

A COMPUTATIONAL MODEL  
OF THE CORTICAL MECHANISMS  
INVOLVED IN PRIMATE GRASPING

by

Andrew H. Fagg

-----

A Dissertation Presented to the  
FACULTY OF THE GRADUATE SCHOOL  
UNIVERSITY OF SOUTHERN CALIFORNIA

In Partial Fulfillment of  
the Requirements for the Degree  
DOCTOR OF PHILOSOPHY  
(Computer Science)

May 1996

Copyright 1996

Andrew H. Fagg

*This dissertation is dedicated to my parents,  
Willis and Dottie,  
who have been a constant source of teaching and support;  
and to my sister,  
Maryann (ow,la,ne),  
who has always been there.*

## Acknowledgments

In my time at USC, I have had the tremendous pleasure to work with Michael Arbib and George Bekey. Both have shared significant amounts of their time and experience, from which I have benefited greatly. In addition, they have given me the freedom to explore many different issues, but also the guidance (and puns) to push me through to the end.

Besides allowing me to run large processes and use massive amounts of disk space, Scott Grafton invited me to spend time working in his PET laboratory. There, I learned experimental design, PET analysis, and a beginnings of human functional neuroanatomy.

Several individuals in particular have had various levels of responsibility in helping and encouraging me as I was beginning the research game. At CMU these include Jay McClelland, Jonathan Cohen, Kevin Singly and Scott Fahlman; and at GD[P|E]: Lealon Watts, Patrick Simpson, Russell Deich and Harold McBeth.

Several Professors and senior Ph.D. students served as valuable mentors in my first few years at USC: Thea Iberall (who continues to do so), Deliang Wang, Bruce Hoff, Peter Dominey, Reza Shadmehr, Irwin King and Ken Goldberg.

Thanks to Michael "el hombre de muchos nombres" McHenry, Gaurav Sukhatme, Joel Hoff, Jim "Armstrong" Montgomery, Man-Wai "Jason" Chu (who is aiding in the webification of this work) and Nicolas Schweighofer for their many useful (and other) comments on various versions of this research and associated documents.

Many others in the robot and brain simulation laboratories have contributed to and influenced my work and life -- including: David Lotspeich, Tony Lewis, Parag Batavia, Scott Cozy (who assisted in construction of the grasping device for the PET experiment), Amanda Bischoff, Jim Liaw, Fernando Corbacho, Jozsef Fiser, Mathew Lamb, Michael Crowley, Jean-Marc Fellous, Lucia Simo, Yuri Pryadkin (who aided in the PET behavioral analysis), Jacob Spoelstra, Osman Qamar, Alex Guazzelli, and (from CS in general) Benjamin Smith.

I have also been afforded several opportunities to interact with experimentalists outside of USC. Giacomo Rizzolatti welcomed me into his laboratory for a number of weeks in the Spring of 95. There, I gained a much greater appreciation for what goes into extracting data from even a single cell. Those at the Parma lab who were particularly influential include: Luciano Fadiga, Vittorio Gallese, Leo Fogassi, Massimo Matelli, Maurizio Gentilucci, Giuseppe Luppino. Visits to the labs of Hideo Sakata and Marc Jeannerod have also been tremendous learning experiences. In particular, I thank Masato Taira and Akira Murata for their valuable interaction.

The old CMU gang has also been a source of encouragement and diversion: Ivan Yanasak, David Sandler, Jeffrey "mover of pdp11s" Meckler, Jeff Snyder, Bill Wilcox, Ira Jaslow, Hiroshi Saito and Phil Kragnes.

Special thanks are also due to my family for all their support, patience, and encouragement over these years.

Funding for this work has been provided in part by the Human Frontiers Science Project, and by a fellowship from the University of Southern California, School of Engineering, and the Computer Science Department. The PET imaging experiment was supported in part by Public Health Service grant KS-08 NS01568. The dataglove used in the experiment was provided by Michael Clark at Apple Computers.

## Table of Contents

Acknowledgments .....	iii
List of Tables .....	ix
List of Figures .....	x
Abstract.....	xv
Chapter 1: Introduction.....	1
Outline of the Thesis.....	4
Chapter 2: Biological Background.....	5
AIP.....	5
Visual versus Motor Cell Responses.....	7
Object Specific Activity .....	9
Size Selectivity.....	11
Position Selectivity .....	11
Temporal Characteristics.....	11
Summary of AIP Responses Properties.....	16
Inferior Premotor Cortex.....	16
Summary of F5 Responses.....	22
Chapter 3: Computational Building Blocks .....	24
Schema Theory.....	24
Opposition Space and Virtual Fingers.....	25
Chapter 4: A Region-Level View of the Reaching and Grasping Process .....	28
The Reaching/Grasping Task .....	28
Model Goals.....	29
Development of the Reaching/Grasping Model.....	30
Visual Processing (The How, What and Where of Objects in Space) .....	30
Interaction of the How/What Pathways.....	32
Adaptation of the Visual/Motor Mapping.....	34
High-Level Model of Visual Processing .....	35
Cortical Mechanisms for Object and Grasp Representation.....	37
The Dorsal System in Monkey.....	37
The Monkey Ventral System.....	39
High-Level Representation of the Grasping Program .....	39
Transformation from a Visual Representation to a Grasp .....	40
Expanded Model .....	41
Execution of the Grasp.....	43
Cortical Mechanisms of Low-Level Grasp Control .....	44
Model of Grasp Control.....	46

Interaction of AIP and F5 During Execution of the Grasp.....	48
Remaining Computational Issues .....	50
Object/Grasp Mapping is Many-to-Many .....	51
Modulation of the Reach Program by the Grasp Program .....	53
Parameter Coding .....	53
The Full Region-Level Model .....	54
Further Biological Background.....	54
The Complete Model.....	56
Progression of Activity During Task.....	57
Summary .....	59
Chapter 5: Low Level Details of the Grasping Model .....	61
$\mu$ -schemas: An Intermediate Computational Mechanism .....	61
The Columnar Structure.....	63
Column Dynamics.....	65
Connectivity of Regions.....	68
Object Representations in PIP and IT .....	68
AIP.....	71
Classes of Neurons in AIP .....	71
Projections from PIP.....	71
Associations from IT.....	73
Schema Cooperation/Competition Via the Basal Ganglia .....	74
F5 .....	75
Representation of Grasp .....	75
Schema Priming by F6 and Cooperation/Competition Via the BG.....	76
Task Specific Biasing of the Selected Grasp.....	78
Interaction between F5 and AIP.....	78
Phasing Via BG .....	81
Interaction of the AIP Active Memory and the BG Phasing Systems .....	84
Grasp Execution.....	86
SI and SII.....	86
MI .....	87
MI Activation.....	89
Internal Timing Model of Current Hand Extension.....	90
Full Execution Circuit.....	91
Summary .....	93
Chapter 6: Model Simulation Results.....	97
Introduction .....	97
Overview of Model Behavior.....	98
Comparison to Observed Biological Data.....	101
F5 Phasic Responses .....	101
AIP Phasic Responses.....	101
AIP Visual/Motor Distinction .....	102
Population Analysis.....	109

Population Analysis of Phasic Behavior.....	110
Population Analysis of Visual/Motor Properties.....	112
Model Predictions .....	115
Prediction #1: Object Coding in AIP .....	115
Population Analysis of Object-Related Responses.....	118
Prediction #2: When an Object Affords Multiple Grasps.....	121
Population Analysis Comparing Conditional and Non-conditional Tasks.....	129
Prediction #3: Object Size Coding.....	134
Prediction #4: Object Size Perturbation.....	139
Prediction #5: Object Parameter Coding.....	145
Object versus Grasp Coding in the Monkey AIP .....	148
Summary .....	150
Chapter 7: Linking PET Imaging in Humans to the Grasping Model .....	153
Synthetic PET Imaging.....	153
Computation of the Synthetic PET Measure.....	154
Synthetic PET Imaging for the Grasping Model .....	155
Conditional Task .....	156
Precision versus Side Opposition.....	157
Human Grasping Experiment .....	160
Materials and Methods.....	160
Subjects.....	160
Tasks.....	160
Performance Measures .....	161
Imaging.....	162
Image Analysis.....	162
Statistical Tests.....	162
Results.....	163
Behavioral Results.....	163
Imaging Studies .....	163
Comparison of PET and synthetic PET Results.....	167
Precision versus Side Opposition/Power Grasp.....	167
Conditional vs Non-Conditional Task.....	167
Summary .....	169
Chapter 8: Grasp Configuration Learning.....	172
Introduction .....	172
Neural Dynamics.....	173
Learning Dynamics.....	176
Reinforcement Prediction.....	179
Experimental Results .....	181
Summary .....	183

Chapter 9: Conclusions and Future Work .....	185
PET Imaging.....	187
Grasp Configuration Learning.....	188
Related and Future Work .....	188
Control of Wrist Orientation .....	188
Planning of Arm Movements.....	189
Coordination of Reach and Grasp.....	189
Learning of Motor Programs .....	190
Encoding of Grasp Programs in F5 .....	191
Perception of Grasp .....	191
Appendix A: Model Implementation Details.....	193
Network Structure Description .....	193
Number and Distribution of Neurons .....	195
Network Wiring Rules .....	198
PIP -> AIP .....	198
IT -> AIP .....	198
AIP -> AIP.....	199
AIP -> F5.....	199
F5 -> AIP.....	200
F5 -> F5 .....	200
F5 -> Mcx.....	200
F5 -> SII.....	201
SII -> F5 .....	201
SI -> SII.....	202
SI -> Mcx .....	202
F5 <-> Area 46.....	202
AIP <-> BG.....	202
F5 <-> BG.....	202
F6 -> F5/AIP.....	203
F6/F2 -> BG.....	203
Protocols.....	203
Vanilla Protocol.....	203
Grasping in the Dark.....	203
Object Fixation.....	204
1 Object/2 Grasps (conditional case).....	204
1 Object/2 Grasps (non-conditional case).....	204
References.....	205
Index .....	212

