (cIPS) which represents the shape and orientation of observed objects(Shikata et al., 2003). cIPS then projects to the anterior intraparietal (AIP) area which contains cells that fire when performing specific grasp actions: for instance, one neuron might only fire if the animal performs a 'power grasp', while another neuron might only fire when it performs a 'precision pinch'. AIP also contains neurons that only fire when the animal observes an object that **affords** a specific action: for example, one neuron might fire if an object that affords a 'precision pinch' is observed (e.g. a pencil), while another neuron might fire if an object affords a 'power grasp' is observed (e.g. a cup). The AIP area then projects to the premotor area F5, an area which contains **mirror neurons** which are discussed in more detail below. The reach and grasp pathways are shown in figure 2.8.

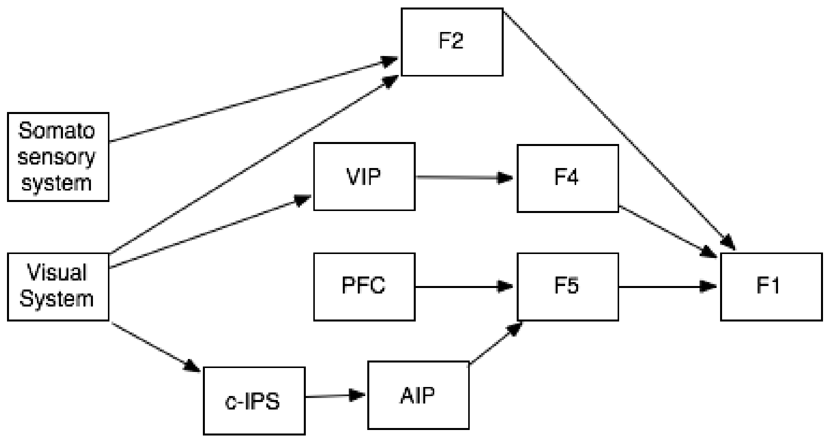



Figure 2.8: Proposed model of mirror neuron circuit involved in action recognition and action execution. The top pathway is the reach pathway through F2 and F4, the bottom pathway through the caudal and anterior intraparietal region is the grasp pathway. Adapted from Knott (2012)

**Mirror neurons and the action recognition pathway** Mirror neurons are neurons which discharge when an agent executes a particular grasp action — for example, using their whole palm to grasp an object — but also when the agent observes that action being performed by another agent. This class of neurons were first discovered in macaque monkeys by di Pellegrino et al. (1992). The experimenters noticed that some neurons in the premotor area F5 which responded when the monkey performed grasping actions also selectively responded when the monkey observed an experimenter performing the same action. Since this discovery, there has been a huge amount of research invested into mirror neurons, and many theories exist of how mirror neurons may relate to understanding intention, empathy, language, and autism among other things. They have also been suggested as the common representation of action plans and observed events by Hommel et al. (2001a).

Since mirror neurons were first discovered in the premotor cortex of the macaque they have also been discovered in the inferior parietal lobule (Fogassi et al., 2005). The mirror neuron system is now considered to comprise the ventral premotor cortex (F5) and the rostral part of the inferior parietal lobe (PF). Neurons with similar properties have also been found in the superior temporal sulcus (STS) but these only fire when observing the actions of others and never when performing that same action. STS also displays activity for a much wider variety of actions than F5. For a review see Rizzolatti and Craighero(2004).

The neural pathway for **action recognition** (shown in figure 2.9 appears to begin in the STS. The STS receives input from the primary visual cortex and sends its output to PF which then sends its output to F5. As well as these regions the caudal and anterior intraparietal sulcus are shown (cIPS and aIPS) which also project to F5. These form the grasp pathway and other evidence also suggests that aIPS also represents the goal state of an grasp action when it is observed or when it is performed. The dorsolateral prefrontal cortex (PFC) projects to F5 as well and is believed to represent multiple alternative action plans.

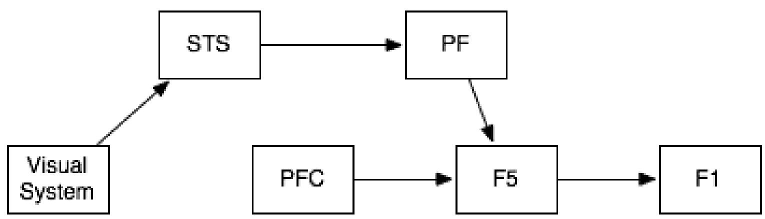



Figure 2.9: Proposed neural pathway of the action recognition pathway, adapted from Knott (2012)

### 2.2.2 Experimental evidence of effect representation

In this section I will review neuroscience studies that present evidence for an effect based representation of actions. Most of the studies that I review have searched for the neural correlates of goals. A goal can be a prepared action, or the intended effect of an action; in this thesis I will obviously attempt to concentrate on studies of the latter.

The first half of this section presents single-cell recording studies performed with monkeys; the second half presents imaging studies from humans. Although imaging studies are less clear than cell-recordings they still provide important evidence from humans which usually cannot be gathered using invasive methods.

**Umiltà et al. 2008** A study by Umiltà et al. (2008) cleverly dissociated between the effects that a motor action bought about and the physical action itself. The task involved using two different designed sets of pliers shown in figure 2.10. The normal pliers required the monkey to open its hand to open the pliers and close its hand to close the pliers; while the reverse pliers required the monkey to close its hand to open the pliers and open its hand to close the pliers. The monkeys were made to pick up an object with a set of pliers. Recordings were taken from neurons in the premotor area F5 (recall from 2.2.1 that F5 seems to encode grasp types) and the primary motor cortex F1 during action.

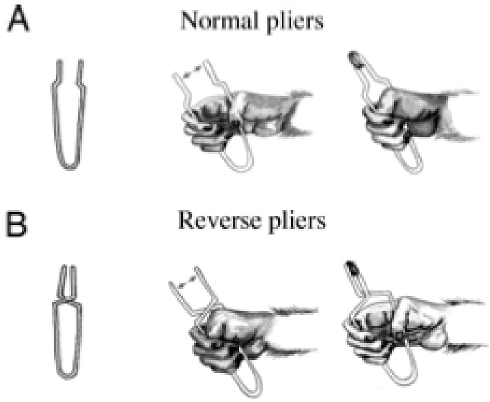



Figure 2.10: **A:** Normal pliers require the monkey to close its hand to grip objects. **B:** Reverse pliers require the monkey to open its hand while gripping the pliers to grip an object. Reproduced with permission from Umiltà et al. (2008), Copyright (2008) National Academy of Sciences, USA.

All neurons recorded during action discharged in response to the monkeys using the plier in grasping actions. The grasping action can be broken into three phases: opening the pliers, closing the pliers, and holding the object with the pliers. Regardless of the type of pliers the monkeys used, normal or reverse pliers, the same pattern of neural activation was observed in the neurons that were recorded even though distinct hand motions were required to perform these actions. 32.7% of recorded neurons were found to begin discharge during the opening of the pliers and peaked prior to the closing of the pliers, 50.9% of recorded neurons discharged exclusively during plier closure, 12.7% discharged during plier closure and continued firing during holding of the object, and 3.7% of neurons discharged only during opening of the pliers.

The fact that the neurons always discharge with the same pattern, regardless of the grasp actions the hand performs, shows that these neurons do not encode the hand

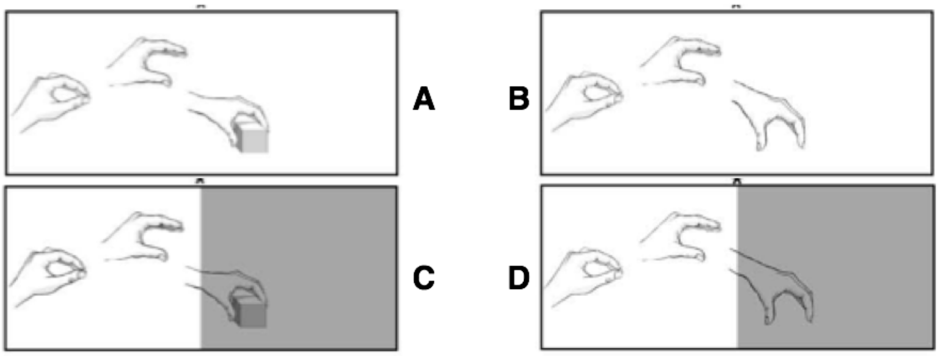



Figure 2.11: Experiment from Umiltà et al. (2001), reproduced with permission from Umiltà et al. (2001).

movements that are required to perform action. This may mean that the neurons encode effects bought about by each phase of action.

**Umiltà et al. 2001** The mirror system is believed to be involved in the recognition, and perhaps imitation, of actions performed by another agent. However actions observed in the world are frequently occluded by other objects and we are still able to recognise the actions agents perform. A study by Umiltà et al. (2001) investigated this by performing reach-to-grasp actions and recording neurons from the F5 premotor area of two macaque monkeys.

The experimenters performed four different experiments as shown in figure 2.11. Monkeys were shown a hand performing a reach-to-grasp action leading to a stable grasp on an object (figure 2.11:A) and the imitation of a reach-to-grasp action without an object (figure 2.11:B). They were also shown these two actions with the condition that the second half of the action (the grasping phase) was occluded by a screen (figure 2.11:C&D). All recorded mirror neurons discharged when viewing the experimenters performing reach-to-grasp action A. 51% of the mirror neurons recorded also responded when viewing action C. None of the mirror neurons discharged when viewing actions B or D.

The results of this experiment show that all mirror neurons require the final state of the a reach-to-grasp action to be achieved to discharge. It also shows that some mirror neurons are able to infer the grasp type when the subject is unable to view the final grasp as long as the goal object has been observed prior to this. These neurons need to 'know' or be able to infer the final effect of a grasp action in order to discharge and may be encoding this state or the achievement of this state. This study highlights the importance of the final goal state in the mirror system and the grasp pathway of movement.

**Iriki et al. 1996** Another study of tool use by Iriki et al. (1996) focussed on the reach pathway rather than the grasp pathway, but under similar conditions to Umiltà et al. (2008). This study sought to discover the neural correlates of the proposed **body schema**, an internal model of the body's motor capabilities which can be temporarily expanded through the use of tools. The experimenters believed that a body schema would involve information from the visual and somatosensory system, integration of these streams occurs in the area of intraparietal-sulcus (IPS). Iriki et al. performed an experiment targeting the anterior region of IPS which previous work has suggested contains a schema of the hand in space(Colby and Goldberg, 1999).

The experiment was performed with two monkeys which were trained to sit still on a primate chair. Food pellets were placed on a table and the monkeys were trained to retrieve these with a small rake. 59 neurons in anterior IPS which responded to both somatosensory and visual stimulation were recorded during the experiment. Some of these neurons (15) fired when the experimenters bought a pellet of food within reaching distance of the monkey. This area was expanded after the monkey had used a tool to retrieve a distant pellet. After one to five minutes without using the tool these neurons would revert back to their normal firing patterns.

This suggests that these neurons do not encode the areas that the monkey can reach, rather they seem to encode the areas that the monkey is able to affect by hand movements.

**Matsumoto et al. 2003** Matsumoto et al. (2003) investigated the mechanisms involved in goal-based action selection. They trained monkeys to perform a visually cued, asymmetrically rewarded MOVE/NO-MOVE task with **reversals**. The experimenters attempted to force the monkeys to make action selections during the task based on the anticipation of a reward.

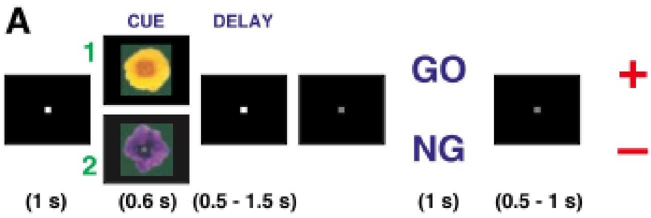



Figure 2.12: Presented cues in Matsumoto et al. experiment, reproduced with permission from Matsumoto et al. (2003)

Figure 2.12 shows the experiment. The monkeys were first forced to fixate a point on the screen after which they were presented with a stimulus (CUE phase). After another period of fixation the monkeys performed a GO (move a joystick) or NO-GO (hold joystick without movement) action based on the stimulus. After another fixation period a reward was provided if the response was correct. During all stages of the experiment neurons from the PFC were recorded.

The experimenters found that some neurons in the mPFC of the monkeys responded selectively to a particular anticipated reward condition (reward + or reward -) during the fixation period after the stimulus and before action. This activity suggests that medial PFC encodes an intended action effect (the monkey's action will bring about a reward).

**Chaminade et al. 2002** A **positron emission tomography** (PET) study by Chaminade et al. (2002) attempted to determine the neural basis of specific aspects of imitation in adult humans. These aspects were the ability to imitate by observation of the **means** (the motor movements involved in the action) and imitation by observation of the **goal** (the final position of the end-effector in a motor action).

The neural bases of these two factors where investigated by recording the **regional cerebral blood flow** (rCBF) of human subjects. The experiment was done in two phases, an observation phase and a testing phase. In the observation phase of each trial, subjects were shown a video clip of an experimenter constructing a small structure from lego. For example, the hand could pick up a small block and place it on top of another block. The final goal state differed in each video and the block was not always placed onto another block. Each trial was varied along two factors: how much of the motor action was shown in the observation phase, and the required response to the video in the test phase. In the test phase subjects were required to perform the motor action shown in the video or an action of their own choice using the blocks.

Videos shown to the subject could show the whole act of the movement, a static image of the final goal state or the first portion of action without the goal state (the means of action from which the observer could infer the final goal state). The subjects were also required to respond to the act in one of two ways, either freely (ignoring the video shown and performing some random movement of any lego block) or imitation of the action observed in the video. Imitation of the action always required the subject to attempt to perform the whole action even if they only observed the means of action or the goal state. In these cases, the subject was required to infer the goal state or the action sequence respectively. During the test phase subjects rCBF was recorded with a PET scanner.

Chaminade et al. predicted that imitation would involve the parietal and prefrontal cortical regions, since these are thought to be involved in action control and higher level cognitive tasks respectively. They also predicted that imitation of the goal would particularly activate areas involved in representing the movements of others (premotor area), since there movements have to be reconstructed through inference. Likewise, they predicted that imitation of means would activate areas involved in identifying the intended goal (medial prefrontal areas) for similar reasons.

PET images were categorised along the following conditions: IW (imitation of action/whole action shown), IG (imitation of action/goal shown), IM (imitation of action/means shown), FW (free response/whole action shown), FG (free action/goal shown) and FM (free action/means shown). By subtracting the common areas involved in different scans it is possible to infer the parts of the brain that are only involved in a specific condition.

In general imitation, after removing areas also involved in any observation or action ((IW - FW) + (IG - FG) + (IM - FM)), activated the areas shown below. The areas seem reasonable as most of them contain mirror neurons or are believed to be part of the mirror system circuit.

- Left superior temporal lobe
- Inferior parietal lobe bilaterally
- Left posterior superior temporal sulcus
- Left superior temporal gyrus
- Left lateral orbital gyrus

| Brain areas activated | |
|---|---|
| Imitation after observation of only goal | Imitation after observation of only means |
| Right dorsolateral PFC | Right dorsolateral PFC |
| Left PFC | Medial PFC |
| Left lateral cerebellum | Left lateral cerebellum |
| Right medial cerebellum | Right medial cerebellum |

Table 2.1: Brain areas activated during imitation of observed means and goals

Table 2.1 shows areas that were activated in the brain exclusively when imitating an observed goal or observed means of action. The authors state that the right dorsolateral PFC (right dlPFC), which is believed to be involved in creation of multiple alternative action plans, may contain a representation of the goal. This is because the right dlPFC was strongly activated when observing the goal than when observing the means suggesting that it is encoding the goal in some way. I would argue that this interpretation is erroneous as they have also factored out the parts of the brain used in the imitation of the whole action. This would mean that the right dlPFC would only encode the goal if the rest of the action is not present which makes little sense.

The left premotor cortex is thought to be involved in the preparation of actions. The authors argue that the absence of observation of the means would require the complete construction of the required action while, in the case of imitation of means, the observation would have already prepared the required action. The medial prefrontal area is thought to be involved in reading the intentions of others, its activation during means has been interpreted as an attempt to extract the intention (the end goal state) from the observed partial action.

The finding of Chaminade et al. that the right dorsolateral PFC has a representation of the goal is quite reasonable considering that it is thought to produce alternative action plans. The fact that this area is not activated during observation and imitation of an entire action could mean that if there is no ambiguity to the action that you want to perform then we obviously do not need to generate alternative plans. However we would still need a representation of the goal so dlPFC cannot be the only place that goals are represented in the brain.

#### Goal representation in the human brain

An experiment by Hamilton and Grafton (2006) used **functional magnetic resonance imaging** (fMRI) habituation to determine the parts of the human brain that represent action goals. **Repetition suppression** is an effect observed in many parts of the brain in which neural activity is attenuated by repeated presentation of a stimulus. This effect is believed to occur in neural populations that are encoding any information from a stimuli that is not changing.

The experimenters searched for parts of the brain where representation suppression occurred during repeated presentation of a goal. Hamilton and Grafton predicted that these areas would be involved in the representation of goals in the brain.

Subjects were presented with a sequence of nine movies, each 2.5 seconds long. Each movie was of the arm of a woman reaching for one of two similarly shaped objects. Each movie varied by two factors: the trajectory of the reach action (left or right) and the goal of the reach action (left object or right object). The first movie presented was designated as 'new' and in each successive movie the trajectory and/or the goal could differ (novel) from the previous movie or could be the same (repeated) as the previous movie as shown in figure 2.13. Subject's **blood-oxygen-level dependence** (BOLD) was recorded at all times while watching the movies. The goal objects were different but were chosen to be of very similar size and require the same sort of grasp to pick up.

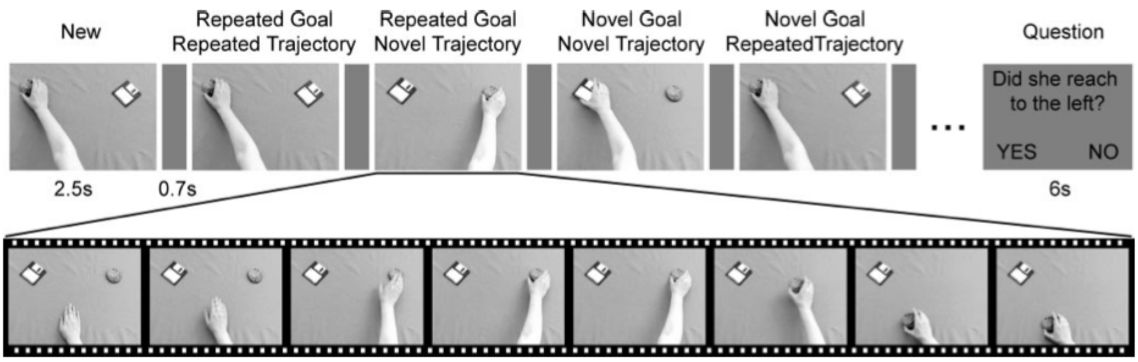



Figure 2.13: An example of the movies that were shown to subjects. Reproduced with permission from Hamilton and Grafton (2006).

Repetition suppression predicts that observation of a repeated goal will cause a systematic reduction in activation in an area coding goal states. These areas will not show the repetition suppression effect when the trajectory is repeated in videos. This pattern of activation was observed in all subjects anterior intraparietal sulcus aIPS). Damage to this area was already known to impair patients ability to interpret the actions of others.

Hamilton and Grafton's result agree's with the results of an experiment by Tunik et al. (2005). Patients with lesions to the aIPS had shown deficiencies in pre-shaping their hands during reach-to-grasp actions. Although these patients could not pre-shape their hands correctly the grasp movement they performed was unaffected. Tunik et al. used **transcranial magnetic stimulation** (TMS) to disrupt activity in aIPS during adaptation of a reach-to-grasp action in order to discern the cause of this effect.

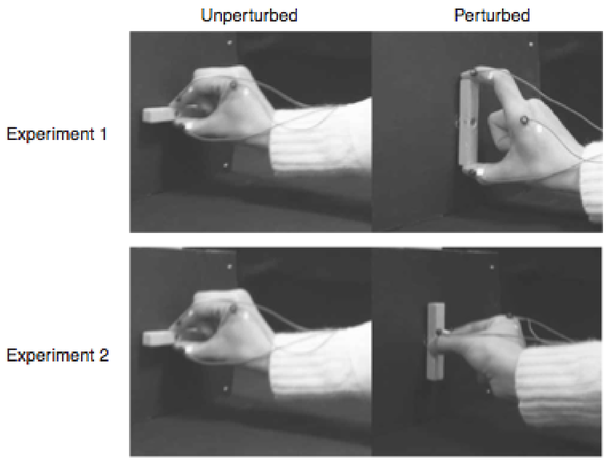



Figure 2.14: In experiment one, subjects were able to vary the size of their grip aperture. In experiment two subjects were able to vary the orientation of their hand. Reproduced with permission from Tunik et al. (2005).

Subjects were made to grasp a long block with a pincer grasp, the block was oriented either vertically or horizontally along its long axis. In the first experiment (See 2.14, experiment one) subjects had to keep their thumb and index finger aligned along the vertical axis of the block. In the second experiment subjects always had to grip the block on the narrow side of the block (See figure 2.14, experiment two). During the subject's reach-to-grasp action the block could be perturbed by $90\,^{\circ}$. Subjects would then have to modify their grasp by aperture size in experiment one and orientation of their grasp in experiment two.

TMS administered to aIPS during a reach-to-grasp action when the target was perturbed caused a deficit in adjustment of aperture size in experiment one and a deficit in orientation adjustment in experiment two. During experimental phases when the target was not perturbed TMS of the aIPS did not affect the reach-to-grasp action at all. This suggests that what is being represented and disrupted in aIPS is a representation of the intended goal of a planned action which is updated in real time. This is consistent with the experimental findings of Hamilton and Grafton.

The results of Hamilton and Grafton and Tunik et al. suggest that aIPS may encode the goals of others (as shown in Hamilton and Grafton) but may also encode the subjects own goals as shown by the disruption of action adaption in Tunik et al.

The results of Iriki et al. described earlier in this section show that aIPS may be where monkeys store a dynamic representation of their effector. If we consider this we could argue that the reduced reaction times found by Tunik et al. are caused by difficulties updating the internal model if it is used to dynamically drive actions. In Hamilton and Grafton's experiment the different goals were chosen to require the same type of grasp to pick up the object. Repetition suppression did not always occur in aIPS even though the grasp action was always the same. If aIPS was representing an internal model then we would expect repetition suppression to occur if the grasp type does not change. Hamilton and Grafton were also studying the observation of another persons actions which makes it unlikely that aIPS only represents an internal model of the observer.

If aIPS does encode goals it seems more plausible that they are intended effects rather than prepared sequences of actions. The results of Tunik et al. make more sense if we consider that reactions to a changing internal representation of goal state are being disoriented by TMS than a sequence of planned actions that need to be updated. Tunik et al. suggest that aIPS may iteratively make comparisons between the current motor command and sensory input and use that to adjust future motor commands. For this to work aIPS must have access to a representation of the final intended effect or it would not be able to adjust motor commands correctly.

### 2.2.3 Summary of neuroscience literature

In this section I have reviewed evidence from neurophysiological studies that illustrate neural representations of the effects (or goals) of action. The studies show no clear evidence of a single region of the brain which encodes action effects; instead they show several areas (medial prefrontal cortex, the anterior intraparietal sulcus, and the premotor area F5) which may represent the different information about intended effects that is required for distinct cognitive tasks.

**medial PFC** Two studies that I have reviewed suggested that action-effects are represented in the medial prefrontal cortex (mPFC). The study by Matsumoto et al. recorded mPFC neurons in monkeys and determined that some of these neurons encoded anticipated reward conditions in a trained task. These reward conditions are the final effect of the action that the monkeys perform, thus the results of the study show that action-effects may be encoded in mPFC.

The other study, by Chaminade et al., involved PET scanning of humans during the observation and imitation of a simple action. In some cases the subjects were shown most of the action but not the final goal state. They then had to imitate the action and try to infer the final goal state without having observed it. The mPFC of each subject was strongly activated under these conditions which the experimenters attribute to the inference of the final goal state of the observed action.

The results of these two studies nicely agree that the mPFC has something to do with the expectation of the effects of an action. Matsumoto et al. find that the mPFC seems to represent the expected effects of the monkeys own actions and Chaminade et al. find that the mPFC seems to represent the expected effects of another observed agents actions.

**aIPS** Three studies that I have reviewed in this section suggest that the anterior intraparietal sulcus (aIPS) encodes a effect representation of actions. Iriki et al. recorded neurons from the aIPS of monkeys whilst they performed reach actions with and without the use of a tool. Some of the neurons in aIPS were found to discharge when an object was within reaching distance of the monkey. When grasping a rake, which allows a longer reaching action, the neurons would discharge when the rake could be used to reach the object, even if it was outside the arm-length of the monkey. These neurons may then represent the area that an agent can influence with its effectors, rather than nearby objects it can directly reach.

Hamilton and Grafton scanned human subjects using fMRI in order to determine neural areas that represent goals. Subjects watched reach-to-grasp actions which were varied by trajectory and goal. The experimenters looked for areas that responded to repeated presentation of the goal with repetition suppression. The aIPS was found to undergo repetition suppression when the goal of an action was repeated but not when the trajectory of action is repeated suggesting that aIPS encodes the goal state of action.

The study by Tunik et al. used TMS on the aIPS of subjects during a reach-to-grasp action in order to disrupt activity. This caused a deficit in the subject's ability to reshape their hand in response to an object that changed as the reach action took place. This suggests either a model of the goal or the agent's effector is represented in the aIPS.

These three studies suggest that the aIPS encodes some sort of dynamic model of the object that is the goal position for a grasp action. aIPS is less likely to represent an agents hand as the results of Iriki et al. show that neurons in aIPS discharge when an object is within reach of the monkey. This suggests a representation of an object rather than the hand.

**F5** Two studies that I have examined suggest premotor area F5 encodes another representation of actions. The first, by Umiltà et al. (2008), recorded neurons in the F5 area of monkeys while they performed grasp action with two different pairs of pliers. The two pairs of pliers required two different hand actions to perform the same grasp actions with the pliers. The neurons discharged with the action the monkeys were trying to perform with the pliers (e.g. closing the pliers) rather than the hand motion that was required. These results suggest that neurons in the F5 region encode the effects of actions rather than the hand movements required.

The second study, by Umiltà et al. (2001), showed that some mirror neurons in F5 also discharged when the final grasp state of a motor action was hidden. This only occurs if the monkey had observed the object that was hidden before being grasped. These neurons seem to encode the inferred grasp type of the motor action.

These two studies suggest that some F5 neurons discharge when a particular part of a grasp action is completed or a successful grasp is achieved on an object. This could be a representation of the sub-goals that are needed in order to reach a final goal state, i.e. some sort of action plan.

The evidence presented here seems to show goals being encoded in various regions of the brain. This may reflect the different types of information about goals that are required for distinct cognitive tasks. A recent fMRI study by Majdandžić et al. (2007) found some evidence in humans that different levels of goals are encoded in separate regions of the brain.

## 2.3 Linguistic Models of Causative Actions

In this section I will give evidence for a class of actions defined by their effects from one more source: linguistics. The evidence comes from the analysis of the way that these actions are represented in human language. I will only discuss the basic evidence of this, the linguistic ideas that I present here will be developed further in chapter 6.

The linguistics evidence that actions are represented by their effects comes from a syntactic phenomenon called the **causative alternation**. Alternations are alternative syntactic ways of expressing a sentence which are roughly semantically equal. For instance, an alternation of *I gave a book to the teacher* is *I gave the teacher a book*. Different verbs undergo different alternations, and they can be categorised semantically by the types of alternation they undergo (Levin, 1993).

The causative alternation is a property of some transitive verbs that allow them to be rephrased as intransitive verbs. These verbs are known as causative alternation verbs. Examples (1-a) and (1-b) show the causative alternation pair of *open*. Notice that in example (1-a) *door* appears as the patient and in example (1-b) it appears as the agent.

This is unusual as we do not want this word to have two separate meanings. Linguists often propose that *John opened the door* really means *John caused [that] the door opened*. This form has the same structure as its causative alternation partner.

(1) a. John opened the door
    b. The door opened

This example shows a piece of evidence from linguistics of an action which is defined by its effects. In this case the action is defined by the effect of causing the door to open. I will examine this phenomena more thoroughly in section 6.1.

## 2.4 Summary of literature

In this chapter I examined studies that show that some actions are defined by the effects that they bring about. The behavioral studies that I examined in the first section show that the reduced reaction times in various tasks were based on the location of the goals of action rather than the motor actions that are required.

I then examined neuroscience studies that suggest that action-effects are encoded throughout the brain in areas such as premotor area F5, the anterior intraparietal sulcus and the medial prefrontal cortex. This is by no means an exhaustive summary of all neural areas that have been proposed to encode action-effects, this shows the major importance they seem to play in primate motor action.

Linguistic theory also provides some evidence for action-effect representation in the form of VP-shells and the causitive alternation. A modern analysis of language has displayed a clear distinction between action-effects and their causes, even in language!

The first three sections provided evidence which motivates the fact that some actions are defined by their effects and the importance that they seem to play in human action.

By now the reader should understand the importance of action-effects in human action; I will now pose the question, what are the mechanisms that allow us to learn these effect-based representations of actions? In chapter 4 I will explore this idea and leverage some of the ideas from experiments and models that I have reviewed in this chapter.

# Chapter 3

# A Review of Simulations of Reaching and Grasping Actions

## 3.1 Introduction

This chapter reviews some other models of motor learning from the literature. I am unaware of any other (biologically plausible) models that have the same goal as this project but many of these simulations use similar underlying mechanisms to learn motor actions and perform movements.

In this chapter, I will first explain some of the common techniques used in models of human action. Next, in section 3.3, I will explore some of the simulations from the literature. Finally, in section 3.4, I will examine the model that I have used to construct our simulation and the shortcomings that it has.

## 3.2 Key Concepts in Computational Motor Control

### 3.2.1 Rigid Body Dynamics

Simulations of human motion make heavy use of robotic techniques. In order to discuss some models of human motion from the literature I will first explain some of the more common methods. The robotic models of humans that I am interested in (and which are used in my simulations) are made from **rigid bodies**.

A rigid body is an idealisation of a solid object which does not any undergo any deformation (stretching, compression or shear deformation) from forces. These simplifications are perfectly reasonable in the situations that these simulations operate under.

A model is made up of a collection of rigid bodies which are connected together by **joints**. These joints allow the different rigid bodies in the model to move relative to each other along different axes. These joints give the total **degrees of freedom** (DOF) of the model which describe how it can move itself into a position. On the end of this chain of rigid bodies is the **end effector**. This is the part of the model that is designed to interact with the environment the robot is in. In the case of these simulations, the end effector is typically a simple human hand[1].

### 3.2.2 Motor State

The model of the arm that I have used is parameterised by the angles that the joints can be moved to. The set of the joint angles at the current time will be referred to as the **current motor state**. The goal position of a movement can also be parameterised by the set of joint angles required to achieve it and will be referred to as the **goal motor state**. Unlike the current motor state, there are many goal motor states for a goal hand position; all of these are valid final states of the arm that will bring about this position.

In order to move the model of the arm we will make use of a traditional technique used in engineering: kinematics. Kinematics is the study of motion without the consideration of the forces which cause that motion. In robotic simulations, the two most common kinematic techniques used for motion control are **forward** and **inverse** kinematics. These algorithms work in **configuration space** (rather than in Cartesian space); this is the multidimensional space which defines the values of each of the joints in a model[2].

#### Inverse Kinematics

Inverse kinematics involves the computation of a set of joint angles that will achieve a given end effector position and orientation. This is far more useful than forward kinematics as the desired position is commonly known and the joint angles that are required to reach that position need to be solved. It is easy to see that, for a human arm, there are many combinations of joint angles that would place the end effector at a goal position. This means that there needs to be other conditions that must be

[1]Since the mechanics of the human arm are quite complex the models that are presented here use fewer degrees of freedom than that of a real human arm.

[2]This is the same as the motor state that I have defined above.

fulfilled to choose the set of joint angles for the arm. For simple situations this is the set of angles which move the end effector through a straight trajectory.

### 3.2.3 Controllers

A controller is used to control and alter the dynamic behaviour of the system as it moves towards a given goal configuration. Controllers fall into two broad categories: feedback and feedforward controllers. Only feedback controllers have been used in this project and the papers presented here so feedforward controllers will be ignored.

In general a controller alters the accelerations of the components of a model as it moves towards a goal location. A feedback controller takes the current and goal motor states of the system and computes a force which will move the system towards the goal state.