## List of Tables

Table 5.1: Description of AIP cell behavior .....	72
Table 5.2: F5 grasp descriptor.....	76
Table 7.1: Significant regions of activity ( $p < 0.001$ ): conditional - non-conditional (combined precision and power) .....	166
Table 7.2: Significant regions of activity ( $p < 0.005$ ) : precision pinch - power grasp.....	166
Table A.1: Description of grasp.....	194
Table A.2: Description of objects.....	194
Table A.3: Number of cells in each model layer .....	196
Table A.4: Distribution of F5 and AIP cells for two experiments.....	197

## List of Figures

Figure 2.1: Cortical regions in the macaque involved in reaching and grasping.....	6
Figure 2.2: The Sakata protocol.....	6
Figure 2.3: Behavior of a motor-type cell under different conditions.....	8
Figure 2.4: Cell that is active during lighted conditions.....	8
Figure 2.5: Visual/Motor cell.....	9
Figure 2.6: Dark enhancement cell.....	9
Figure 2.7: Object-specific activity.....	10
Figure 2.8: A cell that is selective for 'flat' objects.....	11
Figure 2.9: Selective toward small objects, somewhat independent of object type.....	12
Figure 2.10: Activity independent of object size.....	12
Figure 2.11: Modulation of activity as object position changes.....	12
Figure 2.12: Small increase during key phase, with high activity during movement and hold phases.....	13
Figure 2.13: Same neuron as in Figure 2.12, except during movement in the dark condition.....	13
Figure 2.14: Ramp up of activity during key phase.....	13
Figure 2.15: Late onset of movement-related activity.....	13
Figure 2.16: Reduction during key phase and longer extent of activity.....	14
Figure 2.17: Increase, then decrease of activity during key phase.....	14
Figure 2.18: Suppression during key phase. Active during movement, hold, and release phases.....	15
Figure 2.19: Primarily a movement-related neuron.....	15
Figure 2.20: Visual and somatosensory receptive fields of an F4 neuron.....	17
Figure 2.21: Response of a single F4 neuron modulated by the region of peripersonal space that the monkey is reaching toward.....	18
Figure 2.22: An F5 neuron that codes for precision grip.....	19
Figure 2.23: F5 neuron responsive to precision grip.....	22
Figure 2.24: Another F5 cell that is responsive to the precision grip.....	22
Figure 3.1: The three primary oppositions and their corresponding virtual fingers.....	26
Figure 4.1: Experimental setup for the Gentilucci grasp adaptation task.....	35
Figure 4.2: A high-level model of cortical visual processing and its support of the reach/grasp motor program.....	36
Figure 4.3: The expanded seeing/reaching/grasping model.....	42

Figure 4.4: Motor cortex cell whose firing correlated with cutaneous stimulation of the lateral side of the index.....	46
Figure 4.5: The cortical regions involved in the low-level control of the grasping motor program.....	47
Figure 4.6: Interaction between AIP and F5 populations (circles) during execution of the preshape/grasp/hold program.....	48
Figure 4.7: AIP computes the set of affordances for an attended object .....	52
Figure 4.8: The complete model.....	57
Figure 5.1: Column architecture.....	64
Figure 5.2: Convention used for drawing columns .....	68
Figure 5.3: Illustration of PIP coding. Every object class is represented by a population of units .....	69
Figure 5.4: Example coding of three objects (pencil, stick, and a ball) in PIP and IT.....	70
Figure 5.5: Affordance mapping from PIP to AIP .....	73
Figure 5.6: Affordance mapping from IT to AIP .....	74
Figure 5.7: Cooperation and competition in AIP .....	75
Figure 5.8: Priming of the general grasping schema, and selection of a specific grasp to execute.....	77
Figure 5.9: F6 and F2 biasing of the grasp schema selection process .....	79
Figure 5.10: Interaction between AIP and F5 during execution of the reaching/grasping/holding task.....	80
Figure 5.11: F5 interaction with BG implementing the cascade of activity in F5 as execution of the motor program progresses .....	81
Figure 5.12: BG phasing when an F5 unit is active during both the extension and flexion phases of the reach/grasp/hold program .....	83
Figure 5.13: Abstract representation of the interaction between the AIP active memory and the BG phasing loops .....	84
Figure 5.14: Representation of hand state information in the somatosensory cortices (SI and SII).....	87
Figure 5.15: Three classes of MI columns.....	89
Figure 5.16: Internal model of grasp extension state.....	90
Figure 5.17: The low-level circuit involving F5 extension columns.....	92
Figure 5.18: Low-level circuit involving flexion columns.....	93
Figure 5.19: Implementation of the hold phase of the motor program .....	94
Figure 5.20: Low-level circuit for the release phase of movement .....	95
Figure 6.1: F5 activity during execution of a precision grasp.....	99
Figure 6.2: F5/AIP interaction during execution of the grasp.....	102

Figure 6.3: A set of F5 cells with a variety of temporal properties.....	103
Figure 6.4: Variety of AIP cells.....	104
Figure 6.5: A pure motor-related AIP cell.....	106
Figure 6.6: A pure visual AIP unit.....	107
Figure 6.7: A visually-modulated AIP unit.....	108
Figure 6.8: Population analysis of F5 phasic behavior during a precision grasp.....	111
Figure 6.9: Normalized AIP responses over the entire population.....	113
Figure 6.10: Visual/Motor coding in AIP.....	114
Figure 6.11: Two objects that map to the identical grasp.....	116
Figure 6.12: Two visual AIP cells that show significant object-specific modulation.....	116
Figure 6.13: Two visual-related (with some motor) AIP cells.....	116
Figure 6.14: Two motor-related (set phase) AIP cells.....	117
Figure 6.15: Two motor-related AIP cells that are not modulated by the type of object.....	118
Figure 6.16: Comparison of population responses towards two different objects (but identical grasps).....	119
Figure 6.17: Visual-related AIP receive object-specific inputs; motor-related cells receive recurrent inputs from F5.....	120
Figure 6.18: A single object mapping to two possible grasps.....	121
Figure 6.19: Two F5 units in response to the four conditions.....	122
Figure 6.20: A pair of set/extension cells in F5.....	123
Figure 6.21: A pair of extension-related F5 cells.....	124
Figure 6.22: F5 flexion and hold units.....	125
Figure 6.23: A pair of release phase F5 units.....	126
Figure 6.24: A pair of set-related F5 units.....	127
Figure 6.25: A visual/motor AIP cell (more motor-related).....	128
Figure 6.26: A visual/motor AIP cell (primarily visual).....	129
Figure 6.27: A primarily motor-related AIP cell.....	130
Figure 6.28: A pure motor-related AIP cell.....	131
Figure 6.29: Comparison of set-related F5 responses towards two different grasps.....	132
Figure 6.30: Comparison of set-related AIP responses towards two different grasps.....	133
Figure 6.31: Cylinders of different widths map to a precision grasp of varying aperture size.....	134
Figure 6.32: Thumb and index finger temporal behavior as a function of cylinder size.....	134
Figure 6.33: F5 cell responses during precision grasps of different apertures.....	135

Figure 6.34: Visual-related AIP cell that demonstrates object size specificity.....	136
Figure 6.35: Motor-related AIP cell.....	137
Figure 6.36: Visual-related AIP cell that demonstrates little modulation due to object size.....	138
Figure 6.37: Visual/motor AIP cell that is inhibited during the hold phase of movement for mid-range cylinders .....	138
Figure 6.38: F5 movement-related cell (A) and a hold-related (B) cell during the perturbation experiment.....	140
Figure 6.39: F5 cell responses (of cells with preferred aperture between 20 and 25mm only) during execution of the hold phase of a 20mm precision grasp.....	142
Figure 6.40: AIP visual cell that is active for all four conditions.....	142
Figure 6.41: AIP motor-related cell that is most active during the movement phase of 30mm aperture grasps .....	143
Figure 6.42: AIP visual/motor cell that is active in all conditions .....	143
Figure 6.43: AIP motor-related cell .....	144
Figure 6.44: AIP (primarily) visual cell .....	144
Figure 6.45: Four boxes of different dimensions.....	145
Figure 6.46: Comparison of AIP visual responses .....	146
Figure 6.47: F5 population responses for two objects of the same width.....	147
Figure 6.48: Response of a single monkey AIP cell to four objects.....	149
Figure 6.49: Response of a primarily visual-related AIP cell.....	150
Figure 6.50: Response of a visual/motor AIP cell.....	151
Figure 7.1: Neural elements involved in computing the synaptic activity for a simulated brain region.....	155
Figure 7.2: Functional interactions of F2, F5, and AIP .....	157
Figure 7.3: Predictions of relative synaptic activity changes.....	158
Figure 7.4: Positive and negative synapse contributions to the synaptic activity measure .....	159
Figure 7.5: Apparatus used in PET experiment.....	161
Figure 7.6: Reaction times for both the release of the switch and the depression of the next switch .....	165
Figure 7.7: Updated functional model.....	168
Figure 7.8: Significant activity ( $p < 0.005$ ) for conditional - non-conditional tasks .....	170
Figure 7.9: Significant activity ( $p < 0.01$ ) for precision pinch - power grasp.....	171
Figure 8.1: Schematic view of the architecture for the grasp configuration learning model .....	174
Figure 8.2: Responses of several feature detector units.....	182

Figure 8.3: Responses of four additional feature detector units.....182

## Abstract

The act of reaching out, grasping, and manipulating an object involves the integration of information from a variety of sources--from vision of the object of interest, to task requirements, to tactile and proprioceptive information as the grasp is executed. In this thesis, we investigate the cortical mechanisms involved in 1) the translation of a visual description of an object and a representation of the task into an appropriate hand configuration, and 2) the unfolding of this description in time in order to execute the preshape, enclose, grasp, and ungrasp phases of movement. On the basis of behavioral, cell recording, and anatomical data from human and monkey, a computational model of the grasping process is proposed. This model focuses on the roles of the intra-parietal areas (AIP, PIP, and VIP), inferior premotor cortex (F4 and F5), pre-SMA (F6), frontal cortex (area 46), inferiotemporal cortex (IT), and the secondary somatosensory cortex (SII).

In the model, AIP serves a dual role of first computing a *set of affordances* that are appropriate for the object being attended, and then maintaining an *active memory* of the single affordance as the corresponding grasp is executed. F5 integrates visual, task and memory information in order to select one of the several possible grasps. This brain region then drives the high-level execution of the grasp and monitors its progression. Based upon the hypothesized computational roles of AIP and F5, the model make several key predictions about the encoding of grasp and object information at both the single unit and population levels.

In addition, we present results of a PET (positron emission tomography) study that 1) compares brain activity during the performance of the precision and power grasps, and 2) examines the brain regions responsible for processing an abstract instruction stimulus. Through a

technique referred to as Synthetic PET Imaging, we are able to compare the human PET results to the global behavior of the model. We show that this technique can also be used to further constrain the model structure.

## Chapter 1: Introduction

The act of reaching out, grasping, and manipulating an object involves the integration of information from a variety of sources (MacKenzie & Iberall, 1994). From the information available just prior to the start of movement, the monkey (or human) must be able to select a general grasping strategy. Visual information allows the monkey to extract the identity of an object as well as an estimate of its shape, size and localization in space. This information, combined with the behavioral context in which the monkey finds himself affects the choice of grasp that is to be made. The behavioral context can include task constraints (such as the goal of manipulation after the grasp is established), the position of obstacles relative to the target object, instruction stimuli provided by an experimenter or teacher, or memories of recently executed grasps. This plan is then refined depending upon the details of the situation, some of which may become available only as execution progresses. As contact is made with the target object, tactile and kinesthetic information yields an even more detailed description of the shape and locality of the object. This information is used by the monkey to further fine tune his arm and hand movements during the manipulation phase of the motor task.

In this thesis, we investigate the cortical mechanisms involved in:

1. The translation of a visual description of an object and a representation of the task into an appropriate hand configuration.
2. The unfolding of this description in time in driving execution of the preshape, enclose, grasp, and ungrasp phases of movement.

Two cortical regions have been studied in monkey that are particularly compelling in their neural responses during the execution of grasp. These are the rostral inferior premotor cortex (referred to as F5), and the anterior intra-parietal cortex (AIP). Cells in F5 often demonstrate selectivity towards the type of grasp that is made by the monkey (Rizzolatti, Camarda, Fogassi, Gentilucci, Luppino, & Matelli, 1988). In addition, cells fire in relation to specific phases of the executed preshape and grasping movement. Within AIP, cells demonstrate a variety of selective responses, including those towards the type of object presented to the monkey, the size of the object, and its position in space (Murata, Gallese, Kaseda, Kunimoto, & Sakata, 1993; Taira, Mine, Georgopoulos, Murata, & Sakata, 1990). What is striking about many of these cells is that their firing rate is also modulated during execution of grasping movements.

In this thesis, we propose a computational model of the grasping process, which is derived from behavioral, cell recording, and anatomical data from both human and monkey. This model focuses on the roles of the intra-parietal areas (specifically AIP, PIP, and VIP), inferior premotor cortex (F4 and F5), pre-SMA (F6), frontal cortex (area 46), inferiotemporal cortex (IT), and the secondary somatosensory cortex (SII).

In the development of the model, we hypothesize specific computational roles for several cortical regions:

1. AIP is responsible for extracting from the visual representation of an object the appropriate set of *affordances* for that object. By *extracting an affordance*, we mean that AIP highlights the properties of the object that are explicitly relevant for physically interacting with it in a particular way.
2. The continued activity of many AIP cells through the execution of the preshape, grasp, and holding of an object maintains an *active memory* of the grasp that is being planned or that is being executed. This active memory is important for the organization of the global aspects of the grasp during execution. After a secure grasp

is established, this memory can also be used to recalibrate the visual-motor mapping through an associative process between the visual representation of the object and the active memory of the grasp that was actually executed (as opposed to the grasp that was originally planned).

3. F5 is the site responsible for integrating the many constraints that are used to decide on the single grasp that is to be executed. These constraints include visual information (from the affordances extracted by AIP), task information, instruction stimuli, and working memories of recently executed grasps.
4. F5 is responsible for the high-level execution and subsequent monitoring of the planned preshape and grasp.
5. We explicitly address the issues involved in the neural-level encoding of object, grasp, and phasic information, and show how the populations of cells can capture the critical information that is required to perform the necessary computations.

The model makes several key predictions relating to:

1. The encoding of object- and grasp-specific information in AIP and F5 at both the single-cell and population levels.
2. The behavior and timing of the grasp decision process when an object affords multiple grasps.
3. The behavior and timing of neural responses when there is a discrepancy between the visually-programmed grasp and the actual size of the object.
4. The changes in regional synaptic activity as a function of the task that is performed (as measured by synthetic PET imaging).

Finally, we present results of a PET (positron emission tomography) study in human that:

1. Compares brain activity during the performance of a precision pinch and a power grasp.

2. Examines the brain regions involved in the processing of an abstract stimulus that instructs the subject as to which of the two grasps must be performed.

The results of this study are related to our model through the synthetic PET imaging technique.

## **Outline of the Thesis**

Chapter 2 will focus on the key aspects of the monkey neurophysiological data from which we derive our model. The computational background of this work (schema theory and opposition spaces) are examined in Chapter 3.

The core model described in this thesis is developed in Chapter 4; the low level details of the model are described in Chapter 5; Chapter 6 is dedicated to the demonstration and validation of the model behavior, as well as the testing of the model with several novel tasks.

Chapter 7 focuses on relating the model (derived primarily from monkey data) to human PET experiments. We first describe the synthetic PET imaging technique, through which it is possible to make predictions from the model about the what we can expect in human PET experiments. We then present a set of real human PET results, and show how they may be used to further constrain our modeling work.

Chapter 8 steps back from the primary model, and asks how the grasping motor programs that are represented in AIP and F5 might be acquired automatically through a trial-and-error learning procedure. A model is presented which learns (via reinforcement learning) a mapping from object and task parameters to a description of a grasp that is appropriate for that object.

Finally, Chapter 9 summarizes the key aspects of the presented models, and concludes with a discussion of future directions for model development.

## Chapter 2: Biological Background

---

In this chapter, we present a detailed analysis of several key regions in monkey that are involved in the planning and execution of the grasping program. Neurons in the anterior intra-parietal area (AIP) are selective to a variety of object properties (including shape, size, and orientation in space), but also show significant movement-related modulation as the grasp is executed. The rostral region of the inferior premotor cortex (F5) contains cells that fire when the monkey makes specific grasping movements with his hand and/or mouth. Cells in the caudal part of the inferior premotor cortex (F4) appear to encode the final endpoint position of reaching movements.

---

### **AIP**

One area of concentration of the Sakata group has been the posterior bank of the anterior intra-parietal sulcus (Murata, et al., 1993; Sakata, Taira, Murata, & Mine, 1995; Taira, et al., 1990). This region (referred to as AIP) receives input from other areas of the posterior parietal cortex (see Figure 2.1), which provide information about the parameters of the object of interest, including location, orientation, shape, and size (Sakata, 1994). In addition, this region has very significant recurrent cortico-cortical projections with area F5 of the inferior premotor cortex (Matelli, 1994; Sakata, 1994).

Figure 2.2 demonstrates the protocol used by Sakata in these recording studies. The ready signal is given by the turning on of light 1 (L1 in the figure). The monkey responds by placing

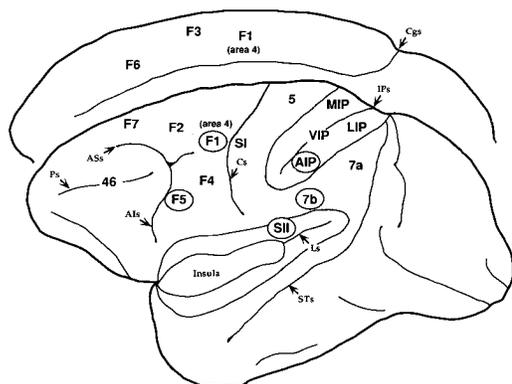


Figure 2.1: Cortical regions in the macaque involved in reaching and grasping. Figure adapted from (Jeannerod, Arbib, Rizzolatti, & Sakata, 1995).

#### HAND MOVEMENT

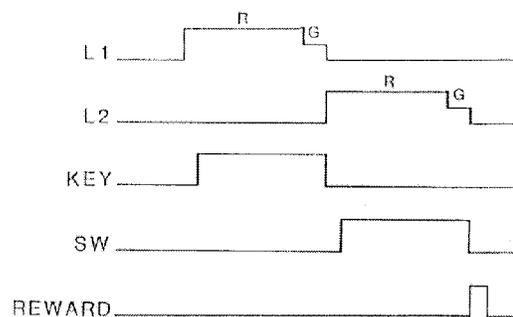


Figure 2.2: The Sakata protocol, as described in (Taira, et al., 1990). Figure from (Sakata, 1991).

his hand upon the touch pad that is located directly in front of him. Contact with the touch pad is indicated by the KEY trace. When L1 changes color (the trigger stimulus), the monkey begins the execution of the reaching movement toward the object that is located in front of him. A second light (L2) turns on as the monkey pulls away from the touch pad. The monkey grasps the object and manipulates it by pulling or pushing in the appropriate direction, as indicated by SW (also referred to as object phase). This position is held until L2 changes color (the secondary trigger stimulus), at which time the monkey releases the object and moves his hand away. If the task is performed properly, the monkey is then rewarded with a squirt of juice.

In the experiment, a variety of objects were presented to the monkey. For each object, four different conditions were used: movement in the light (to manipulate the object), movement in the dark <sup>1</sup>, fixation of the object in the light, and fixation in the dark. The first set objects included buttons that had to be pushed, knobs to be pulled, and a joystick that had to be grasped and pushed in a particular direction (Taira, et al., 1990). A second set of objects that have been

<sup>1</sup> At the time dark trials are performed, the monkey has already seen the object and manipulated it in the light.

studied include various sizes of plates, cones, cubes, cylinders, spheres, and rings (Murata, et al., 1993).

Muscimol-induced lesions of AIP lead to a significant deficit in the monkey's ability to grasp objects (Sakata, 1994). The grasping movements are clumsy and uncoordinated, and as a result the monkey is unable to shape his hand or orient his wrist appropriately for objects that are presented to him. However, the monkey is still able to execute the basic sequence of the task: wait for the go signal, move the hand to the object, and then return from the object when the secondary go signal is given.