A simple feedback controller can be used to move the arm towards a goal at a fixed velocity attempting to minimise the distance between the goal and the current position of the end effector. The error calculated by the controller will be the difference between the expected position of the end effector and its actual position at that time. Because the controller feedback will be slightly delayed simple controllers like this tend to overshoot the target and then oscillate around it before reaching a stable configuration. This motion is unnatural, so more complex controllers are used to produce more accurate behaviour.

#### PID Controllers

A common feedback controller is the PID controller (proportional-derivative-integral controller). This controller has three components to do this, a proportional, integral, and derivative component. These are summed at each time step and used to adjust the velocity of the movements.

The proportional component is proportional to the current error term. A controller which only makes use of the proportional component (i.e. the integral and derivative components are zero) will act like the simple controller discussed above. Because this may never reach the target another component (the integral) is added to the controller.

The integral component gives a sum of the past errors that should have been corrected. This accelerates the movement towards the goal as time passes, but may cause the controller to overshoot the goal.

The derivative component calculates the rate of change of the error as a prediction of the future error. This slows the rate of change of the controller and attempts to counteract some of the overshoot that may be caused by the integral component.

The sum of these three terms is used to adjust the dynamics of the motion. All of the individual terms are often multiplied by an individual constant to increase or decrease the relative importance of each term in the result. This controller may not suffer from the oscillations which plague more simple controllers, however the controller is sensitive to the constant terms which are multiplied with the PID values. These have to be optimised for each problem the controller is applied to in order to get good results.

## 3.3 Other simulations of infant motor action

So far in this chapter I have discussed the purely mathematical solutions to motion. In this section I shall discuss how these techniques have been leveraged to create biologically plausible models of human motion and motor learning. The simulations I discuss here have particular relevance to this project, more examples of models of motor learning and control can be found in the literature.

A notable early neural model of motor learning is Kawato et al. (1987). The model is a hierarchical neural network model that is used to learn and produce motor commands. The model is broken into parts as motivated by the literature and represents internal models of the dynamics of the arm.

The model is goal-driven; after selecting an object it chooses a trajectory which brings the hand to the chosen object, and then creates a sequence of motor commands which moves the hand along the selected trajectory. This trajectory is a sequence of points for the hand to follow through space and is produced by the algorithm from Uno et al. (1989). What is learned is the feedforward component of the motor controller that generates these commands. This model of planning is typical of many models of motor function and it used to good effect by Uno et al. to produce skilled motor action. However, more recent evidence has shown that, at the very least, motor planning and control is a far more iterative process which relies on feedback. Evidence discussed by Cisek (2005) suggests that complete trajectories are not created prior to action, instead they evolve as the action is performed. Models of the type produced by Uno et al. and Kawato et al., while highly successful, do not capture this aspect of motor learning.

Many recent computational models of motor learning have involved learning through imitation. Billard (2001) created a biologically inspired robotic model which learnt motor actions by imitating another agent in a simulated environment. This is of particular interest to this project as the model learns and is tested on its ability to perform reach and grasp actions.

Billard models several stages of motor action: action recognition, motor control, and motor learning. These are also broken into smaller modules which Billard associates with areas of the brain that he is modelling. Recognition involves a simple conversion of cartesian coordinates to a reference frame relative to the observer. Motor control and learning are broken into stages representing the spinal cord, brain stem, premotor cortex, primary motor cortex, and supplementary motor area.

In Billard's model, learning involves the imitator attempting to replicate the imitatee's neural activity. The model is able to learn a reach and grasp movement using this principle to a high degree of accuracy. The emphasis on imitative learning to learn such basic actions as reach and grasp in this paper unusual. The emergence of reach and grasp actions in humans occur very early in development and are assumed to be mostly learnt from a mixture of feedback of the infants own actions. It is this type of early motor learning that I will focus on.

The Infant Learning to Grasp Model (ILGM) (Oztop et al., 2004) is a model of motor learning which is guided by neuroscience and child development literature. ILGM is mostly a model of how infants learn a trajectory which causes a hand to grasp a target object and differs from many previous models of grasp learning because of its open-loop action execution. ILGM requires only two things to learn actions: the agent is able to sense the effects of it's actions, and the agent is able to use feedback to adjust its movements.

ILGM makes the assumption that reaching, which always appears before grasping in infant motor development, somehow subsumes grasping in real motor learning. The basic reach action is assumed to have been learned by the model and implemented using traditional engineering techniques. The model learns how to grasp in steadily more complex situations from this baseline.

Oztop et al. developed a model which learnt to grasp from goal-directed reaches. In the first instance they learnt to grasp an object while the orientation of the wrist is automated by the system. After learning this, the experiment was performed again, but the orientation of the wrist was learnt as well as the original parameters.

In these simulations, ILGM generates parameters for its joints to reach a goal object (with a small chance of generating a random plan for exploration); from this a motor plan is generated and performed. The model is trained using a reinforcement scheme termed **the joy of grasping**: in this scheme rewards are produced by touch sensations in the fingers and hand, and punishments are produced in the absence of such sensations. As input, the model takes a goal motor state, and after training the model was able to generalise to nine different object locations with 85% success rate.

The ILGM provides computational support for the hypothesis that a simple action (in this case reaching) can lead to a more complex action (grasping). This is done without a preprogrammed concept of minimising the error between the hand and object shape or a function matching the hand shape to an object. The idea of graded reward (the joy of grasping) for a successful or partially successful action will be leveraged to learn causative actions in my model. The separation of the learning of reach and grasp actions will also be used in the model I present; the learning of grasp actions will also feed back into refining reaching actions.

These models make good attempts to explain many aspects of the acquisition of motor skills and some of the behavioral markers that we see as infants acquire them. However they are missing one particular aspect that requires us to produce our own model which is discussed in the following section. Too much of the trajectory of the motion is precomputed by the model.

## 3.4 Grasp Project Model

The models that I have discussed so far precompute trajectories in order to reach a goal-arm-state. Traditionally, theorists have assumed that control of an action can be broken into two separate pieces: planning the trajectory, and computing the motor commands required to produce this trajectory. Obviously this scheme would require that these operations are performed in sequence. However, research into biological motion control has shown that this highly detailed trajectory is not computed prior to the initiation of action (Cisek, 2005). A general trajectory is used at the beginning of motion (which appears to a vector in the general direction of the object) which is then refined using reafferent sensory feedback until the agent achieves the intended goal state.

In order to produce a model that uses default trajectories and reafferent feedback to produce a trajectory I expanded an existing model of reach-to-grasp actions designed and built by Tim Neumegen (Neumegen, In progress). This model, called Grasp, is a simulated human arm built from simple rigid bodies using the JMonkey[3] physics engine.

[3]www.jmonkeyengine.com

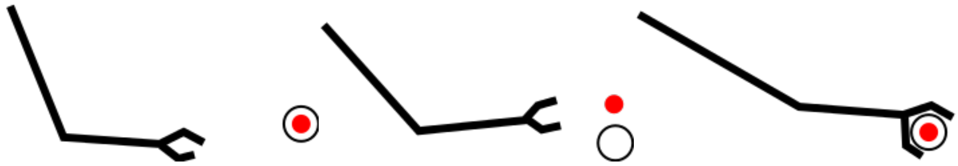
There are two interesting aspects of this model. Firstly the trajectories of movements are not precomputed, instead these are computed as the motion unfolds. Secondly the trajectories of complex movements are built on top of those of simple movements.

Early in development, the model learns simple actions that achieve touches, using Oztop's idea of the joy of grasping. Later in development, the control signals which generate simple reach-to-grasp actions are **perturbed**, to produce more complex reach-to-grasp actions. It is important to note that this term is used slightly differently than in the motor control literature. A perturbation is normally a change in some aspect of an experimental subject's motor situation: for instance, a change in the location of a target object when a reach action is under way Jeannerod (1996) or an experimentally-induced change in the subject's motor plant(Wolpert et al., 1998). In Neumegen (In progress) and in this work a perturbation is a temporary alteration of the goal-motor-state of the end-effector which is removed under some condition such that the end-effector follows a particular trajectory through space.

### 3.4.1 The use of perturbations in a model of motor control

In Neumegen's model a perturbation is an alteration of the **goal motor state**, i.e the set of joint angles, that must be achieved in order to bring the hand into contact with the target objects. The reader might also imagine that a perturbation is also a function of the particular action that the agent is attempting to perform. For example, a relatively small perturbation is necessary to grasp an object, but a large perturbation may be required to squash an object with a lot of force.

An example of a perturbation is shown in figure 3.1. In this simplified picture the object is shown by the black circle and the goal motor state of the hand is shown as a red ball. The red ball actually shows the final position of the hand wheras the real goal motor state is in the joint coordinate system of the arm but the idea is similar. This simple image assumes that we want to grasp the black circle from above. Figure 3.1a shows the arm beginning to reach for the target. At this point the arm is simply moving towards the centre of the target object. Figure 3.1b shows some midpoint of the reach-to-grasp action. At this point the goal motor state (in joint coordinates) is temporarily perturbed to shift above the object, this means that the trajectory of the arm changes. In figure 3.1c the perturbation of the goal motor state has been removed and the hand has shifted back towards its target and achieved a grasp of the object. This example leads to the question of when perturbations of the goal motor state should be applied, and when they should be suppressed. This is dealt with in section 4.3.2



(a) Beginning Reach (b) Mid-Reach (c) Grasp Achieved

Figure 3.1: A perturbed goal motor state

### 3.4.2 Considerations for a model of causative-actions

At the level of implementation, the mechanism that is used to learn motor control in the Grasp Project is the recording of successful reach and grasp actions during training. These are stored with a reference to the 'score' of the action, this is analogous to the intrinsic joy of grasping that is discussed in Oztop et al. (2004). When learning or performing an action, high scoring past trajectories are taken and modified in an attempt to produce a new higher scoring trajectory (they can also be randomly made more or less complex at this point) which is added to memory if it is good. Low scoring trajectories are discarded as memory becomes full, this reinforcement learning allows the model to learn to grasp objects.

Although this model does not require trajectories to be computed prior to motor action, the mechanisms that are used in Neumegen's system are not biologically plausible. Therefore, one of the goals of my project is to produce the dynamic behaviour of the Grasp model using a neural network model. An intriguing concept introduced by the Grasp project is the idea of a perturbation of the goal motor state. In chapter 2 I discussed evidence for an effect-based representation of motor actions. I want to produce a model able to learn these causative-actions using a mechanism which does precompute the entire trajectory prior to movement; this mechanism will be the perturbations discussed above. The question for this project is: can we use these perturbations to produce the complex actions that we are interested in?

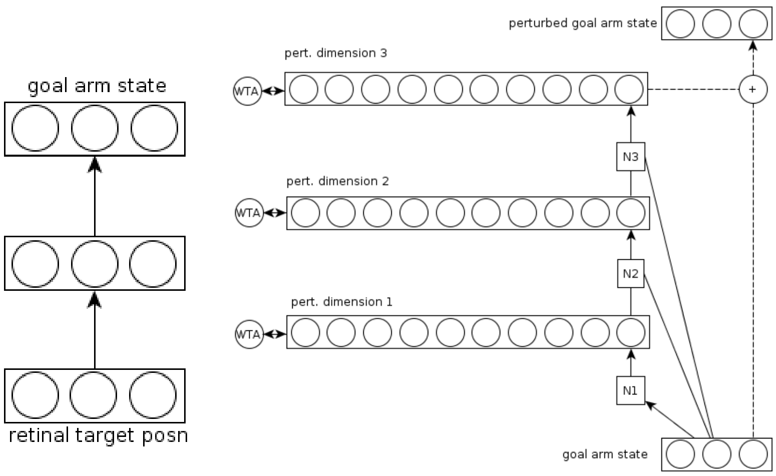
# Chapter 4

# Modeling simple reach-to-grasp actions

In this chapter I will present a computational model of reach-to-grasp actions. This is a novel model in its own right since it avoids precomputing detailed trajectories prior to action. This shows that the idea of perturbations of the goal-motor-state is viable to learn simple actions, however the main point of this initial model is as a basis for the model of causative actions which I present in chapter 5.

## 4.1 Introduction

In order to build a computational model of a human arm that is able to perform the class of actions that were described in chapter 2, I first considered what underlying mechanisms the model must have. The model must be able to perform reaching motions towards a target and manipulate its digits in order to perform an action. In the model that I have produced (shown in figure 4.1) this process and the learning that is involved has been broken into several stages.




(a) Reach Network (b) Perturbation Network

Figure 4.1: Completed computational model. a) shows the reach network; this network converts an object in retinal-coordinates to a vector of joint-coordinates which will bring the end-effector into contact with the object. This output is given as input to the perturbation network which calculates a perturbation of the goal motor state which will cause the model to grasp the object. Each of the components of the perturbation vector must be calculated individually and are interdependent so are calculated in series here. Since there are many possible perturbations which could lead to a successful grasp a discrete distribution of possible perturbations is calculated in which a winner-take-all (WTA) method is then used to select the perturbation with the highest activation.

These stages are partially motivated by existing computational models of action control (see chapter 3), and partially motivated from child development literature. Gordon (1994) provides a review of the ages that certain components of simple action manifest in infant development. In early stages of infant development the hand of the infant is always open before a reach action and the fingers do not move as the hand moves towards its target. Any attempts by the infant to pre-shape their hand for the object are crude. By four to five months, the infant can produce a more accurate reach trajectory towards objects. The infant is learning a function mapping visual information about a target object onto a motor movement. At about nine to ten months the infant is able to orient and pre-shape their hand to an object so the visuomotor function that is learned is somewhat more complex. At this point an infant's ability to grasp an object is still largely based on tactile feedback and stimulation of the palm will cause the infant to close their hand.

This model attempts to capture some of this structure in the sequence that it learns parts of the simple reach to grasp action. This is broken up into a reach/orientation-learning and grasp-learning stage. These will be referred to as stage one and stage two respectively.

In stage one this model learns to map the retinal location of a target object on the retina onto a set of joint angles of its arm that will bring its fingers into contact with a target object. I will call this set of joint angles a **goal motor state**. It does this using a type of learning known as **motor babbling**. After the model has achieved a certain level of competence it begins to learn how to orient its palm in relation to the object it is trying to touch. The second stage of this development involves the model learning to grasp the objects in its surrounding space.

### 4.1.1 Learning

Our model uses a neural network to learn the different parts of reach to grasp actions. The model trains the neural networks it uses in ways that differ from traditional training methods. While neural networks are usually trained with a complete set of training data, our model generates its own data through a sequence of **exploratory movements**. In each exploratory movement, the arm begins in a fixed position and a target object is presented producing retinal information. This information provides input to a neural network to be trained; the network generates a prediction of the goal motor state. To allow exploration of the full range of possible motor actions, noise is added to the output, and the resulting motor action is performed. The amount of noise is reduced during training, as in a standard reinforment learning paradigm (Sutton and Barto, 1998).

This model gets data by attempting to interact, in a simulated environment, with objects in front of it. In early stages of the model's 'development' a simple touch event on an object will generate data and reset the simulations state. Before the next iteration of the simulation begins the neural network will be trained on the data that it has gathered so far.

This model also operates a **sliding window** over the training data. This means that the model only learns from the most recent $N$ items of training data. This means that as the model's performance improves during learning the older, generally lower quality, data is discarded eliminating it from the training set and gradually reducing its effect on generating actions over time. The quality or score of the data will be discussed in section 4.2.3.

### 4.1.2 A progression of rewards for stages one and two

Our model also incrementally learns from only better data as it progresses through stage one and two. This can be motivated by the well known phenomena in which dopamine neuron activation predicts the expected reward signals for an action when presented with a cue (Schultz, 2002). For example, when a new reward is presented dopamine neurons are activated. As this reward is repeatedly observed the reward begins to move backwards through time to the earliest event that predicts that the rewarding event is going to occur. The activation of the neurons is also dependent on the strength of the reward signal. Positive activation of dopamine neurons occurs when an unpredicted reward occurs or when a predicted reward is better than expected. Negative activation occurs if a predicted reward does not occur or the actual reward is worse than expected.

Learning can occur using this mechanism and is known as the **blocking-effect** (Kamin, 1969). This states that a stimulus that is fully associated with a fully predicted reinforcer cannot be learned. As our model learns to reach and grasp for objects we assume that the predicted reward (touching or grasping the object) would increase so we begin to learn only data that produces a higher reward. Note that this only captures the positive activation of dopamine neurons that occurs when a reward is better than predicted. It does not capture the negative activation that occurs when a reward is worse than predicted.