### **Visual versus Motor Cell Responses**

Figure 2.3 shows a cell that is active primarily during conditions where the monkey makes a movement toward the object (as compared to fixation of the object). For this reason, Sakata classifies this cell as a motor-related cell. This behavior contrasts with the visual-related cell shown in Figure 2.4. Here, the cell is only active during conditions in which the monkey is able to see the object. Note, however, that the cell response in the light movement condition is modulated by the phase of movement. In addition, during this condition, we see an increase in activity at the beginning of the key phase of the task, then a suppression before the trigger stimulus. (Sakata, 1994) attributes the fluctuation during the key phase to the fact that the monkey does not continuously fixate on the object during this phase. This effect was better controlled during the second set of experiments, in which L1 and the object were superimposed upon one-another. During the fixation condition, the activity level of the cell (Figure 2.4) remains relatively constant over the entire presentation of the visual stimulus, indicating that the monkey continued to fixate on the object during the key phase.

Many cells exhibit both visual and motor properties. This class appears to represent a continuum between the visual and motor response types. In Figure 2.5, the cell responds more heavily to the motor condition, but shows significant activity during the visual fixation case.

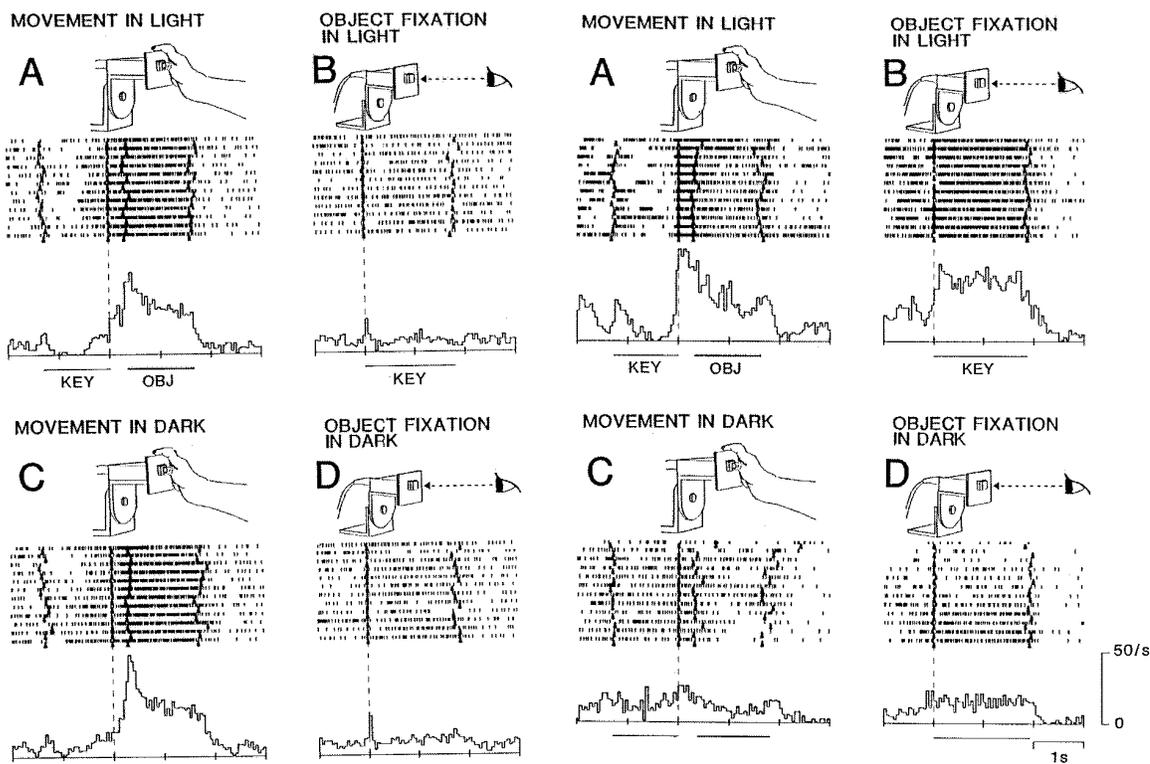


Figure 2.3: Behavior of a motor-type cell under different conditions: light / dark and fixation/manipulation. *Key* indicates the period during which the monkey has his hand on the touch pad. *Obj* indicates manipulation of the object. From (Sakata, 1991).

Figure 2.4: Cell that is active during lighted conditions. From (Sakata, 1991).

Within the data presented by Sakata, there are also cells that respond more favorably to the visual condition.

The cell shown in Figure 2.6 demonstrates an enhancement in activity for the movement in the dark condition (as compared to movement in the light). The cell is apparently inhibited (either directly or indirectly) by some visual features belonging to the joystick. This illustrates an important point about distributed representations of motor programs: the behavior of individual cells does not necessarily correspond in a logical manner to the computation that is performed by the network as a whole. However, what is important is that for any particular

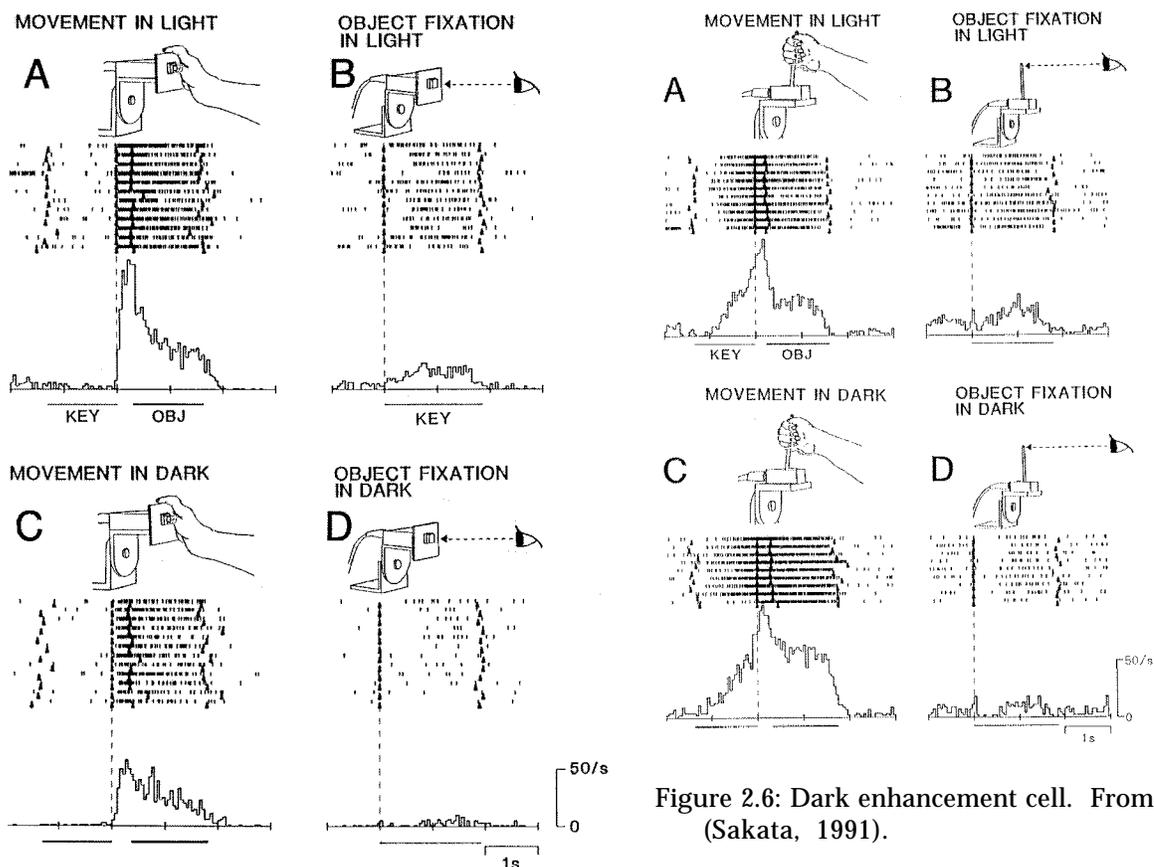


Figure 2.6: Dark enhancement cell. From (Sakata, 1991).

Figure 2.5: Visual/Motor cell. From (Sakata, 1991).

situation, enough cells agree about the output that is to be generated. Then, through a process of cooperation and competition, the correct output is generated while conflicting outputs are shunted.

### Object Specific Activity

Many cells in this area of the posterior parietal cortex also demonstrate object-specific activity. However, it is important to note that some cells actually respond to several different objects, with varied activity levels. The cell depicted in Figure 2.7 responds heavily to a pull-knob that is recessed into a box, and partially responds to a normal pull-knob and a push-button. In addition, Figure 2.8 demonstrates a cell that has been labeled as an *angular selective cell* by

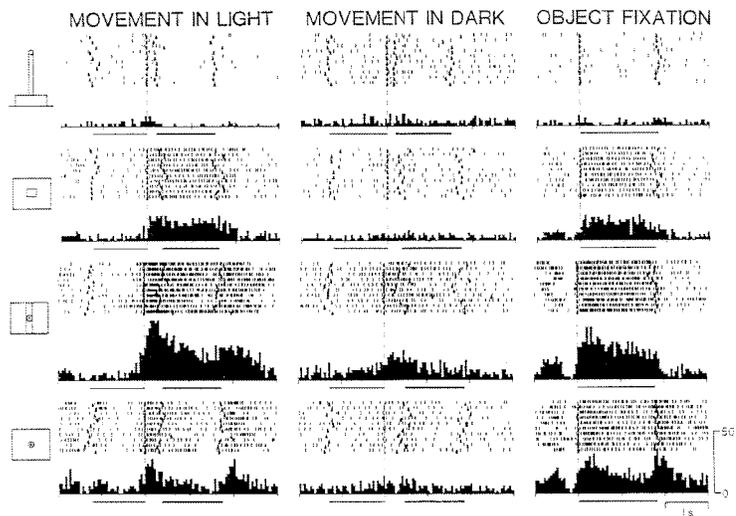


Figure 2.7: Object-specific activity From (Sakata, 1994).