### 4.1.3 Simplifications of this model

Before continuing to the structure of this model I will first address two simplifications I have made. First, this model restricts the degrees of freedom of the model of the arm that I use. Second, the model uses traditional neural network models for neural circuitry rather than more modern methods (such as spiking network models).

The restriction of the degrees of freedom in learning largely came about to decrease the time of training in this model. Simulations that I conducted occurred in real time (partially due to problems with the physics). This meant that I was forced to simplify in order to decrease the time simulations would run to a realistic time. This restriction should not detract from the results, the model is able to perform the actions that I require. More active degrees of freedom will simply increase the size of input data and the required number of neurons used to learn. This simple model may not scale to more degrees of freedom as the search space for 'correct' actions grows large. Real infant motor learning probably involves a mixture of haptic and visual feedback. In order to learn motor actions with more degrees of freedom this model might require the implementation of a simple visual feedback system to reduce the search space. With the addition of forms of feedback to aid learning there is no reason to think that this model would not scale to increased degrees of freedom.

The second choice I made when modeling is to use a traditional multilayer perceptron (MLP) as opposed to a more modern spiking neural network. This is in part due to the level of circuitry that I wish to model; I am only interested in coarse grained population codes of the motor circuit rather than the representation involved at a per neuron level. A second reason for this choice is the relative complexity of the interconnected parts of the model. Instead of using many computationally intensive spiking networks using MLPs allows for faster training without adversely affecting the results of learning.

## 4.2 Reach Network: Structure, Training, and Results

### 4.2.1 Introduction

The first stage of this computational model allows the simulated model to learn how to reach and touch an object based on a simple representation of its retinal location (we assume a fixed retina - i.e. one which cannot change in position relative to the agents arm). Only one object is presented at a time, and there are no obstructions in front of it. This is a relatively simple problem which involves learning mappings between the visual coordinates of an object and a set of joint coordinates that would allow the arm to reach that point. Note that the reach network only calculates the goal motor state for a visual location, it does not control the motion that is performed to reach that goal state. The motion is controlled by a PID controller, as discussed in section 3.2. When a goal motor state is calculated by any of the networks in this model the PID controller will calculate a force to move the arm towards that state. This is done continuously until a state producing a reward occurs (in this case a touch event).

### 4.2.2 Structure of neural network model

Figure 4.2 shows the structure of this stage of the model. The three input units encode the $x$, $y$, and $z$ positions of the object in a retina-centered coordinate system and the three output units encode three joint angles of the arm: two for the shoulder ball joint and one for the elbow. The $x$ and $y$ (horizontal and vertical) inputs are read from a simple visual routine that computes the centroid of the object on the retina. The $z$ coordinate (distance of the object from the eye) is read directly from the physics engine; this routine is a placeholder for a more complex depth-perception routine, e.g. stereopsis.

The input layer neurons accept values between -1 and 1 where 0 corresponds to the centre of the visual field. The hidden layer and output layer both use linear activation functions.

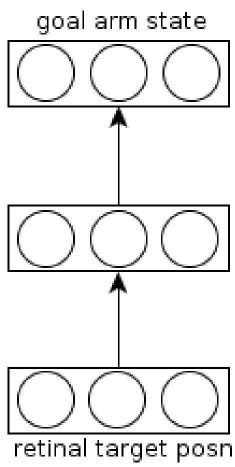



Figure 4.2: Structure of the reach neural network model.

### 4.2.3 Training the network

Training data is gathered by allowing the simulated human arm to motor babble (move randomly) within its environment until the arm makes contact with an object within its visual field. Objects can be generated anywhere within the receptive field of the model, but within reaching distance of the arm. The visual location of the object is used as input data to the neural network. The output data, which is a set of joint angles from the simulated arm, is collected from the model when a touch event occurs.

When a touch event occurs the visual coordinates and set of joint angles are logged to files and the model is reset to its initial conditions. A **touch score** is also logged with the set of goal joint angles (these make up the goal motor state). This score is a measure of how good the touch was, a large surface area of contact between the object and the fingers will generate a large score, and a small surface area of contact will generate a small score. This score is used to determine the learning rate of the neural network when learning each piece of training data. A piece of training data with a low touch score will have a low learning rate, and a piece of training data with a high touch score will have a higher learning rate. As discussed in section 4.1.2, as time passes data with a very low score will have it's learning rate tend to zero. Before the simulation begins again the neural network is trained, for a single epoch, on all previously gathered data. Training of the neural network is done using back propagation (Bryson and Ho, 1969).

As the model experiences more touch events it gains more data and begins to explore its environment in a less random manner as the random noise is reduced and the network begins to predict more accurate goal joint angles. The sliding window also throws away old data, which should improve the performance as new data should have a higher touch score.

### 4.2.4 Results

In order to test the performance of the reach network we need a way to produce a test set. Because all of our data is generated from random simulations the best way to do this is to train the network, and use the learned training data as a test set for an untrained network. Figure 4.4 shows the error of predicting an unseen test set as an untrained model experiences touch events. Before the model attempts to reach a target in each epoch it first tests itself on a set of unseen data that has been gathered from a prior run of the model. The average sum squared error is then logged to a file. As shown in the figure this quickly asymptotes to a value of around 1.5 and then doesn't improve. Obviously, the system's performance quickly improves, but what is the remaining average error? It is not unusual to expect the prediction error of a neural network to approach zero if enough training data is provided. This is unlikely to occur in our situation as the objects that the model is interacting with allow for a variety of acceptable goal motor states to bring about a touch event. For a given object, whose visual location is $x$ there may be several different joint configurations capable of reaching $x$ so $error > 0$ for training. Instead the network will probably predict a joint configuration which 'averages' all the acceptable training joint configurations.

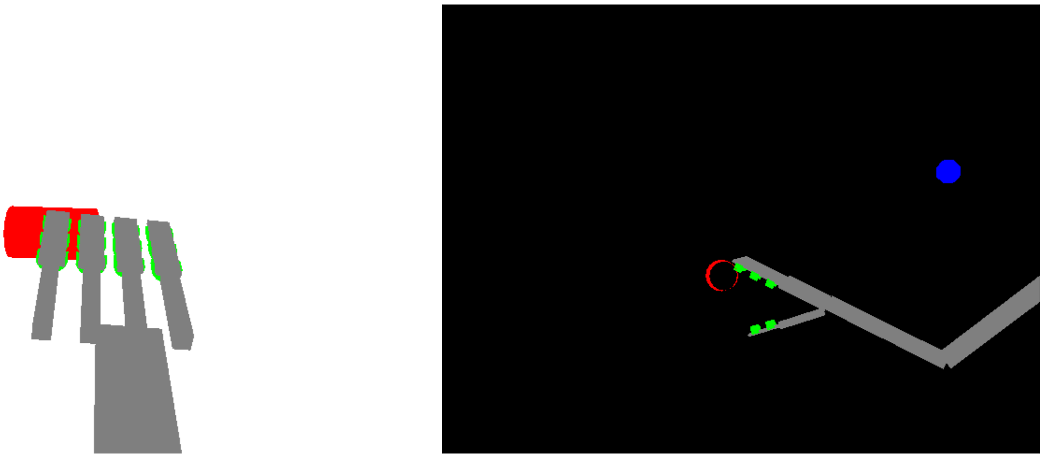


Figure 4.3: A typical touch event from two angles.

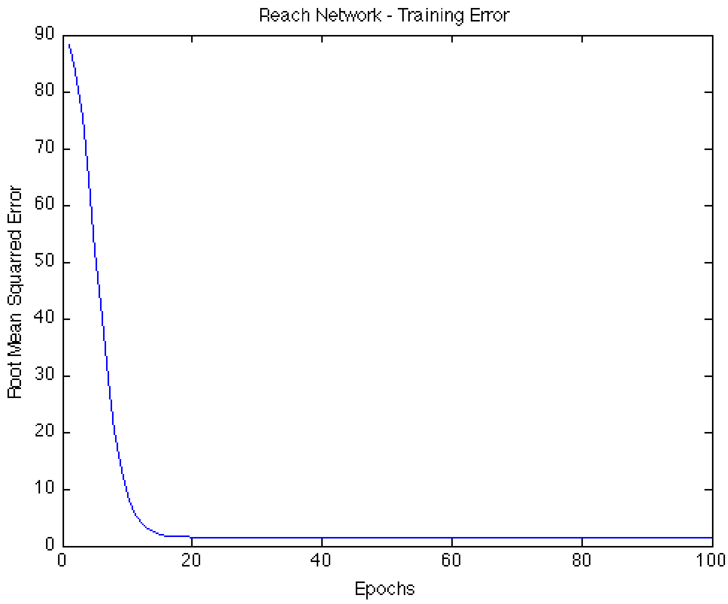



Figure 4.4: Error of testing the neural network.

Figure 4.5 shows the average error of the network that is produced when it is trained on the existing data after each touch event. Unusually the error spikes upwards during various epochs of training, this is because the training set is increasing as the training continues.

The last part of the testing of this stage of the model is testing its ability to actually reach and touch an object that it is presented with after a period of training. After being trained on a data set of 200 visual locations and joint angles the models ability to touch a presented visual object was tested 30 times. The results of this test are

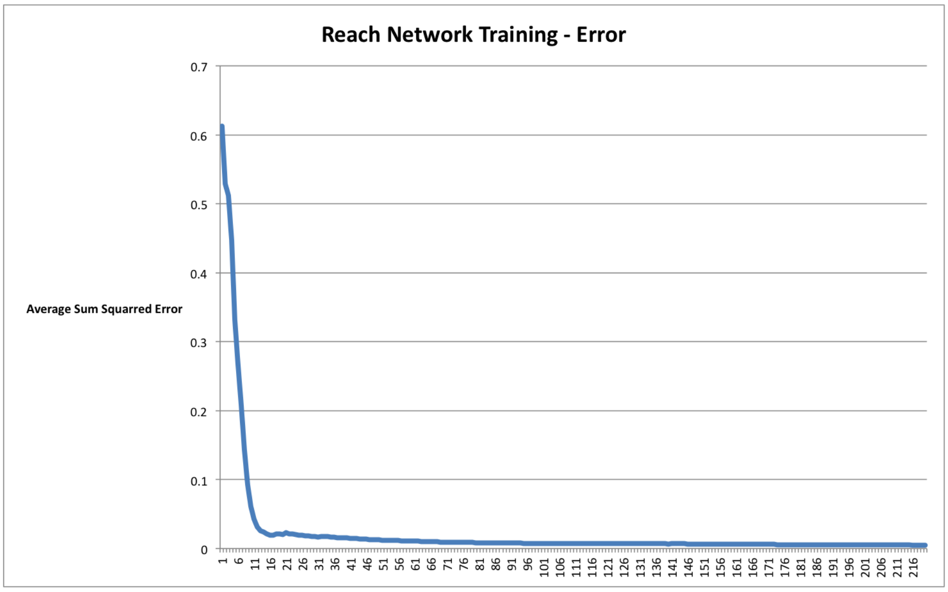



Figure 4.5: Error of the training of the neural network. Redo this in matlab for consistency.

shown in table 4.1. The results of this are good, the model is able to reach and touch an unseen target in 80% (24) trials. Although the model does not touch the object in the remaining 20% (6) trials, it does get very close to the object in most of these unsuccessful trials as shown in the last entry of the table.

Table 4.1: Verification check.

| **Reach-to-touch testing (30 trials)** | |
|---|---|
| Successful touch | 80 % |
| Unsuccessful touch | 20 % |
| Average distance from target in unsuccessful trial | 3.55 cm , STDV 1.3 cm |

Figure 4.6 shows the successful reach-to-touch locations that are produced by this model.

The model's real performance is not as good as the error rates suggest, this can probably be attributed to a single factor: some visual locations will be greatly under-represented in the training data. This is due to the fact that some visual locations are mechanically harder for the model to reach. As the random noise the model uses to explore decreases over time it also means that unexplored areas are less likely to be explored in future trials. This means that some areas may not have any successful touch events. For example, if the model doesn't touch an object on the left and the

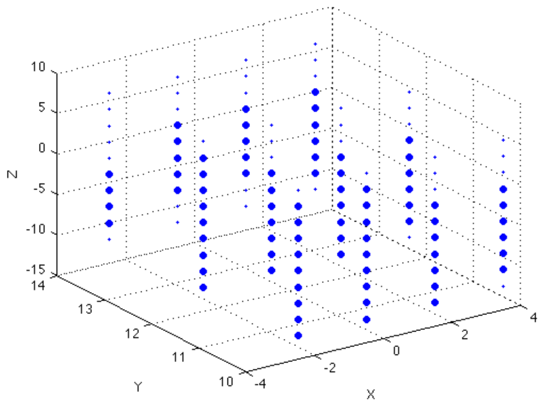



Figure 4.6: Successful Reach-to-touch locations after training. Successful locations are given by large points, unsuccessful locations by small points.

randomness is decreased, then the neural network is unlikely to predict a correct set of joint angles for any object on the left which it has no data to interpolate from.

In order to test this the trained model was tested by giving it objects one-by-one in a uniform grid spanning the entire space. The results of this are shown in figure 4.6. These results roughly validate this theory as the areas which the model is unable to reach are clustered in groups rather than distributed across the space. They are also consistent across different trained networks suggesting that these unreachable areas are difficult for the model to explore.

## 4.3 A network for learning reach-to-grasp actions: the Perturbation Network

### 4.3.1 Introduction

The model that has been proposed so far produces rigid behavior based on a few variables (the location of the object and some simple visual information such as the orientation of the object). In order to produce reach-to-grasp actions (and more complex actions) we have introduced a notion of perturbations of the goal hand state which was discussed in section 3.4.

In the context of this model a perturbation is a small change to the location that the model perceives the object to be in. This alters the goal motor state that the simulation will choose to move the arm to. Using a series of perturbations of the goal arm state allows the model to produce trajectories other than the straight movement that would normally be produced. This section describes how these perturbations are generated and some of the problems that are presented in using these to create more dynamic behavior.

### 4.3.2 Issues with the perturbation network

This model represents a perturbation as a three-dimensional delta which is applied to the goal hand state generated by the reach network. Representing them as changes to the goal state means that they can be suppressed when we want the arm to return to its original goal state (as discussed in section 3.4.1).

The question of when a perturbation should be applied and removed from an action has not yet been addressed. When learning perturbations that allow the arm model to grasp we found that we can simply remove the perturbation when the hand moves within a certain distance of the object. If the perturbation is be applied immediately to the goal motor state the will cause the end-effector to travel through a parabolic trajectory rather than a straight one.

Changing the trajectory of an action is likely to involve feedback from the visual system. Distance is the simplest thing that seems to allow the removal of perturbations to produce a trajectory that reaches the intended target. However, this has problems dealing with more complex actions in which case a more complex mechanism of removing and adding perturbations may be appropriate. Another problem with the way we use perturbations is that the distances that we use are not learned, they are static distances set as part of the model. These issues will be addressed in section 7.3 as work that can be done at a later date.

Another problem is with the range of the data that is produced to train the model. This meant that a traditional neural network architecture was unsuitable for learning perturbations, this will be discussed in section 4.3.3.

### 4.3.3 Structure of the network and training

The original plan for this neural network was to use an **SRN** (Simple Recurrent Network)(Elman, 1990). This network would allow the model to learn sets of perturbations which lead to particular actions being performed on an object. The structure of this model is shown in figure 4.7.

To gather training data the reach network (see 4.2) was given the location of the target object in visual coordinates and produced a goal motor state to reach that target. A perturbation of the goal motor state was then generated to cause the arm to move through a different trajectory to the target. This perturbation was removed when the arm came within a fixed distance of the target and was logged as training data if this resulted in the hand grasping the object.

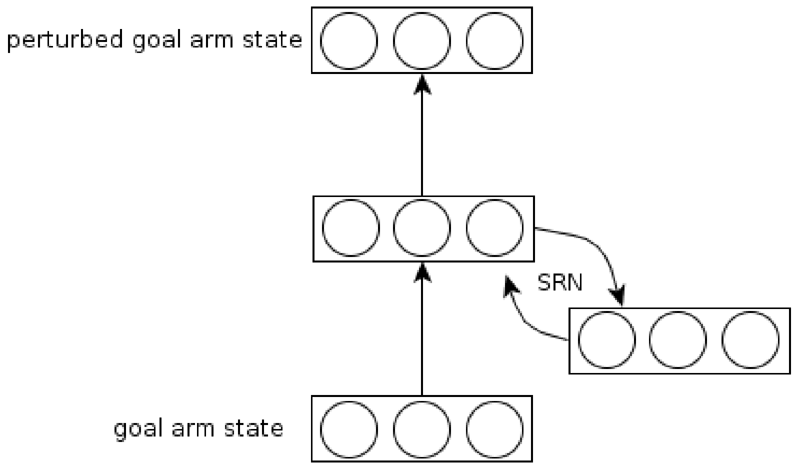



Figure 4.7: Original Structure of the Perturbation Network.

Problems arose when training this model as the successful perturbations that were produced for a particular goal motor state were very variable. This meant that the SRN learnt some sort of average of the total set of perturbations for a given goal motor state. This average goal motor state would not cause a perturbation that would allow the arm to grasp its target. In order to fix this the neural network structure was changed to 4.8.

In this new structure, each of the components of a perturbation are calculated in a serial structure with the resultant perturbations being fed into the input layer of the next network in order to calculate perturbation components that are dependent on each other. The output was also changed to a localist structure of 40 output neurons. Each of these neurons corresponds to a particular value in the range -1 to 1 with the same interval spacing between them. The neural network was then trained using one-hot encoding. This meant that the network produces a perturbation based on the perturbation, or similar perturbations, that it saw most often. This could also be adapted to use a softmax output activation to produce a probability distribution of the possible perturbations. These changes stop the model from averaging successful perturbations and instead causes it to produce the trajectories that it has

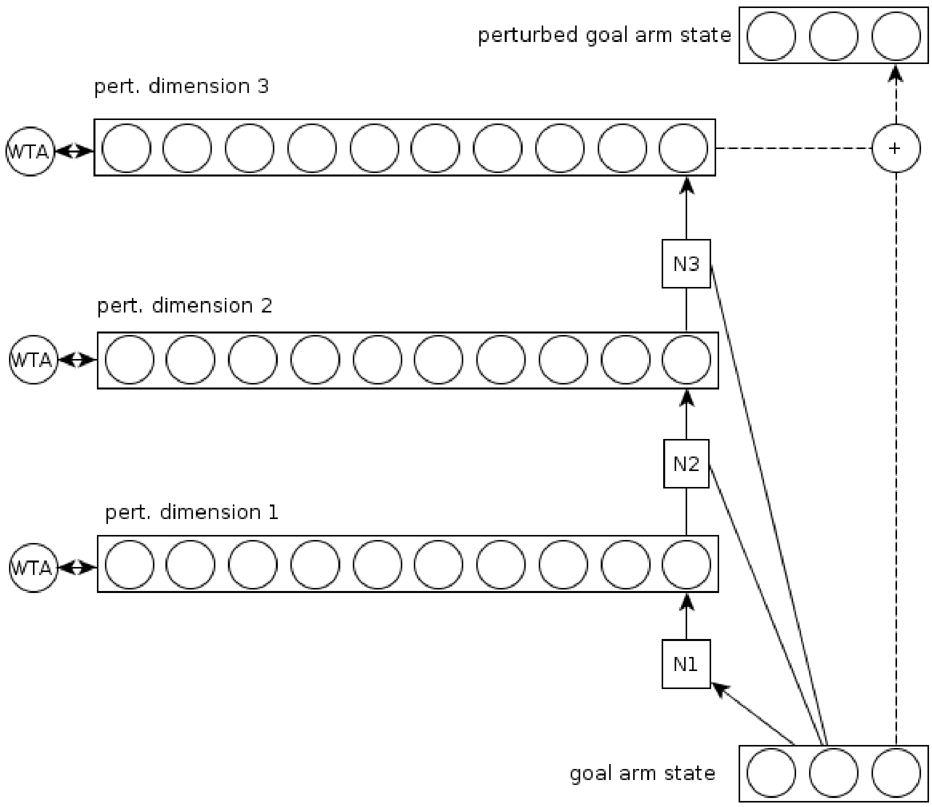



Figure 4.8: New Structure of the Perturbation Network

found successful most frequently.

This structure is more accurate at producing grasp actions than the original plan for this neural network, but it still had some problems. The issue with this neural network is that the localist output is too coarse to produce correct values for most perturbations. In order to fix this the output of the network was changed to a mixture of localist and distributed coding. The winner of the output neurons was taken, then its value was changed slightly based on the activation of the output neurons on either side of it. This sort of coding is similar, albeit simpler, to the population encoding of direction observed by Georgopoulos et al. (1986). An example of the output from my model is shown in figure 4.9. In this case the winning neuron would be the fourth from the left, but the value that is produced would be slightly biased by those in its local region (a few neurons from the left and right).

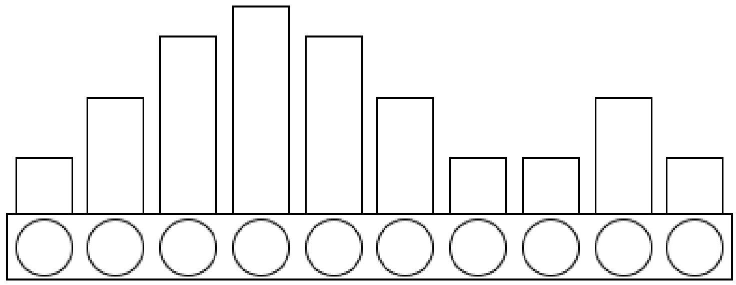


Figure 4.9: Population Coding for the perturbation network output

### 4.3.4 Training the Network

Data for the perturbation network was gathered in a similar fashion to the reach network. An object was placed at any reachable location in the world in front of the arm and then the model attempted to reach and grasp it. A predicted goal motor state was produced by a trained reach network. This location was then perturbed by some random amount.

This perturbation was suppressed according to a distance metric from the original goal motor state. This would then cause the arm to move back towards its original target. If the inside of the fingers touch the object then the hand attempts to grasp the object. At this point training data is logged in the form of the original set of the joint angles and the random perturbation that led to a grasp. As well as this a value is also depending on whether the grasp that occurred was stable or not. A stable grasp is defined as one where the object and hand are in the same relative motion.

After a successful reach-to-grasp event the model is trained on a sliding window of successful training data. As the number of simulations progress the learning rate for non-stable grasps approaches zero, this speeds up the training for the network and can be partially motivated by likening this to the reward prediction mechanism described in 4.1.1.

After training a validation of the model is performed to see if it can reach and grasp a target object while using a perturbation, the results are shown in table 4.2. In trial one the models performance poor, we decided that the issues with performance could be due to insufficient distribution of training data across the space or noise between the learned joint perturbations. Noise between the joints seemed more probable as two of the joints (the shoulder and the elbow) move through the same axis, therefore any combination of these two joints which could bring about a certain hand location should be valid training data. A function such as this should be very difficult for the neural network to learn. To test this, I plotted the goal-hand-state training data against the perturbation that had bought about this state. If the vectors had similar magnitude and direction for similar positions then the noise is not likely to be an issue (as the model is always producing similar joint perturbations for similar object locations). The graph also tells us what the distribution of training data is across the space, one such graph is shown in figure 4.10. This diagram shows similar vectors for similar locations in the search space, this means that it is unlikely that two perturbations in the same axis causes much noise in the learning. The diagram also shows that the distribution of training data is very localised, a more even distribution of training data across the space is likely to improve performance.

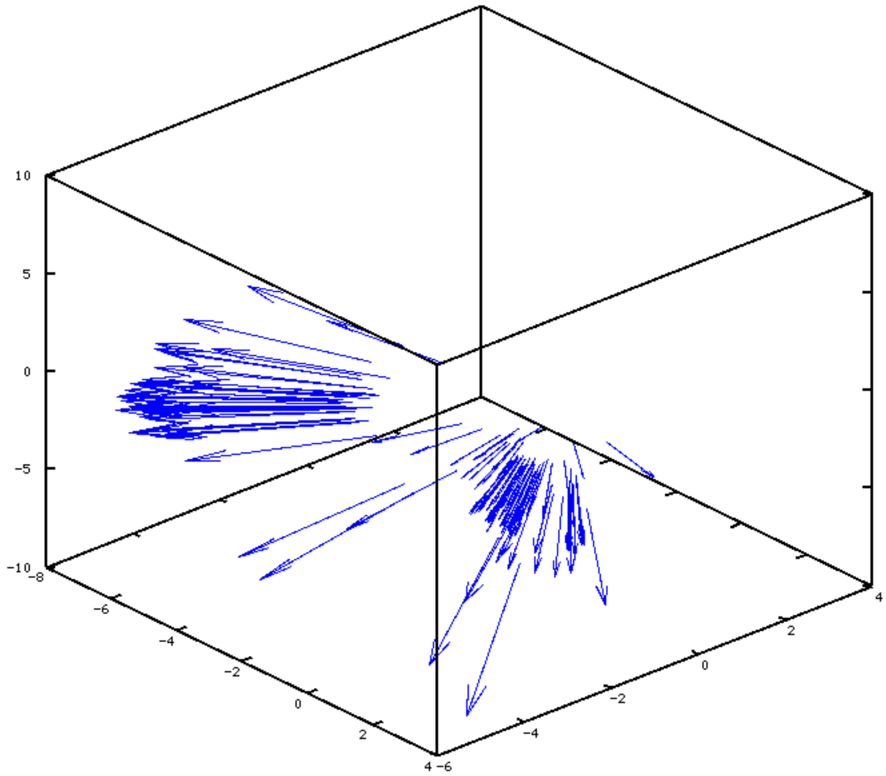



Figure 4.10: Vectors of training data in cartesian space.

In order to improve the models performance and test its ability to interpolate the next experiment was run in a less plausible manner. A grid of evenly distributed and spaced objects was created and then presented to the model at random. The model then attempted to reach and grasp the objects until it was successful or a certain number of trials elapsed (30 trials). The training data gathered from one such experiment can be seen graphed in figure 4.11. This graph shows a more even spread of data although there are still large gaps in parts of the space. This can be explained by analysing the performance of the model and observing it gathering training data. The model's performance is shown in table 4.2 as trial number 2. The performance of the model to reach unseen locations is much better than it was, however there is still a large number of object locations that the model is not able to successfully grasp.

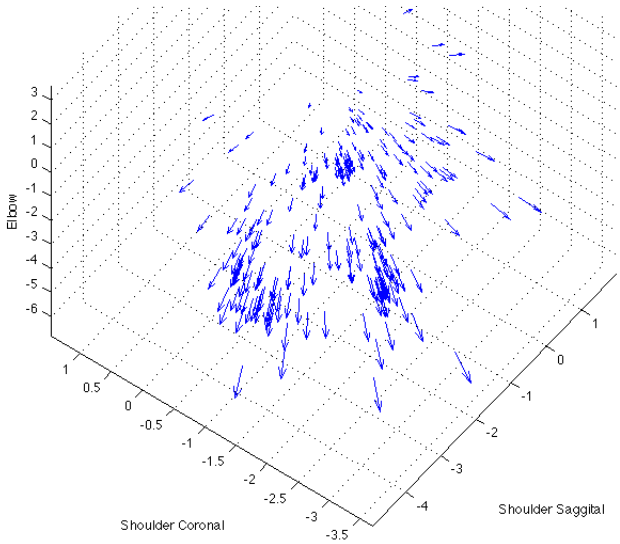



Figure 4.11: Vectors of training data in cartesian space gathered from grid experiments.

Table 4.2: Verification Check of perturbation network

| **Reach-to-grasp Testing** | |
|---|---|
| Trial Number | Percentage Successful Grasp (30 trials) |
| 1 | 25 % |
| 2 | 50 % |

Figure 4.12 shows the results of a more exhaustive test of object locations to the model. This shows that almost all of the successful grasps occur on the right-hand side and middle of the search space. By observing the model gathering data it is possible to see that it is much more difficult for the hand to grasp some object locations than others. There are a few possible reasons for this, one is that the degrees of freedom of the arm are too few to easily reach these locations (there are no degrees of freedom in the wrist), another is that the distance that the perturbations of the goal motor state are removed will only work at some of the locations in the search space.

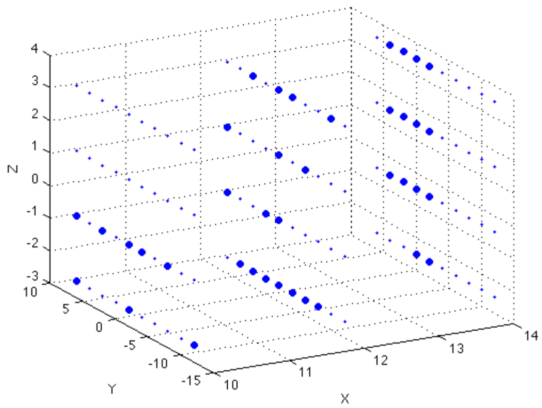



Figure 4.12: Objects that can be successfully reached in Cartesian space. Successful locations are shown as large points, unsuccessful locations as small points.

## 4.4 Summary of the model

In this chapter I have explained the model that we have used to simulate the development of simple human actions. This model has allowed us to produce a model that can perform a few primitive actions: it can reach towards a target, and it can manipulate its digits in order to grasp a target. Although its performance is not perfect the reasons for developing this model were to create a platform with which the ideas of action-effects can be examined. This model can be used for that so the remaining issues with the model will not be addressed in this thesis. It should be easy to visualise how the model could be improved by having more degrees of freedom and a more complex system of visual and sensory feedback to decide perturbations of the goal motor state. With these flaws, the model will be used as a base for the exploration of more complex actions as part of chapter 5 of this thesis.

# Chapter 5

# Implementation of causative actions

Recall from chapter 2 that we have chosen to define a class of actions which we call causative actions. In this chapter I will discuss the causative actions that we have chosen to investigate: squash and bend.

In order to do this I will use the model of trajectories (perturbation of the goal motor state) that was used in chapter 4. This will be used to learn trajectories which bring about particular causal effects. It is important to note that this model of causative actions does not depend on this model of trajectories: any other model of trajectory representation could have been used. In particular I could have used Oztop et al. (2004) model of via-points. It is also worth noting that this work makes no detailed reference to neuroscientific experiments, there are currently no detailed models of how causative actions are represented: this model is a first step into this area. As such, this model will make novel predictions that could be tested in neuroscientific experiments.

## 5.1 Model of causitive actions

### 5.1.1 Introduction

In order to create a model capable of performing causative actions I have expanded the model that was presented in chapter 4. Here I will discuss how the model was expanded and how it was trained and tested. Videos of the trained model interacting with objects can be found at *http://graspproject.wikispaces.com/Movies*.

### 5.1.2 A proposal about the reinforcement signal that trains causative actions

In section 3.3 I discussed Oztop et al's idea about the joy of grasping: that touch/grasp sensations are intrinsically rewarding, and therefore function to teach the motor system to develop grasp actions. How can this idea be extended to a system which must learn arbitrary causative actions?

My proposal is that causative actions are learned by a completely separate reward system, which is hard-wired to deliver rewards for a different class of sensory stimuli. I suggest that this reward system is *the agent's own event perception system*, which the agent uses to recognise arbitrary distal events in the world, involving arbitrary external agents and objects. Of course agents have a system for recognising external events in the world. And we know quite a lot about this system, at least as it recognises observed actions of external agents - see for instance Keysers and Perrett (2004). The system is implemented in a specialised neural pathway, the 'action recognition pathway' which was discussed in section 2.2.1, running from primary visual cortex through superior temporal sulcus and inferior parietal cortex to mirror neurons in premotor cortex.

My suggestion is that when the agent is executing an action on a target object, this action recognition pathway is engaged, *in a special mode which causes it to generate a reward signal if an action is recognised.* When the agent is executing an action on a target object, the agent's attention will be on this object. The action recognition pathway normally requires the agent's attention to be on the entity performing the action. But in the case where the agent is also *interacting* with the attended object, any action which it performs is plausibly one which the agent has actually *brought about.* If the action recognition system axiomatically delivers a reward whenever it recognises an action under these circumstances, this will function to teach the motor system about movements which bring about particular causal effects. This is the proposal which I will implement in the current chapter.

### 5.1.3 Structure of the model

One perturbation of the goal state is enough to perform simple actions, such as reach-to-grasp, but it is not enough to perform more complex causative actions. Hitting an object would require a large perturbation to the goal arm state (with a lot of force). But a single perturbation would only allow the arm to create a non-linear intial trajectory; once that perturbation is removed the arm would move to touch the object.

In order to perform a squash action it is not too difficult to imagine that you would need, at least, two perturbations: one to place the hand above the object, and another to move the hand in a trajectory through the object's position towards a point below it, so that the PID controller causes it to make contact with sufficient force. Although it is possible to think about actions which might require more than two perturbations to be applied, such as throwing, the network I describe here will only allow two perturbations. This should allow the arm to perform most actions, although I will only investigate a few in this report.