Sakata (Sakata, 1994). Note that in all cases where the cell is active, a side opposition is involved in the grasp of the object.

In general we hypothesize that the more similar two objects are, the more their representations will overlap. However, the term *similar* is a difficult one. In this particular area, similar can mean a couple of things: similar in visual pattern, or similar in the required hand configuration. Most likely we will see both types of constraints reflected: we want to classify objects as easily as possible (i.e. using features that are easy to discriminate), but we must separate objects that require different motor programs. From this, we would expect two objects that are very distinct visually and require different hand configurations to use different sets of cells in AIP, and two objects that are very similar and require the same motor program to use a large number of overlapping cells. But how does the system deal with convergent and divergent mappings? At this time, the data does not give us a clear view due to the fact that very little has been done to compare similar objects. These issues will be addressed from the perspective of the model in Chapter 6.

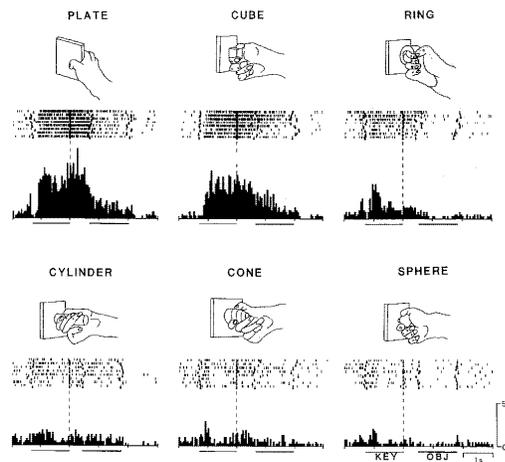


Figure 2.8: A cell that is selective for 'flat' objects. From (Sakata, 1994).

### Size Selectivity

A number of AIP cells demonstrated selectivity toward the size of the presented object. Figure 2.9 shows one such cell, which is selective for small objects, but shows some independence in regards to the type of object. On the other hand, some cells demonstrated independence from the size of the object (Figure 2.10).

### Position Selectivity

A very small number of cells exhibit activity modulation as a function of the position or orientation of the object (e.g. Figure 2.11). This information may be made available to F4 as a mechanism for modulating the position of the wrist as a function of the grasp type that is used.

### Temporal Characteristics

In addition to the various selectivity classes, the neurons in AIP also show striking temporal behavior that is related to the ongoing movement. The identifiable phases are: set (*key phase* in the figures), preshape, enclose, hold (*obj phase*), and ungrasp.

Figure 2.12 depicts a cell that shows a small increase of activity slightly before the key phase and maintains the same level through the duration of this phase. This indicates that

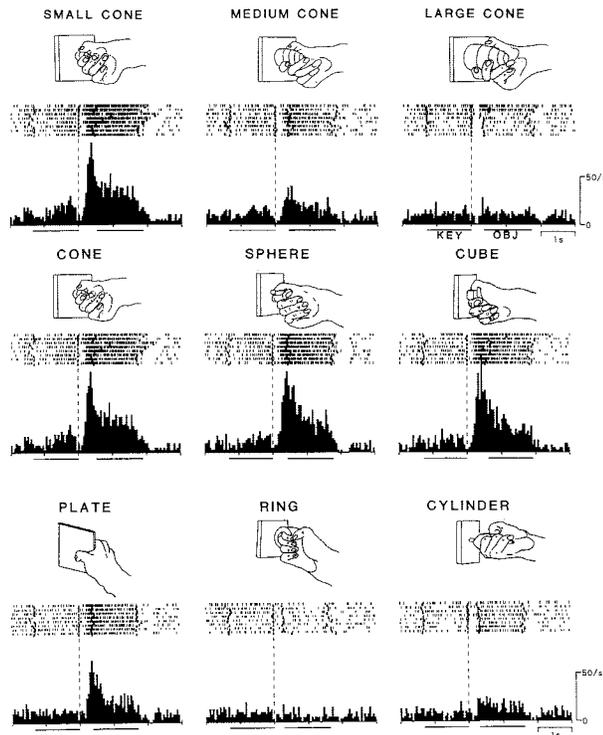


Figure 2.9: Selective toward small objects, somewhat independent of object type. From (Sakata, 1994).

the monkey is fixating on the object before his hand reaches the touch pad. Compare this behavior to the cell's activity during movement in the dark (Figure 2.13). In this case, we only see a small increase in activity at the end of the key phase. Note also that this cell shows very high activity during the preshape and enclose phases and moderate activity during the hold phase of the motor program.

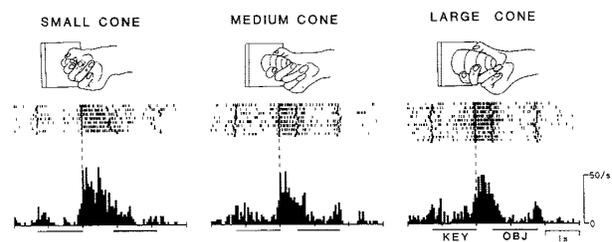


Figure 2.10: Activity independent of object size. From (Sakata, 1994).

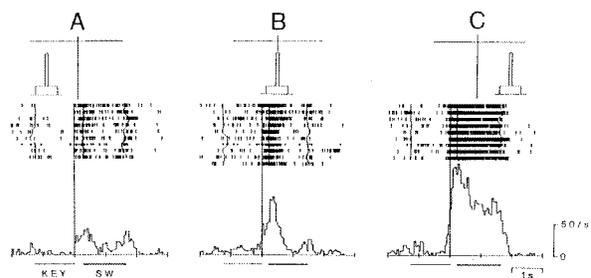


Figure 2.11: Modulation of activity as object position changes. From (Sakata, 1991).

MOVEMENT IN LIGHT

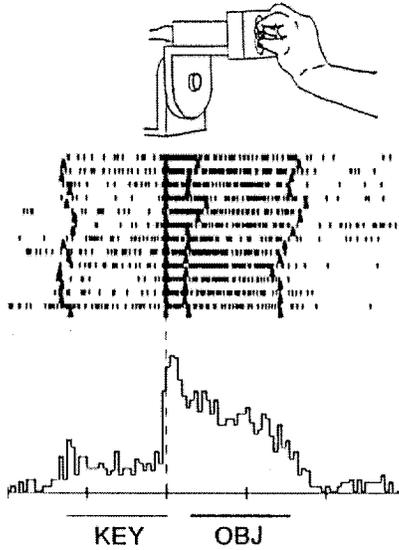


Figure 2.12: Small increase during key phase, with high activity during movement and hold phases. From (Sakata, 1991).

MOVEMENT IN DARK

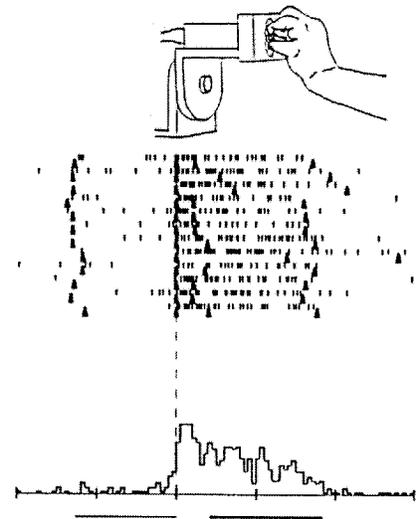


Figure 2.13: Same neuron as in Figure 2.12, except during movement in the dark condition. From (Sakata, 1991).

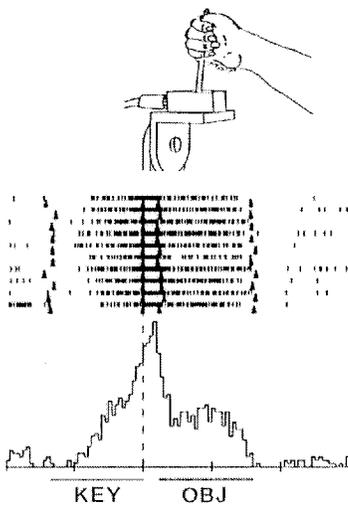


Figure 2.14: Ramp up of activity during key phase. From (Sakata, 1991).

MOVEMENT IN DARK

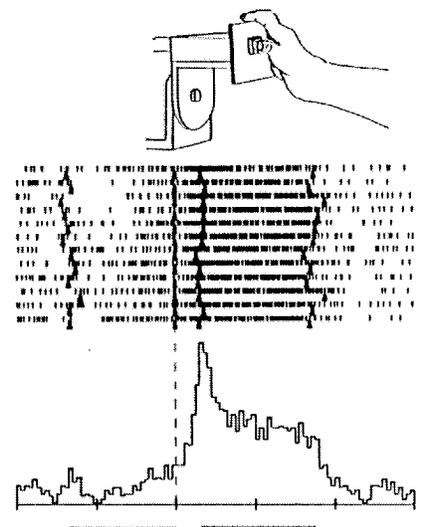


Figure 2.15: Late onset of movement-related activity. From (Sakata, 1991).

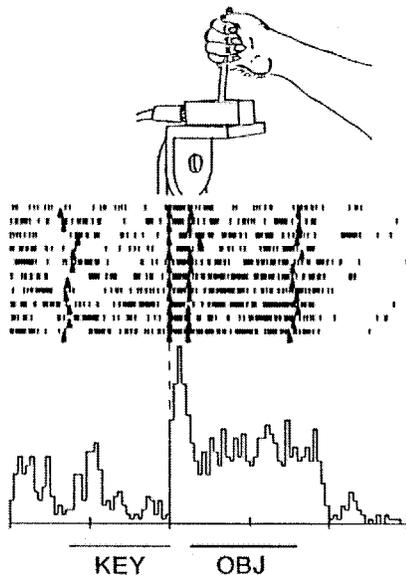


Figure 2.16: Reduction during key phase and longer extent of activity. From (Sakata, 1991).

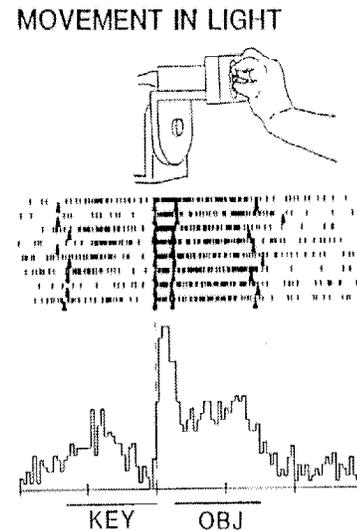


Figure 2.17: Increase, then decrease of activity during key phase. From (Sakata, 1991).

Figure 2.14 demonstrates a significant difference in reaction to the visual stimulus. Shortly into the key phase, we see a rapid increase in cell activity, peaking during the movement. This type of cell is most likely to be directly involved in configuring the motor set for the coming reach and preshape. As in the previous cell, there is also substantial activity during the hold phase of the motor program, ending when the handle is released.

Figure 2.15 depicts a cell that shows a very late onset of movement-related activity. This activity is most likely related to the enclose phase of the grasp.

The cell in Figure 2.16 exhibits a similar behavior to that shown in Figure 2.4, where the activity starts at a relatively high level, but during the key phase, this activity is suppressed until the trigger stimulus. Also note that after the object is released, this cell continues to fire at a significant level, a behavior that is unlike most other cells. This activity indicates possible participation in the ungrasp phase of movement. Similarly, the cell in Figure 2.17

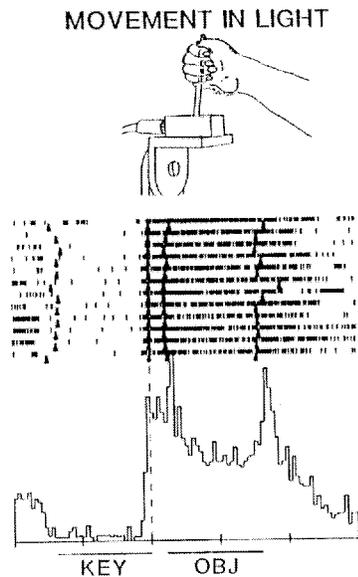


Figure 2.18: Suppression during key phase. Active during movement, hold, and release phases. From (Sakata, 1991).

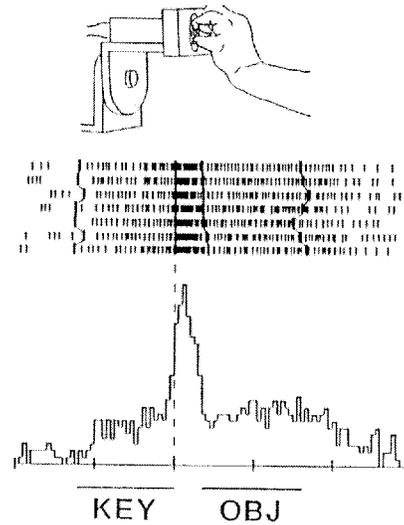


Figure 2.19: Primarily a movement-related neuron. From (Sakata, 1991).

demonstrates an initial increase in activity at the onset of the ready signal, which decreases to almost zero activity at the end of the key phase, and then shows a significant increase when the movement begins. Unlike the previous cell, the activity level drops off as the monkey releases the object.

The cell of Figure 2.18 shows significant activity as the object is being released, which tapers off slowly. This compares with many cells that have very abrupt drops in activity when the object is released. The second activity peak is possibly due to the fact that the monkey occasionally regrips the object after it is released (Sakata, 1994).

Finally, Figure 2.19 demonstrates activity that is related primarily to preshape and enclose, and with only slight activation during key, hold, and return phases.

## **Summary of AIP Responses Properties**

To summarize, the key features of the AIP responses are:

- A few cells (21% in one study (Taira, et al., 1990)) respond to simply fixating an object (visual-related), others (37%) are active only when a movement is being made to manipulate the object (motor-related). However, many cells (37%) fall somewhere between these two extremes.
- Many cells within this area are selective for the type of object that is presented. Most often, cells are actually selective to varying degrees for several objects. Typically, these objects share some common characteristics.
- Some cells demonstrate specificity toward the size of the object to be grasped, but show a certain degree of independence from the type of object.
- A small number of cells show modulation based upon the object's position and/or orientation in space.
- Most cells demonstrate phasic activity related to the motor behavior. The identifiable phases include: the "key" (set activity), preshape, enclose, and hold. Cells participate in varying degrees during different phases of the movement, but are usually most highly active during the preshape and enclose phases of movement. Very importantly, once a cell becomes active, it typically remains active until the object is released.

## **Inferior Premotor Cortex**

The macaque inferior premotor cortex (located ventral from the spur of the arcuate sulcus) has been identified as being involved in reaching and grasping movements (Rizzolatti, 1987). This region has been further partitioned into two sub-regions: F5 (rostral region, located along the arcuate) and F4 (caudal region). The neurons in F4 appear to be primarily involved in the

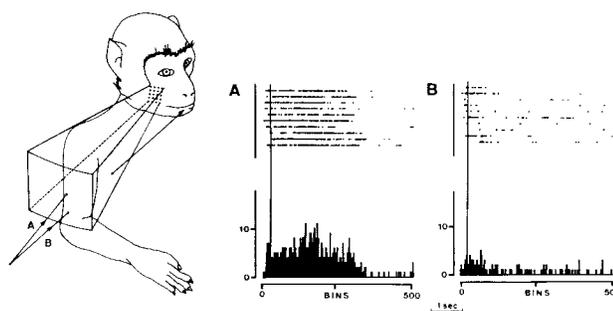


Figure 2.20: Visual and somatosensory receptive fields of an F4 neuron. The two traces demonstrate the neuron's responses to an object moving along two separate trajectories that share the same starting location. The neuron is also responsive to tactile stimulation on the marked area of the face. From fig. 10 of (Gentilucci, et al., 1988).

control of proximal movements (Gentilucci, Fogassi, Luppino, Matelli, Camarda, & Rizzolatti, 1988), whereas the neurons of F5 are involved in distal control (Rizzolatti, et al., 1988).

Many F4 neurons demonstrate responsiveness to both visual and somatosensory stimuli. The somatosensory map is topological and covers both the face and the upper body. Visual responses can be elicited by (among other things) the movement of objects toward the monkey. These visual responses are also topologically mapped according to the moving object's projected point of contact with the face or upper body of the monkey, as shown in Figure 2.20. It is important to note that these two topological maps coincide with one another. In the figure, the cell is responsive to trajectories that terminate on the marked area of the monkey's face (such as trajectory A). This is precisely the same region in which the cell is responsive to tactile stimuli. Even though trajectory B initiates at the same location as A, its point of contact with the monkey's face is significantly different and thus the cell does not respond.

In addition to the passive properties of these F4 neurons, recordings were also taken during the active generation of arm and hand movements. Three primary groups of neurons were observed: movement-towards-the-mouth, movement-towards-an-object, and facial movement. The movement-towards-the-mouth neurons appeared to be involved in any movement that brought the monkey's hand toward the mouth, despite the location of the arm prior to

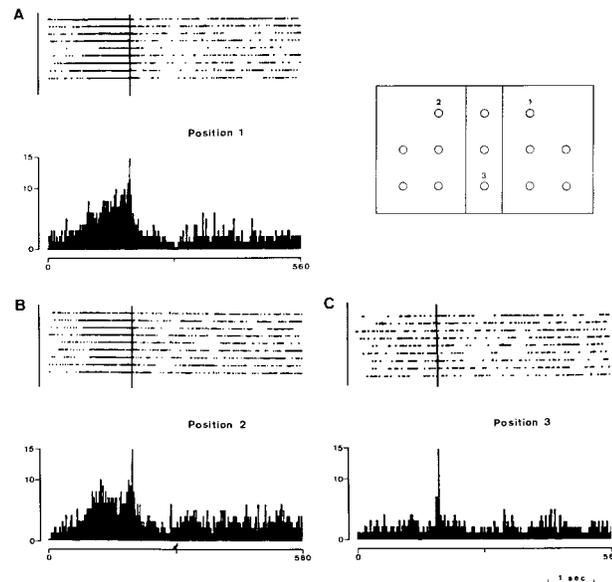


Figure 2.21: Response of a single F4 neuron modulated by the region of peripersonal space that the monkey is reaching toward. This particular cell is responsive to upper areas of the space. The histograms are aligned at the point where the hand made contact with the target object. From figure 11 of (Gentilucci, et al., 1988).

movement initiation. The movement-towards-an-object neurons, however, were selective for specific areas of the peripersonal space, as demonstrated in Figure 2.21. The receptive fields of these neurons contained very large areas of the peripersonal space. The active and passive response properties of individual neurons also show some congruence: neurons active during movement toward the upper region of space were sensitive to tactile stimulation on the face and lips. The sensory properties of these neurons (in response to visual and tactile stimuli) may be involved in either the elicitation of movements toward regions of the body in direct response to the various stimuli, or may actually be involved in the visual- and tactile-based guidance of the ongoing movement.

Recently, cells in the inferior premotor cortex (probably within F4) have been identified which have visual receptive fields in the region around the hand (Graziano, Yap, & Gross, 1994). These cells appear to encode position information of objects relative to the hand's current location. This information could be used to visually guide the hand towards the target object.

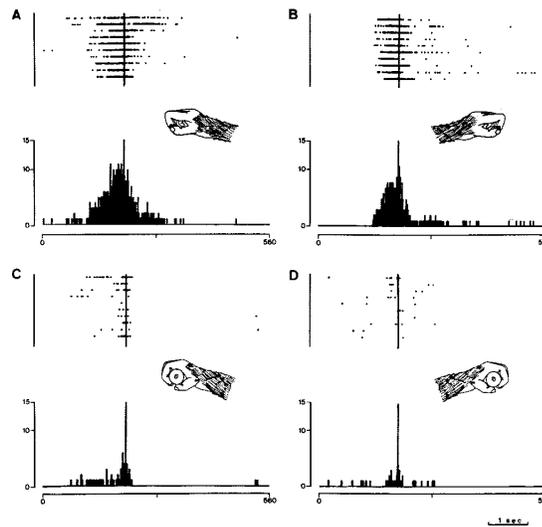


Figure 2.22: An F5 neuron that codes for precision grip in both the contralateral (A) and ipsilateral (B) hands. The neuron does not respond significantly to a palm-opposition type grasp. The histograms are centered at the point where the monkey made contact with the object. From figure 4 of (Rizzolatti, et al., 1988).