Changing the perturbation network to allow the learning of causative actions is a relatively simple process but can be done in a few ways. Firstly, we need something that 'recognises' the event that has been caused. Secondly, we need to be able to produce multiple perturbations which when provided to a PID controller can cause it to generate a complex trajectory through space. Finally, we need a way to allow the model to select different causative actions to execute on a given object (or to choose a simple non-causative reach-to-grasp or reach-to-touch action).

In this simulation the event will not actually be detected by a realistic visual routine; instead we will hard code a simple 'oracle' module that consults the physics engine directly and sends the event that is taking place to the model. For example, if the model causes an object to squash, the event recognition module will send a signal indicating that a squash event has occurred to the model. The set of perturbations that caused this event will then be flagged as a motor routine which 'cause's a squash', and is added to training data.

Producing multiple perturbations could be done using an SRN extension to the existing perturbation network. However, the actions that we have chosen to look at should only require two perturbations to learn, so, for simplicity, we have opted to have an unfolded SRN. This will have two outputs, the first perturbation and the second perturbation. The input to this will be the perturbation that caused a grasp action.

Notice that at no point in this process do we explicitly define a motor action, all of these will be learned by the effects successful motor commands bring about on the object. In this way I produce a model of motor actions where the actions are defined by the events that they cause in the environment as discussed in section 2.4.

### 5.1.4 Learning

The mechanism that we use for learning the perturbations which cause effects are similar to those used in chapter 4. The postulate here is, as the joy of grasping is used

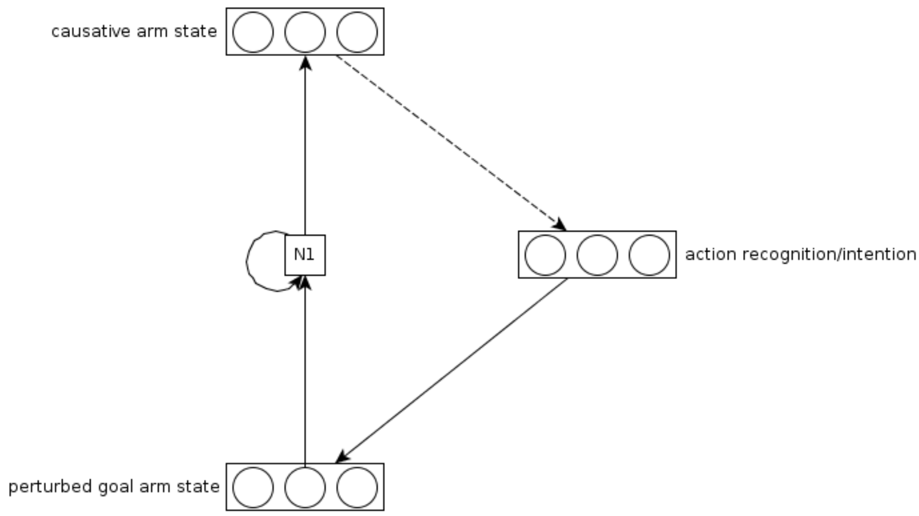



Figure 5.1: Structure of the perturbation network.

as an intrinsic reward to learn reaching and grasping actions, to learn causative actions a reward signal is axiomatically produced by events observed to happen to the target object. This signal is produced by the episode recognition module that was discussed in the previous section.

One difference to notice is that there is no gradient of these effects in these rewards: in my implementation the event signal is on or it is not. For the purposes of this project that is fine, but there may be other mechanisms involved within the neurocircuitry that restrict the event that may be considered to be squashes or bends so as to improve performance and dexterity over time.

### 5.1.5 Creating Objects

In order to explore the idea of causitive actions we need to simulate objects which have certain motor affordances. In my project I have defined three objects: one that can be grasped, one that can be squashed, and one that can be bent. Complex articulated objects are not simple to make inside the jMonkey environment, so I have made some simple objects which have the required behaviours.

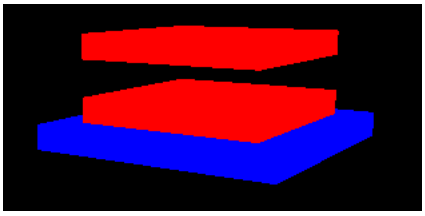


Figure 5.2: Squashable Object

The squashable object shown in figure 5.2 is made of two planes connected together by a joint. This joint is rigid enough that unless an external force is applied it will maintain a state of equilibrium holding the planes apart. If a large enough force is applied to the surface of the top plane it will press into the bottom plane; this is defined as a squash event.

It is easy to see that in order for this event to take place the hand must first be moved above the top plane and then must move towards a point under the bottom plane with enough force that the top plane will be depressed.

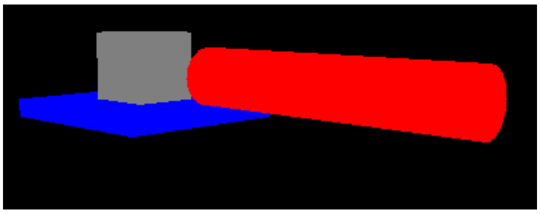


Figure 5.3: Bendable Object

The bendable object shown in figure 5.3 is made of a thin rod or 'lever' that has been attached to the box using a pivot joint, so it can pivot around the box. If a large force is applied in a downward or upward direction on the end of the lever it will rotate around the pivot; this is defined as a bend event.

In order for a bend event to take place the hand must be moved above or below the end of the lever and then must move directly down or up to move the lever through some angle.

The object that will be used in these simulations which has a grasp motor affordance is the same that was used in chapter 4: a tube which is oriented horizontally in space.

### 5.1.6 Training the network

The expanded model is trained in a similar fashion as the simple model in chapter 4. First, an object is generated in the environment. The model then generates a perturbation trajectory, using the trained perturbation network, that would cause the controller to move the hand to grasp the object. Next, the goal motor state is randomly perturbed in order to attempt to produce an effect. If a desirable effect is produced then it is added to training data.

The objects used in this simulation are the three discussed in section 5.1.5 (graspable, squashable and bendable). In order to speed training the area over which the objects have been generated has been restricted to a smaller area than in the simulations in chapter 4. Objects are able to be generated at any position in this area.

The perturbation trajectory that is first generated by this model is a simple reach-to-grasp plan. Recall, from chapter 4 that this is a goal arm state and a perturbation of that goal arm state. This perturbation is used as input to the new neural networks in our model which produce two new perturbations; these are then applied to the motion instead of the orignal perturbations in order to attempt to cause the action afforded by the object in the environment.

If a desirable effect is produced by the model it is detected by the episode recognition module that sends a signal to the model which then logs the new motor plan (made up of two perturbations) and the effect this plan causes as training data. The model trains itself each time a new piece of training data is added as it did in chapter 4 simulations.

### 5.1.7 Results

The results in figure 5.1 shows the number of actions the model performs when tested on random objects (grasp, bendable, or squashable) placed in the trained spatial extent. These tests were performed with 30 trials for each object. Results are shown in table 5.1.

Table 5.1: Results of training with three actions.

| **Causative Action Testing** | |
|---|---|
| Action Type | Percentage Successful (30 trials) |
| Grasp | 86.7% |
| Squash | 93.3% |
| Bend | 73.3% |

The first thing to notice is that the results for grasping using this model are much better than the results that were shown in section 4.3.4. This is obviously because of the restricted area in which objects are presented in the simulated environment (approximately 20% of the total area). This result lends some support to the argument that some areas of the environment are just harder for the arm to learn in due to mechanical problems.

Bend actions are successful less frequently than the squash or grasp actions. This is most likely because the bend action requires the hand to apply force in a precise location of the lever arm. In the bend action the model is required to learn where to apply force on this arm. This is unlike the squash action which requires only that the model learns to apply a certain amount of force to any part of the object.

### 5.1.8 Summary

The results of the simulations I reviewed in this chapter provide a proof-of-concept that perturbations of the goal state can be used model the learning of complex causative actions. Although the simulations in this section constrained the learning to a smaller area of space than was used in chapter 4 I have already shown in the previous chaper that this model can abstract over larger spaces. Admittedly it does perform less successfully over large areas but this does not detract from the results of learning to perform causative actions in this restricted area.

# Chapter 6

# Towards a Syntactic Interpretation of the Model of Causative Actions

In this chapter I will interpret the structure of the model I have presented in chapter 5. I will do this using a recent sensorimotor interpretation of syntax by Alistair Knott (Knott, 2012).

## 6.1 Introduction to the Minimalist Model of Syntax

Knott's analysis of syntax is performed using the generative grammar known as Minimalism(Chomsky, 1995). In this section I will briefly discuss the main ideas of the Minimalist interpretation of syntax that are pertinent to Knott's analysis and my model. I will not cover the evidence that has led to Minimalism's current form; for a complete review of the theory see Carnie (Carnie, 2002).

### 6.1.1 X-bar Schemas

Phrase structures are built from conjunctions of simple **x-bar schemas**, also known as XPs. Each word in a sentence **projects** its own structure in the form of one of these schemas as shown in figure 6.1. These schemas are connected together to create the phrase structure of the whole sentence.

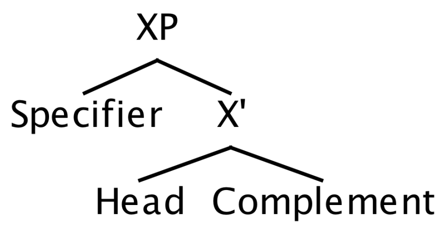



Figure 6.1: An X-bar schema

Each grammatical class of word projects a different kind of x-bar schema, but they all have the same basic structure. The word appears at the **head** position. The two other open slots in an XP, the **complement** and **specifier**, are places where other XP structures can appear. The head provides semantic information for the structure; the complement and specifier provide slots for additional semantic information regarding the head. The phrases or XP structures which occupy these slots are known as **arguments**, this is shown in figure 6.2.

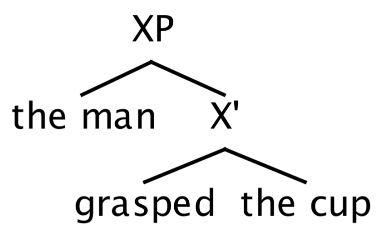



Figure 6.2: The phrases present at the specifier and complement positions (*the man* and *the cup*) of this XP are arguments of the head (*grasped*).

Head positions do not have to be words, they can also contain other elements such as inflections. XPs headed by words are called **lexical projections**, other XPs are known as **functional projections**. Some of the different forms of XP that I will use are: VP (verb phrases), DP (determiner phrases), IP (inflection phrases) and AgrP (agreement phrases).

### 6.1.2 Phonetic Form and Logical Form

Some syntactic theories, such as Minimalism, posit that there is more to language than the surface level. Minimalism uses this other level to try to give a concise explanation of the differences that exist between different human languages. The surface form of language is called the **phonetic form** (PF); this includes details such as word order, intonation and phonology. The underlying structure is called the **logical form** (LF); this is the level that syntax interacts with meaning.

The PF of a sentence is the phonetic representation of that sentence. This phonetic representation is then pronounced by the speaker. The LF of a sentence is represented as a right-branching tree of connected XP structures, this is shown in figure 6.3.

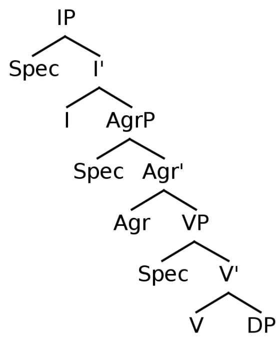



Figure 6.3: The logical form of a simple sentence. IP is the inflection x-bar, AgrP is the agreement x-bar, VP is the verb x-bar, and DP is the determiner x-bar. Evidence for these different projections can be found in Carnie (2002).

The LF of a sentence in all languages is identical if those sentences have the same semantic meaning. The differences in the surface form (or PF) of different languages is attributed to intermediate transformations of the underlying LF, these are called **transformational rules** or **movements**.

### 6.1.3 Movement

There are two kinds of movement that occur in the Minimalist interpretation of natural language that I will discuss, these are **head movement** and **DP movement**. Both of these forms of movement allow the generation of sentences in orders that Minimalism would not otherwise be able to be produce.

DP movement is motivated by the need to explain sentences with the same semantic meaning but with differing surface forms. Minimalism requires that these sentences share the same LF, in order to produce a different PF some of the DPs must undergo movement operations before pronouncement. An example in English are the two sentences in example (1).

(1) a. John seems to walk
    b. It seems John walks

The LF with movement of this sentence is shown in figure 6.4. *Seem* is known as a **raising verb**, not all languages have these but Minimalism must be able to explain these using movement in the LF. This movement operation can produce either of the sentences in example (1) depending on whether the movement is performed before or after the sentence is pronounced from PF.

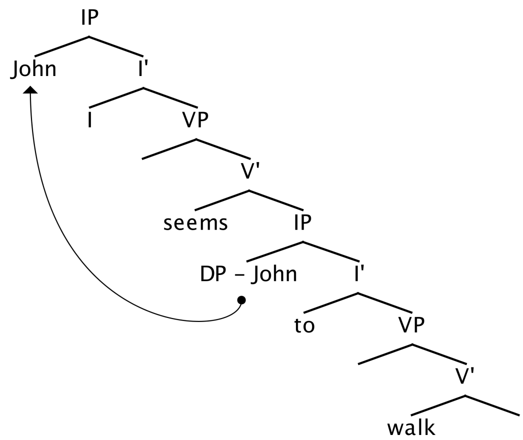



Figure 6.4: DP movement

**Head-to-head** movement (or head movement) allows words at a head position in an XP to raise to a higher head positive whilst leaving a **trace** of their semantic information behind. Figure 6.5 shows an example of head movement in an English sentence.

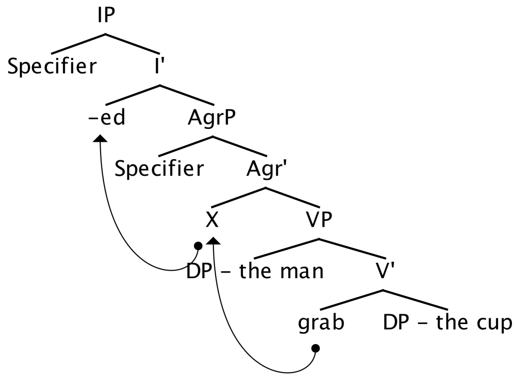



Figure 6.5: A verb head raises to its inflection in the IP projection

The major linguistic argument for this is inflections on verbs. These **morphemes** provide additional information, in particular information which agrees with the verbs subject and objects. Linguistic agreement refers to a covariant semantic property of one element in a sentence with another. An example in English is *-s* as in *lift-s*; *-s* indicates agreement with a simple subject. Verbs in English are not always inflected, in these cases they will appear with an auxiliary verb such as *will lift* or *did grasp*.

In order to explain agreement, we have to assume that at some point, the verb sits in some local syntactic relationship with the subject. For some languages this will not be the same point that the verb will be pronounced, so it must move from this location before pronouncement. This allows us to explain the different inflection and verb orderings in other languages; for example, in French, adverbs come after verbs. In French this leads to head raising above the adverb, while in English the head is lowered below the adverb. This is further explained through the use of different **overt** and **covert** movements in languages.

The idea of overt and covert movements is that all movements occur in the LF regardless of which language is actually being spoken but that these movements can happen before or after the sentence is pronounced. Movements that occur before pronouncement are overt (the movements are noticed in the surface form); those that occur after pronouncement are covert movements. Figure 6.6 shows the overt and covert movements involved in producing the surface form of the English sentence *Alice grabbed the cup.*

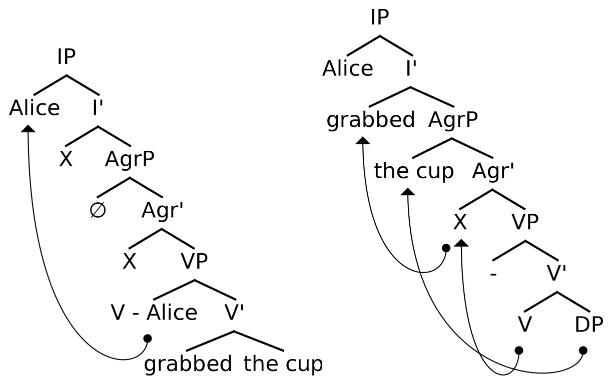



Figure 6.6: Left: shows the overt movements in this sentence. Right: covert movements that take place after pronouncement.

The overt and covert movements that would occur to produce sentences in languages with different word orderings are different; that is what allows the production of different word orderings based on a single LF.