A functional distinction between F4 and dorsal premotor cortex (referred to as F2 by Rizzolatti) has also been made recently through the use of muscimol-induced lesions (Kurata & Hoffman, 1994). In a task in which the monkey had to pair an arbitrary instruction stimulus with the movement of a joystick (requiring an arm movement), a lesion in the arm area of F2 resulted in the monkey's inability to correctly match the correct movement with the stimulus. However, when the lesion was performed in the inferior premotor cortex, the direction of joystick movement was typically correct, but the executed movements were much weaker than normal.

Muscimol lesions of F5 result in a similar behavior as was observed in AIP: the monkey is only able to make clumsy grasping movements (Rizzolatti, 1995). One interesting difference, however, is that the deficit is most clear during execution of the precision grasp (as compared with lateral pinch or power grasp). Among other things, this effect might be due to the encoding of more primitive grasps in regions other than F5.

Individual neurons within F5 are responsive to somatosensory stimuli on the hands and mouth, as well as to visual stimuli (correlating with the size of the object <sup>2</sup>). During movement execution, many neurons code for the specific grip that is made by the monkey, but do not respond during axial or proximal movements made in the absence of distal movements. A neuron that codes for a precision grip made by either the contralateral or ipsilateral hand is shown in Figure 2.22. Other neurons were observed to be active during the formation of whole-hand grips (palm opposition) as well as lateral pinches (side opposition). Additional groups of neurons within F5 are responsive to other motor acts involving the hands. These include neurons that fire while the monkey is grasping with the hand and mouth, as well as neurons that are active while the monkey is tearing an object.

Activity of the F5 neurons correlated with the phase of the ongoing movement. The following phases were identified (Rizzolatti, et al., 1988): preparatory, finger extension, finger flexion, and holding. Cells that show activity prior to the onset of movement are labeled as preparatory cells. Note that these cells would probably have been considered *set cells* (Evarts, Shinoda, & Wise, 1984) if the protocol utilized an explicit a delay period between the viewing of the target object and the initiation of movement (we will use this term here).

Many cells are actually active during more than one phase of movement. The neuron shown in Figure 2.22 initially activates as movement is initiated, and continues to fire during the finger extension and finger flexion phases. Figure 2.23 demonstrates a cell that activates as the hand encloses on the target object, and terminates one second after contact is made. On the other hand, the neuron shown in Figure 2.24 begins to fire shortly before contact is made and continues to fire through the length of the grasp.

---

<sup>2</sup> In these experiments, the monkeys were only presented with two basic types of objects: bars and small raisin-size objects. Thus, it is not possible to distinguish object size from object or grasp type.

The activity patterns seen in F5 contrast significantly with those of the hand area of the primary motor cortex (Rizzolatti's F1) in two significant ways. First, many of the F5 neurons fire over several phases of the movement, even though each phase activates different sets of muscles. In fact, we see some cells that begin to fire before movement is observed. Second, a large number of the F5 neurons fire during movements of both the contralateral and ipsilateral hands, which would not be the case for a muscle-related coding scheme. What we conclude from this is that F5 is involved in some higher-level specification of the distal movement to be performed. However, this is not at such a high level that an individual neuron codes for the entire motor act.

Regarding general neural coding issues, individual neurons in F5 code for only a piece of the grasp that is being executed. By *piece of a grasp*, we mean that the representation of the grasp is distributed across a set of neurons, whose collective responsibility is to produce the appropriate grasp. This distributed coding within F5 is both in time (representing different phases) and in action space (e.g. different grasps). An important point about distributed representations is that individual neurons often take on multiple roles. We see this in the temporal domain in those neurons that are active during multiple phases of movement. This has also been seen in the action domain. Some neurons have been observed in F5 that are active during two distinct motor acts: grasping with the hand and grasping with the mouth. Also, if one looks carefully at Figure 2.22, one can see some activity during the whole-hand grasp of the contralateral hand (panel C), even though this cell has been labeled as being selective to the precision grip. The key point is that despite this not so discrete coding of actions at the neural level, the system is still able to produce discrete actions as a result of the collective behavior of the neurons. At the same time, because we observe discrete classes of behaviors, this does not necessarily imply that the discreteness should be reflected down at the level of the individual neuron.

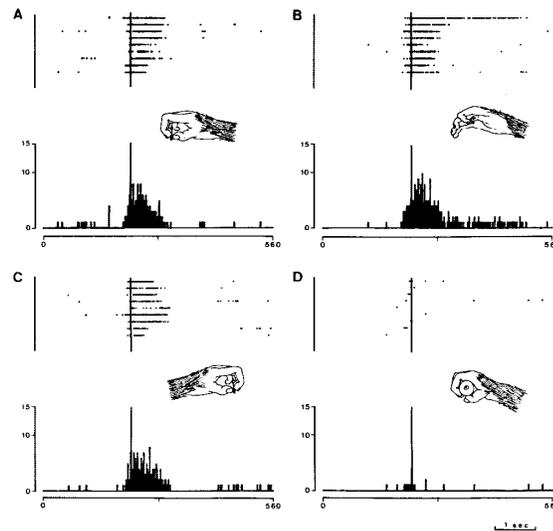


Figure 2.23: F5 neuron responsive to precision grip. Note that the activity of the neuron is not modulated by the orientation of the hand during the grasp. From figure 6 of (Rizzolatti, et al., 1988).

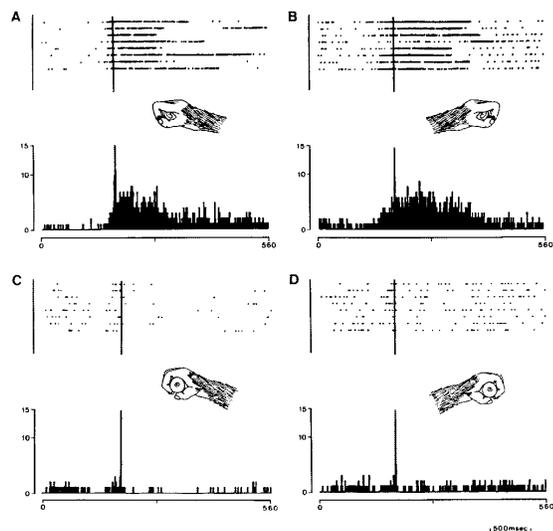


Figure 2.24: Another F5 cell that is responsive to the precision grip. This particular cell continues to fire for the duration of the grasp. From figure 7 of (Rizzolatti, et al., 1988).

## Summary of F5 Responses

The most important observations made of F5's role in the execution of grasping movements are summarized as follows:

- Cells tend to be most responsive for specific grasp types made by the monkey. Cells coding for grasps that are not being executed are sometime inhibited.
- The largest number of cells are dedicated to grasps involving the thumb opposing the fingers (pad opposition or precision grasp); a moderate number of cells is involved in the coding of side opposition; and very few cells are active during power grasps. In one experiment, 39% (56/142) of cells active during the grasp were observed to code a precision grip, 30% coded a finger prehension grasp (all fingers opposing the thumb), 4% coded for palm opposition, and 27% were non-specific (Rizzolatti, et al., 1988).
- Most cells are most responsive to a particular phase of movement. The identifiable phases are: set, extension, flexion, and hold. The largest number of cells are active during the extension and flexion phases of movement. In the same experiment, 34% (35/104) of grasp-related cells were active prior to distal movements, 35% during finger extension, and 32% were active during finger flexion. These grasping units constituted 66% of observed cells; 9% of cells were active while holding the object.

## Chapter 3: Computational Building Blocks

---

Schema theory and opposition space provide languages for describing computation processes and grasps, respectively. In this chapter, we outline their primary features.

---

### Schema Theory

Schema theory (Arbib, 1989) provides a language for describing functional decompositions of sensory and motor processes. An individual *schema* is a parameterizable description of an encapsulated computational element. Schemas are *active* in the sense that not only do they contain state information (data), but also the programmatic description for gathering and manipulating the data. In addition, schemas may be defined hierarchically (composed of a network of *sub-schemas*).

A *schema instance* is a parameterized copy of a *schema* that performs the specified computation based upon the schema description and the provided parameters. These active entities are created through a process of instantiation and activation. Instantiation, also referred to as *priming* in (Dakin, 1986), is performed by the parent schema and involves preparing the schema instance for execution by creating the actual copy of the structure and providing the instance with the appropriate parameters and input/output ports. *Activation* of the schema instance marks the point at which it begins to execute. This event is potentially contingent upon a triggering signal or the state of other schema instances.

This influence of activation based upon the state of other schema instances is what we refer to as schema coordination. Typically, these schemas have been instantiated by a common parent and must work together to perform the designated task (*schema cooperation*). Cooperation can mean one of two things. First of all, two schemas can execute simultaneously, solving different parts of the problem or blending their results together into a single common output. Second, two schemas may execute at different times, each handling separate temporal components of the computation.

Besides cooperation, two primed schemas may represent completely different solutions to a given problem, and a decision must be made as to which is most relevant to a particular situation. This resolution process (*competition*) may be left to the schemas themselves, to be decided based upon locally available state information.

From a biological stand-point, however, schema theory does not provide a sufficient language for mapping between schemas and neurons. Although we allow a schema to be implemented as a neural network and then connect it into a network of other schemas, this is typically done at a functional level and does not address how individual neurons might participate in several schemas in a graded fashion. What is missing is a bridge from the functional level of analysis to an implementational one. We will address this issue as we develop our model in Chapter 5.