### 6.1.4 Minimalist Structure of a transitive sentence

A transitive sentence is one that contains a transitive verb; this is a verb that has a subject and must also have at least one object. Example (2-a) shows a simple transitive sentence, where the verb *grabbed* has a subject, *Alice*, and one object, *the cup*. This verb is known as a monotransitive verb, a ditransitive verb is shown in example (2-b). In this example the two objects are *the cup* and *Bob*.

(2) a. Alice grabbed the cup
    b. Alice gave the cup to Bob

Figure 6.7 shows the LF structure of the transitive sentence shown in example (2-a).

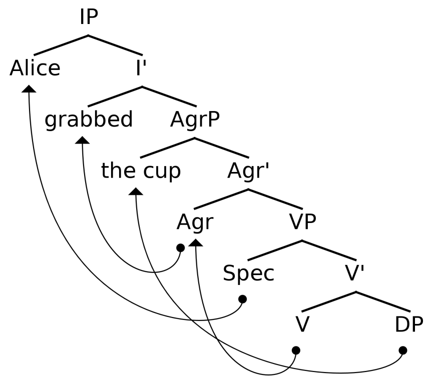



Figure 6.7: LF structure of a transitive sentence

### 6.1.5 VP-shells

The Minimalist model's strong use of X-bar schemata means it is committed to the **binary branching constraint** (Kayne, 1984). This states that any syntactic structure with more than two arguments is invalid, so each node in a syntax tree can have, at most, two children.

Double-object sentences, those with a **ditransitive** verb, pose a prima facie problem for the binary branching constraint. In these sentences the verb needs to have three arguments, the subject, and two objects. The proposed solution to this in Minimalism is that these ditransitive verbs can be decomposed into two semantic parts, the cause and the effect. Example (3-a) shows a simple double-object sentence which can be decomposed to example (3-b).

(3) a. Alice put the cup on the table
    b. Alice CAUSE the cup GO on the table

Notice that we represent CAUSE and GO as abstract heads, these structures have become known as **VP shells**(Larson, 1988). An example of a binary-branching syntax tree of a double-object sentence is given in figure 6.8. In this example the subject, *Alice*, raises to the normal position and the verb *go* raises to the position of *cause* and join it where it is pronounced as *put*.

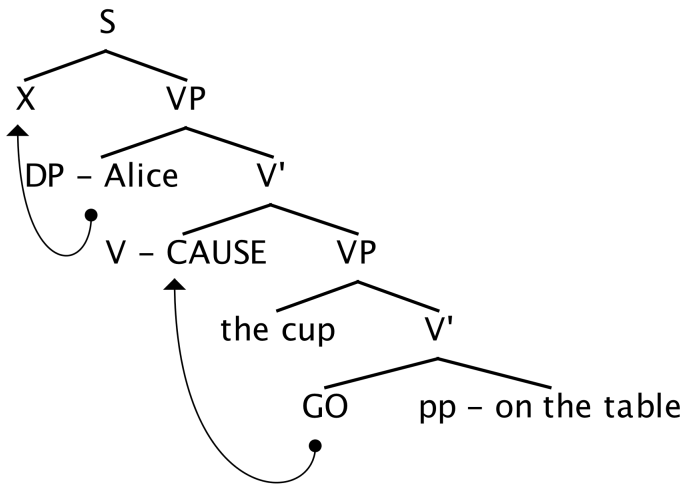



Figure 6.8: LF of a double-object sentence using VP shells

As well as covering double-object sentences VP shells can also accommodate the linguistic phenomenon named the causative alternation which I presented in section

6.1.5. Recall that this allows transitive verbs to be rephrased as intransitive verbs. Example (4-a) and (4-b) shows the causative alternation pair of *broke*. VP-shells allow us to interpret this as example (4-c).

(4) a. Alice broke the cup
  b. The cup broke
  c. Alice CAUSED the cup TO BREAK

Figure 6.9 shows the form of the LF of the sentence *The man squashed the ball* after all covert and overt movement has taken place.

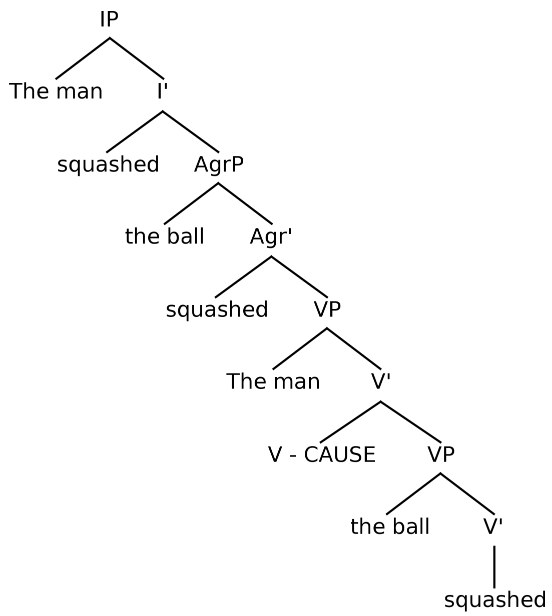



Figure 6.9: LF Syntax tree of *The man squashed the ball*

### 6.1.6 VP-shells and action-effects

VP-shells provide a theoretical solution to the problem of double-object sentences and the causative alternation. In addition to this VP-shells can also provide some evidence for effect-based representation of actions. Language is fundamentally a description of our experiences, the fact that our analysis of language provides a model which clearly separates action-effects and their causes shows how important this separation is.

An even stronger interpretation of this, proposed by Knott (2012) is that the LF of a sentence shows a strong isomorphism with sensorimotor processes. It may be that LF is actually a description of the sensorimotor processes involved in performing that action. A conclusion that you can draw from this is that the agent focuses its attention on itself, and then observes the movements it makes to bring a certain change in the state of an object. In the rest of this chapter I will briefly introduce Knott's sensorimotor interpretation of LF structure, and make some suggestions about how it can be applied to causative actions of the kind I have modelled in my system.

## 6.2 Introduction to Knott's Sensorimotor Interpretation of Logical Form

Recall from section 6.1 that I discussed the right-branching structure of a syntax tree which is constructed from simple X-bar schemas. Recall also that the syntactic theory that I have used (Minimalism) posits both a surface and underlying structure known as PF and LF respectively. The LF of a sentence describing the squash action that I considered in chapter 5 is shown in figure 6.9.

In Knott's analyses of syntax and motor action it is suggested that there is a strong isomorphism between the Minimalist model of LF and the sequence of sensorimotor actions that take place during action. His contention is that the LF of a sentence, like the one in figure 6.9, can be understood as a description of the sequence of sensorimotor operations involved in that action. In Knott's model the LF of a sentence describing a concrete episode in the world can be interpreted as a description of the sensorimotor process through which this episode was perceived. This model is an example of an 'embodied' model of language: an LF structure, which basically encodes the meaning of a sentence, is a simulation of a perceptual process. In Knott's sensorimotor model, a transitive reach-to-grasp action is perceived through a well-defined sequence of sensorimotor operations: first an action of attention to the agent, then an action of attention to the target object, then the activation of a motor program. The episode can be perceived either by the agent of the action himself, or by an external observer: in the former case, the motor program is activated for real, in the latter, it is activated by circuitry in the mirror system. For simplicity, I will only consider the case where the observer is the agent.

In Knott's interpretation of LF, each of the X-bar schemas in the LF tree structure describes a single sensory or motor operation performed by the observer, and the right-branching structure of X-bar schemas describes the sequential structure of these operations. This is illustrated for the LF of 'John grabbed the cup' in figure 6.10. What's more, each part of an X-bar schema has a specific sensorimotor interpretation, as shown in figure 6.11. The XP describes the context in which the sensorimotor oper-

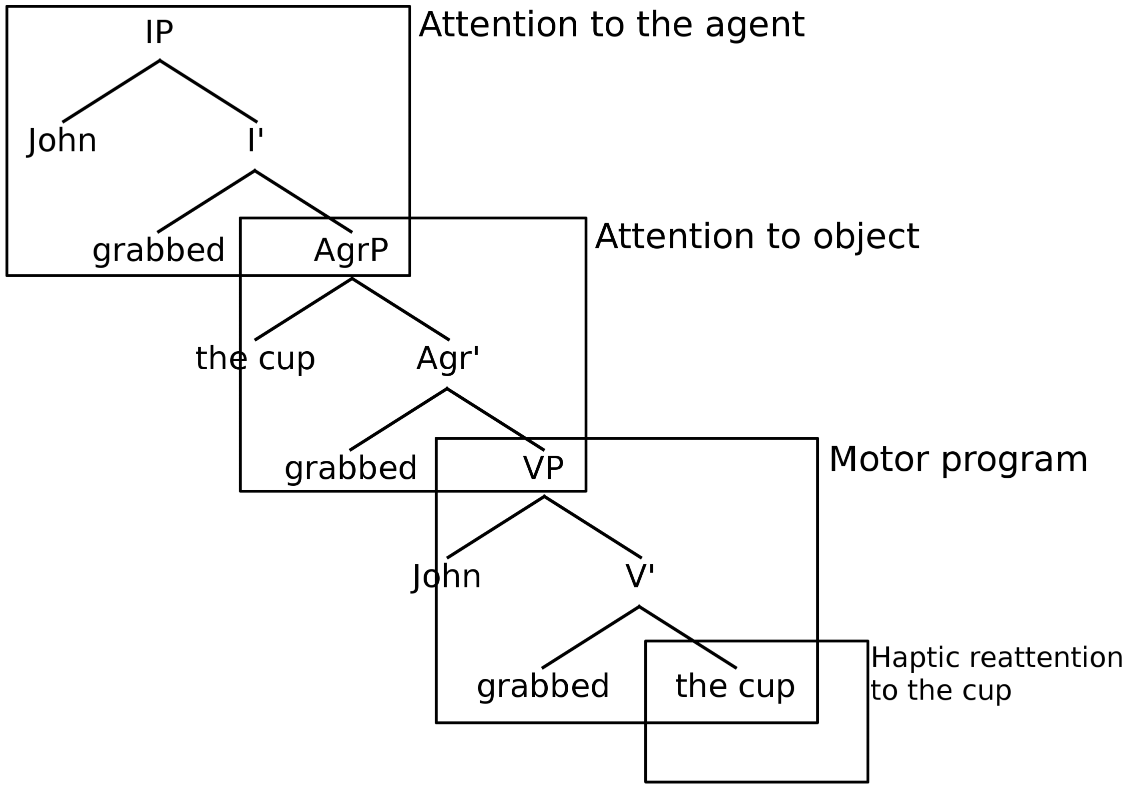



Figure 6.10: Syntax tree interpreted as a description of a sequence sensorimotor operations.

ation is executed. The head, X, describes the operation itself. The specifier describes a reafferent sensory consequence of the operation, and the complement describes the new context which the operation brings about. This means that every position in the LF of a sentence has a sensorimotor interpretation.

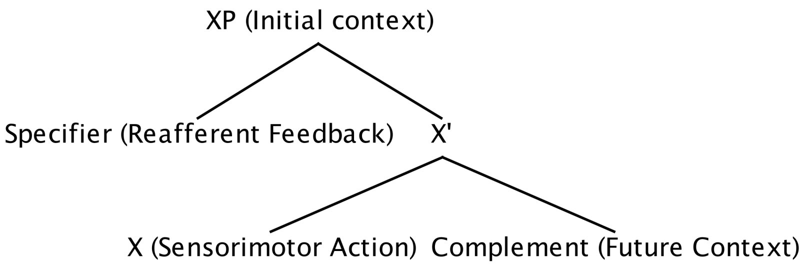



Figure 6.11: Sensorimotor characterisation of X-bar schema

One relevant aspect of Knott's interpretation is that head (X) positions are assumed to describe sensorimotor operations as they are planned, rather than as they actually occur. When a sequence of motor or attentional operations is planned, the area of the brain which holds the plan is the prefrontal cortex, rather than areas representing actual attentional or motor operations. The actions in the plan are active in parallel in this prefrontal area, even though the plan results in a sequence when it is executed (Averbeck et al., 2002). Knott uses this property to explain the phenomenon of head movement. Recall from Section 6.1.3 that verbs and their inflections heads move from lower head positions to higher positions: the inflected verb is actually present at all three head positions in the clause. Knott suggests that the verb stem is read from the planned motor action, and its inflection is read from the planned action of attention to the agent. Both these elements are tonically active while the plan is executed, and so can be pronounced at any time.

## 6.3 Interpretation of My Model

### 6.3.1 Interpretation of VP shells

The LF of a sentence describing one of the actions which I modelled in chapter 5 is shown in figure 6.9. Recall that a VP shell is used to represent a sentence which is able to undergo the causative alternation. This provides an unambiguous LF representation which is consistent across languages.

In Knott's analysis the first actions in the sequence are attending to the agent (self) and then attending to the object. These are described in the LF structure by the heads of the IP and AgrP projections.

Knott's hypothesis that a right-branching sequence of two XPs describe a sequence of two sensorimotor operations makes a prediction about the nested VP structure shown in figure 6.12: the VPs in the tree must describe other sensorimotor operations occurring after the agent has focussed attention on the object it is intending to manipulate.

In the Minimalist model, the suggestion is that the inner VP denotes the action which the target object is caused to perform, or undergo. This is what I have suggested in chapter 2: that there is a class of actions that are learned by the effects that they cause in the agents environment. In the tree the inner VP is dominated by the outer CAUSE VP which captures the fact that the agent is bringing about this effect.

So how does this relate to the structure of my motor model? If the model that I have produced, which can perform these kinds of causative actions, can also be interpreted in a similar structure this may lend computational support to Knott's interpretation. The

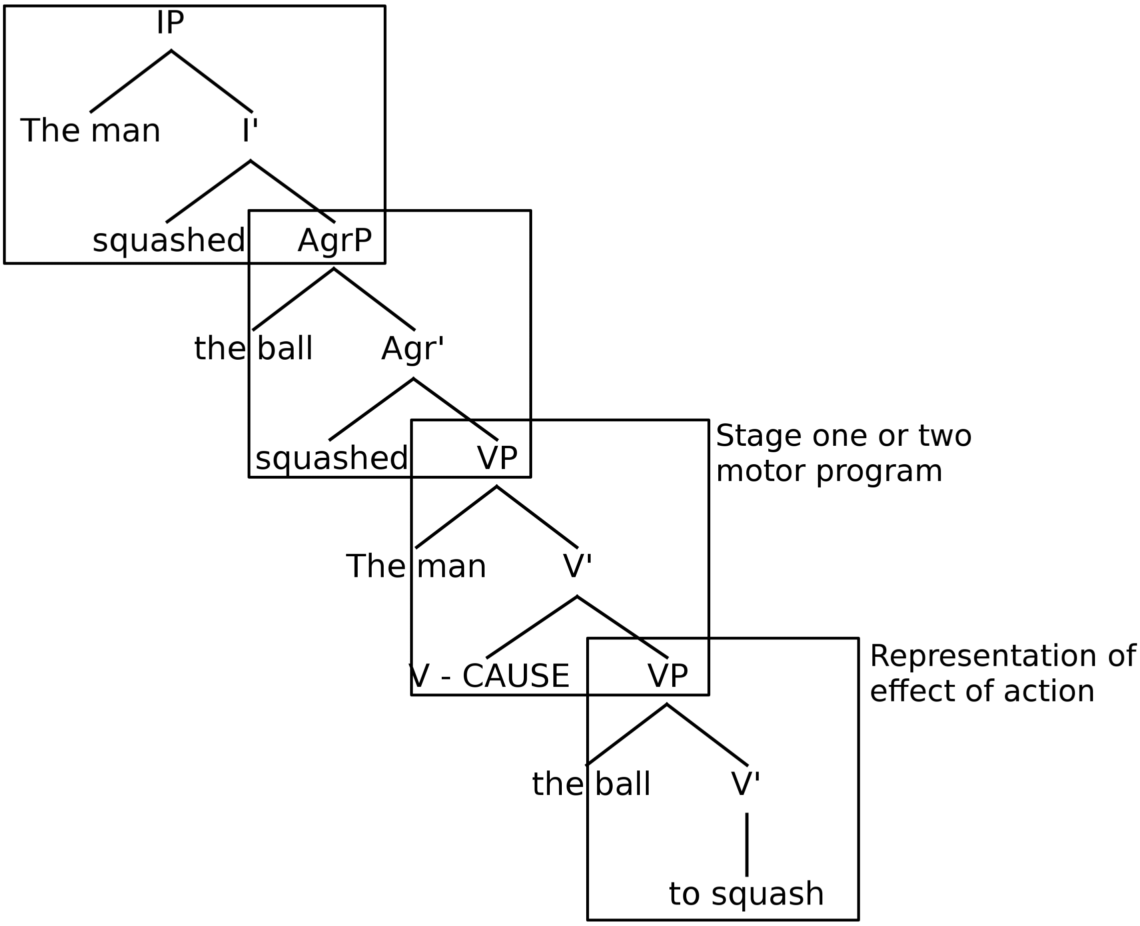



Figure 6.12: VP shells of figure 6.9: *the man squashed the ball*

model that I have produced has three components: the reach network, the perturbation network, and the causative action network[1].

Each of the networks that I have produced corresponds roughly to one of the actions in the motor plan that Knott has proposed.

### 6.3.2 Higher VP

In my analysis I will take the outer VP in figure 6.12 (closest to the top) to be a correlate of stages 1 and 2 of my model (the reach and perturbation networks respectively). The motivation of this is a result of the analysis of this X-bar using Knott's sensorimotor interpretation.

In Knott's interpretation, the higher VP node corresponds to the context in which the causative motor action is initiated; in this case the agent will be attending to the

[1]At this point this network is actually an unfolded SRN, this was done only to make the experimentation easier. A normal SRN could have been used.

target of this motor sequence. The specifier (which I am interpreting as reafferent feedback) is the agent; this makes sense in this situation as the agent is moving and generates a representation of its own body as it moves through space. The head of this VP is the sensorimotor action that the agent is performing, in this case this is *CAUSE*. What does this mean? In Minimalism a word such as *squashed* is just a complex of this head and the head in the inner VP as described in section 6.1.5; this is accomplished by head-to-head movement before the LF is pronounced. The basic idea is that these parts of the word *squashed* are read out from the action plan representation in the PFC.

This idea means that we can actually think of the head *cause* as being a part of the sensorimotor sequence required to bring about this action. It is also a part of the caused episode that is represented in the inner VP. The sentences that Knott has analysed and I presented in section 6.2 showed that grasp actions are represented at this head. I put forward the idea that outer VP schema represents the activation of a simple stage one or two motor program in which the agent attempts to reach or grasp a target ending with the agent touching the object.

### 6.3.3 Lower VP

My other proposal is that the inner VP of figure 6.12 is a representation of the event that the causative action brings about. Once again this is motivated by an analysis of the structure of the X-schema.

The initial context (as represented by the lower VP node) is the agent touching the object of the action. This is motivated by the interpretation I have for the outer VP: the action that it represents ends with the agent touching the object. In this case the specifier is the object of the action, this may seem contradictory to the statement that I made above: that the agent gets reafferent feedback as it moves. This can be understood as instead meaning that the reafferent feedback is the object as it's configuration changes in space. This also makes sense in my argument that this represents the effects of an action since we are now monitoring the object rather than the motor system.

I have proposed that the outer VP head is part of a sensorimotor sequence that makes up a complex causative action. This means that the head of the inner VP MUST be the PFC representation of the other part of this action. This other part of the action is more abstract than a simple reach or grasp; it is defined only by the effect that it brings about on the object. This is exactly analogous to the causative network that I have produced, it is trained only on the effect that is produced on a target object.

## 6.4 Clarification of interpretation

In this chapter I have presented a modern model of syntax and an interpretation of the model that I have produced which provides some evidence that this model is, at least, somewhat accurate in its treatment of human motor action. However, I do not want the reader to take away too much from this proposal, this is a very early idea and the model that I have produced here still has problems. As these could also be related to problems that exist in my syntactic interpretation I will discuss all of these problems in chapter 7.

# Chapter 7

# Discussion

The model that I have created in this project has accomplished the goals that I set out at the beginning of this thesis. This model is able to grasp without precomputing trajectories (as suggested by neuroscience data). The idea of perturbations of the goal arm state was also able to be extended from learning simple actions to learning actions which are defined by the effects that they cause, our so-called causative actions.

This chapter is devoted to the discussion of the structure of the model that I have produced. This will mostly focus on the problems that I have discussed at the end of chapters 4 and 5. I will also attempt to pose experiments or observations that can validate or invalidate this model. Lastly, I will discuss the syntactic interpretation that I proposed in chapter 6.

## 7.1 Main Points of the Model

In order to save the reader effort I will first list what I consider to be the most important theoretical information and aspects of the model that I have constructed. First, I have made a case for a 'class' of actions which seem to be learned based on the effects that they bring about in the world. This was used to motivate the idea that these observed effects of action can be used as reward signals in learning causative actions. Secondly, using perturbations of the goal state I was able to produce a model that can learn to perform actions which do not precompute the entire trajectory which is consistent with recent studies. This idea allowed my simulated arm to learn to grasp an object in its environment. Thirdly, the model that I produced uses a distributed encoding on the output, this allows it to represent multiple alternative trajectories. Lastly, this idea of perturbations was extended using an unfolded SRN. This allowed the model to learn more complex causative actions.

## 7.2 Structure of the model

The structure of my model suggests grasping an object is a more primitive action than the causative actions. This is shown in the child development literature by the early stage that this is learnt. However, I make a stronger claim: the grasp action is an underlying prerequisite for learning more complex actions.

This is obviously true for actions which involve grasping objects, but what about the actions that I have looked at: squashing and bending. There does not seem to be a reason that the information used to learn to grasp an object would be useful to learn these actions.

There are two parts to a reach action in my model: the initial reaching action (reach-to-touch) and the perturbation of this (which creates a different trajectory). It is intuitive to think that the information of the reach action will be used for more complex causative actions. Most of these also involve the agent interactive physically with an object so that agent must be able to make contact with it. We can also imagine that the perturbations that are learnt for a reach-to-grasp action also contain information about what is easy for an agent to do. For example, a perturbation contains information about which direction the arm should approach the target. This is most likely to be the direction that would most naturally orient the palm towards the object. This information is not exclusive to a grasping action, it is also useful for both the squashing and bending actions. While this only means that some of the information is being used in both, this is a difference in granularity of the representation.

## 7.3 Problems with this Model and Future Work

### 7.3.1 Application of Perturbations

The perturbations that I have proposed in this report must be removed from the action as it unfolds in order for the hand to reach the target. When learning to grasp an object this removal is done when the hand reaches a certain distance from the object. This is a fairly arbitrarily defined metric which could potentially be learned by the model. Doing this may dramatically increase the complexity of learning making it a questionable solution.

Perturbations must also be removed when the model is learning a complex action, although this is less important than it was when grasping. However, the actions that I considered in this project (squash and bend) are not easy to learn using the same mechanism of removing perturbations that I discussed above. In the model the first perturbation was only removed when the arm reached the perturbed goal arm state. In these simulations the perturbation is treated more like a waypoint. The distinction to note is that these are not precomputed and that they are all calculated in an object centered reference frame.

Generally, it would be best if the mechanism for removing perturbations was consistant across all my simulations. Because the controller that I used is not needed to produce the trajectories that are produced in experiments, but there is no reason to suggest that this mechanism should suddenly change at a non-discrete point in child development.

### 7.3.2 Circuitry Granularity

Another issue with this model is the point at which I make the transition from learning a grasp action to learning a causative action. The learning that has been done using the perturbation network is only used as a rough spatial guide for the learning that is done in the causative action network. In this case we notice a marked drop in performance of the model during this point that may not be there in real infant action learning.

This problem may just be to do with the granularity of this model and the rate at which new pieces of neural circuitry are 'switched on'. It is more realistic that the pieces of neural circuitry are more finely grained and may become active in a highly non-linear fashion (or may always be on). If the performance of novel motor actions are consistently worse than the learned grasp action coarse separation and activation may be supported.

### 7.3.3 Performance

The last of the problems with this model that need to be addressed is its performance when generalising to objects at arbitrary points in space. Generalisation is good when the model is learning to reach-to-touch, but it is much worse when learning to reach-to-grasp and has not really been tested much when learning causative actions.

Some of the bad performance of the model relates to the difficulty of gathering data in some portions of the simulated environment. The model of the arm that I have used is no where near as dextrous as a real human arm, nor does it have the degrees-of-freedom to move. This means that object positions which would normally be reachable with a wrist movement (which my model does not currently learn) would instead only be reachable by a very small subset of the possible joint orientations. This makes getting much data for these positions statistically unlikely so the neural network is unlikely to learn how to interact with objects in those positions.

## 7.4 Discussion of Syntactic Interpretation

In chapter 6 I suggested a syntactic interpretation of the computational model that was discussed in chapter 5. Here I will discuss some of the problems with this interpretation and propose some tests or observations that could validate or invalidate this model.

One problem with the syntactic interpretation that I posit is that it doesn't account for the precise neural circuitry that I have modelled. Knott's interpretation applied to my model produces four sensorimotor frames of action, one for each of the X-bar structures in the tree. I am only modelling two of these, the sensorimotor operations involving the two VP structures at the bottom of the tree. In order to do this I use three circuits (the neural networks), this suggests I am not modelling this at the same granularity that the LF structure may describe.

The sensorimotor interpretation of the causative alternation that I have put forward makes certain predictions about the areas of the brain that should be active when performing causative actions. This predicts that areas involved in the action recognition pathway (e.g. STS and inferior parietal cortex) will have different activity when performing a squashing or bending action than when they are performing a regular transitive action like grasping. This has the obvious problem that current imaging techniques may not have the granularity to probe this accurately. Given the complexity of the brain if this prediction is true it is not conclusive proof; it merely suggests that this is a possible explanation.

# References


B. Averbeck, M. Chafee, D. Crowe, and A. Georgopoulos. Parallel processing of serial movements in prefrontal cortex. *Proceedings of the National Academy of Sciences*, 99(20):13172–13177, 2002.

A. Billard. Learning motor skills by imitation: A biolocially inspired robotic model. *Cybernetics and Systems*, 32:155–193, 2001.

Arthur E. Bryson and Yu-Chi Ho. *Applied optimal control: optimization, estimation, and control*. Blaisdell Publishing Company or Xerox College Publishing., 1969.

A. Carnie. *Syntax: A Generative Introduction*. Wiley-Blackwell, 2002.

T. Chaminade, A. N. Meltzoff, and J. Decety. Does the end justify the means? a pet exploration of the mechanisms involved in human imitation. *NeuroImage*, 15(2):318 – 328, 2002. ISSN 1053-8119.

N. Chomsky. *The minimalist program*. MIT Press, 1995.

Paul Cisek. Neural representations of motor plans, desired trajectories, and controlled objects. *Cognitive Processing*, 6:15–24, 2005.

C. Colby and M. Goldberg. Space and attention in parietal cortex. *Annual Review of Neuroscience*, 22:399–349, 1999.

G. di Pellegrino, L. Fadiga, L. Fogassi, V. Gallese, and G. Rizzolatti. Understanding motor events: a neurophysiology study. *Experimental Brain Research*, 91:176–180, 1992.

J. L. Elman. Finding structure in time. *Cognitive Science*, 14(2):179 – 211, 1990.

L. Fogassi, V. Raos, G. Franchi, V. Gallese G. Luppino, and M. Matelli. Visual responses in the dorsal premotor area f2 of the macaque monkey. *Experimental Brain Research*, 128:194–199, 1999.

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

A.P. Georgopoulos, A.B. Schwartz, and R.E. Kettner. Neural population encoding of movement direction. *Science*, 233:1416–1429, 1986.

A. M. Gordon. Development of the reach to grasp movement. In *Insights into the reach to grasp movement*. Elsevier Science, 1994.

A. F. Hamilton and S. T. Grafton. Goal representation in human anterior intraparietal sulcus. *The Journal of Neuroscience*, 26(4):1133–1137, 2006.

B. Hommel. Inverting the simon effect by intention. *Psychological Research*, 55:270–279, 1993.

B. Hommel. Event files: feature binding in and across perception and action. *Trends in cognitive science*, 8(11):494–500, November 2004.

B. Hommel. Action control according to tec (theory of event coding). *Psychological research-psychologische forshung*, 73(4):512–526, July 2009.

B. Hommel, J. Muesseler, G. Aschersleben, and W. Prinz. The theory of event coding (tec): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24(5):849–878, 2001a.

B. Hommel, J. Muesseler, G. Aschersleben, and W. Prinz. Codes and their vicissitudes. *Behavioral and Brain Sciences*, 24:910–937, 2001b.

A. Iriki, M. Tanaka, and Y. Iwamura. Coding of modified body schema during tool use by macaque postcentral neurones. *NeuroReport*, 7:2325–2330, 1996.

M Jeannerod. Reaching and grasping. Parallel specification of visuomotor channels. In W Prinz and B Bridgeman, editors, *Handbook of Perception and Action Volume II: Motor Skills*, pages 405–460. Academic Press, 1996.

L.J. Kamin. Selective association and conditioning. In *Fundamental issues in instrumental learning*, pages 42–64. Halifax, Canada: Dalhousie University Press, 1969.

M. Kawato, Kazunori Furukawa, and R. Suzuki. A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics*, 57:169–185, 1987. ISSN 0340-1200.

R. Kayne. *Connectedness and Binary Branching.* Foris, Dordrecht, 1984.

C. Keysers and D. Perrett. Demystifying social cognition:a hebbian perspective. *Trends in Cognitive Science*, 8(11):501–507, 2004.

A. Knott. *Sensorimotor Cognition and Natural Language Syntax.* MIT press, 2012.

Richard K. Larson. On the double object construction. *Linguistic Inquiry*, 19(3): 335–391, 1988.

B Levin. *English verb classes and alternations: A preliminary investigation.* University of Chicago Press, Chicago, Ill, 1993.

J. Majdandžić, M. J. Grol, H. T. van Schie, L. Verhagen, I. Toni, and H. Bekkering. The role of immediate and final goals in action planning: An fmri study. *NeuroImage*, 37(2):589 – 598, 2007.

K. Matsumoto, W. Suzuki, and K. Tanaka. Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science*, 301:229–232, July 2003.

R. E. Morin and D. A. Grant. Learning and performance on a key-pressing task as function of the degree of spatial stimulus-response correspondence. *Journal of experimental psychology*, 49:39–47, 1955.

T. Neumegen. Grasp. Master's thesis, University of Otago, In progress.

E. Oztop, N. Bradley, and M.A. Arbib. Infant grasp learning: a computational model. *Experimental Brain Research*, 158:480–503, 2004.

W. Prinz. Perception and action planning. *European Journal of Cognitve Psychology*, 9(2):129–154, 1997.

L. Riggio, L. de Gonzaga Gawryszewski, and C. Umiltà. What is crossed in crossed-hand effects? *Acta Psychologica*, 62(1):89 – 100, 1986.

G. Rizzolatti and L. Craighero. The mirror-neuron system. *Annual Review of Neuroscience*, 27:169–192, 2004.

A. Schlack, S. J. Sterbing-D'Angelo, K. Hartung, K. Hoffmann, and F. Bremmer. Multisensory space representations in the macaque ventral intraparietal area. *The Journal of Neuroscience*, 25(18):4616–4625, 2005.

W. Schultz. Getting formal with dopamine and reward. *Neuron*, 36:241–263, 2002.

E. Shikata, F. Hamzei, V. Glauche, M. Koch, C. Weiller, F. Binkofski, and C. Bchel. Functional properties and interaction of the anterior and posterior intraparietal areas in humans. *European Journal of Neuroscience*, 17:1105–1110, 2003.

J.R. Simon and A.P. Rudell. Auditory s-r compatibility: the effect of an irrelevant cue on information processing. *Journal of applied psychology*, 52(3):300–304, 1967.

R. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 3(1):9–44, 1988.

R. Sutton and A. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

E. Tunik, S. H. Frey, and S. T. Grafton. Virtual lesions of the anterior intraparietal area disrupt goal-dependent on-line adjustments of grasp. *Nature*, 8(4):505–511, April 2005.

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

M.A. Umiltà, E. Kohler, V. Gallese, L. Fogassi, L. Fadiga, C. Keysers, and G. Rizzolatti. I know what you are doing: A neurophysiological study. *Neuron*, 31(1):155 – 165, 2001.

Y. Uno, M. Kawato, and R. Suzuki. Formation and control of optimal trajectory in human multijoint arm movement. *Biological Cybernetics*, 61:89–101, 1989.

S. A. Verschoor, M. Weidema, S. Biro, and B. Hommel. Where do action goals come from? evidence for spontaneous action-effect binding in infants. *Frontiers in Psychology*, 1(0), 2010.

D Wolpert, C Miall, and M Kawato. Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2(9):338–347, 1998.