## **Opposition Space and Virtual Fingers**

The problem of describing grasping movements has been approached by a number of researchers (Cutkosky, 1989; Napier, 1956). One such formalism, called *opposition space*, begins by identifying the primitive oppositions that can be applied by the primate hand (Iberall, Bingham, & Arbib, 1986; MacKenzie & Iberall, 1994). A primitive opposition consists of a pair of forces that are applied to the target object. The three oppositions available to the primate hand are:

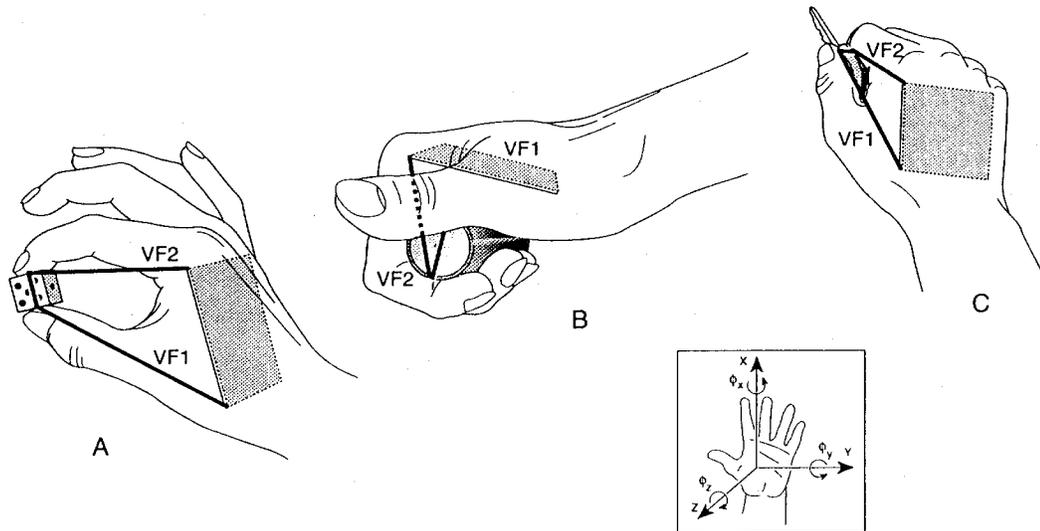


Figure 3.1: The three primary positions and their corresponding virtual fingers: pad position (A), palm position (B), and side position (C). From figure 2.7 of (MacKenzie & Iberall, 1994).

1. The *pad position* (Figure 3.1 A), in which the thumb is opposed by one or more fingers. A precision pinch (usually involving the index or middle finger opposing the thumb) is one example of a pad position.
2. The *palm position* (B), where the fingers oppose the palm (the thumb is not involved). A *power grasp* is one example.
3. The *side position* (C), in which the side of a finger is used to oppose the thumb or another finger. A *lateral pinch* involves the thumb opposing the side of the index finger.

The term *virtual finger* is used to describe (in abstract) the physical entity that is used in applying the force. Thus, each position involves the use of two virtual fingers (VF1 and VF2 in the figure). In the case of the pad position, the two virtual fingers are the thumb and any subset of real fingers. For the palm position, the two virtual fingers are the palm itself and a subset of the real fingers. A virtual finger can also counteract task-related forces or torques

without opposing a second virtual finger (this is referred to as VF3). This is the case for grasps that oppose gravity, such as when carrying a suitcase.

Given these notions of oppositions and virtual fingers, a grasp can then be described (in opposition space) as the combination of:

1. One or more oppositions.
2. The mapping of real to virtual fingers for each opposition.
3. The virtual finger state variables.

It is important to note that a grasp may consist of multiple oppositions (1). For example, a common grasp that is used with a screw driver is one in which a palm opposition (with the ring and little fingers opposing the palm) is used to grasp the handle, and a pad opposition (index finger opposing the thumb) stabilizes and guides the tip. The virtual finger state variables include the VF length, orientation, and width, as well as the orientation of the surface at which contact is made, a measure of the forces that can be applied by the VF, and the sensory inputs that are available for controlling grasp.

A key point regarding this description scheme is that it defines a hierarchy of grasp description. As we proceed from element #1 to #3, more and more details of the grasp are provided. This might suggest a way of thinking about the representation of grasp at a neural level. As we have already seen in Chapter 2, neurons in F5 have a tendency to align themselves with the type of grasp that is executed by the monkey. Further details of the grasp coding might be captured within sub-populations of these grasp-oriented cells (although we have not yet seen this in F5, this is suggested by the size-selective cells observed in AIP). In addition, the opposition space representation may also tell us something about how neurons might be shared between different grasp instantiations. Two grasps that share some components (e.g., an opposition or a virtual finger with a certain set of real fingers) could overlap in their neural representation, sharing neurons that are responsible for controlling precisely these common elements.

## Chapter 4: A Region-Level View of the Reaching and Grasping Process

---

This chapter outlines the computational and neurobiological constraints for our model of primate reaching and grasping. From these constraints, we outline a region-level view of the model, and show how the different brain regions interact to perform a grasping task.

Within the model, the dorsal and ventral visual streams first compute a representation of the object of interest. We hypothesize here that information from both streams converge on AIP, where the set of appropriate grasps (or affordances) is computed. These grasp-centered codes are then passed to area F5. We propose that it is this region that is responsible for first selecting from one of the possible grasps, and then unfolds this representation in time to execute the preshape/grasp motor program. During the execution process, AIP performs as an active memory by maintaining a representation of the grasp that is actually executed.

In the subsequent chapters, we will discuss the neural network implementation of the model, and then demonstrate its behavior in simulations of several experimental contexts.

---

### The Reaching/Grasping Task

The reaching and grasping tasks studied in this work are derived from the protocols used in the laboratories of Sakata and Rizzolatti (Rizzolatti, Camarda, Fogassi, Gentilucci, Luppino, & Matelli, 1988; Taira, Mine, Georgopoulos, Murata, & Sakata, 1990). In these tasks, the monkey sits directly in front of a touch pad or a switch that he must engage with his hand. An

object of some shape and dimension, selected by the experimenter, is positioned just within reach - directly in front of the monkey. The following outlines the sequence of events that take place during one such protocol (of Sakata):

Step 1: The monkey, sitting in the monkey chair, focuses his attention on an LED.

Step 2: A ready signal (turning on of the LED) indicates the start of a trial. The monkey responds by placing his hand on a touch pad and fixating on the LED, which will indicate when movement may begin. The LED is positioned (using a horizontal half-mirror configuration) such that it appears superimposed on top of the object to be grasped. The object is lighted in such a way that it is also visible.

Step 3: The LED changes color (the *GO signal*). The monkey responds by removing his hand from the touch pad and reaching towards the target object. During the reaching phase of movement, the monkey preshapes his hand in anticipation of making contact with the object.

Step 4: Contact with the object is made, and the monkey secures a grasp. A manipulatory movement is then made, such as pulling on the object or pushing it in some specific direction. The resulting position is held for a random delay period.

Step 5: The LED changes color again (the *secondary GO signal*). The monkey responds by releasing the object and moving his hand away.

If all steps of the task are performed properly, the monkey is rewarded (in this case with a drop of juice), after which the next trial is begun.

For the most recent set of experiments performed by Sakata, objects have been one of: cylinder, plate, torus, cone, sphere, and cube. The dimensions of these objects are also varied. Earlier experiments have made use of objects such as joysticks, buttons, and pull-knobs.

## **Model Goals**

Even the simple task presented here provides a number of interesting challenges that we approach in the model below. These include:

1. What mechanisms are involved in the transformation of a visual description of an object into a grasp that is appropriate for interacting with that object?
2. How may a variety of sources of information be used to select a single appropriate grasp? This information includes visual (from both the how and what pathways), context-related (what is the current task?), and abstract stimuli (e.g., an instruction stimulus).
3. How can the decision to execute a particular sub-program modulate decisions made for other components of the program? For example, how does the selection of the grasp affect the positioning of the wrist, and how can we implement this computation neurally?
4. How is the program coded at different levels? In AIP, what are the primary parameters being represented; how are they coded? In F5, how are the grasps represented and then played out further down stream?
5. What mechanisms provide the neural activity patterns seen experimentally, especially within AIP and F5?

## **Development of the Reaching/Grasping Model**

### **Visual Processing (The How, What and Where of Objects in Space)**

When a human or monkey directs attention towards an object, what types of information are extracted from the resulting retinal image that ultimately will be useful in somehow interacting with that object? Goodale and Milner have argued that two very different classes of information are extracted by distinct visual pathways (Goodale, Meenan, Bulthoff, Nicolle, Murphy, & Racicot, 1994b; Grady, Haxby, Horwitz, Schapiro, Rapoport, Ungerleider, et al., 1992). The *what pathway* (or ventral stream) passes from the visual cortex to the inferotemporal cortex (IT), where information about the object becomes available for access by "higher" cognitive centers. The *how(or where) pathway* (the dorsal stream) channels the

visual information from visual cortex to parietal lobe. It is this pathway that extracts parameters that are critical for physically interacting with the object.

One patient studied by (Goodale, et al., 1994b), referred to as D.F., suffered from a bilateral lesion of the ventrolateral occipital region (affecting the what pathway). Although she was capable of grasping objects just as normals would (preshaping her hand during the reach and placing her fingers on the objects in locations that facilitated a stable grasp), she was unable to discriminate between objects. In a second task, D.F. was repeatedly presented with a rectangular slot placed at various orientations (Goodale, Jakobson, Milner, Perrett, Benson, & Hietanen, 1994a). Although she was able to place a card within the slot, she could not indicate with her hand the orientation of the slot (the indicated orientation did not correlate with the actual orientation of the slot).

A second patient, R.V., suffered from a lesion in the occipitoparietal cortex, affecting the how pathway (Goodale, et al., 1994b). This patient showed no deficit in her ability to discriminate complex objects, but was unable to shape her hand appropriately in order to establish a stable grasp on the object. (Jeannerod, Decety, & Michel, 1994) reported a similar case, in which the patient had suffered a bilateral posterior parietal lesion. This individual, referred to as A.T., had apparently lost the ability to translate object size into hand aperture size. Consequently, she had developed a grasping strategy in which she preshaped during the reach movement by opening her hand to its greatest extent, and then clumsily enclosing her hand around the object once the arm had brought it within reach.

These observations suggest that these two visual systems are very distinct in the type of information that they compute and that functionally they solve a separable class of problems. However, it is a significant oversimplification to believe that these two pathways do not at all interact (Horwitz, Grady, Haxby, Schapiro, Rapoport, Ungerleider, et al., 1992). When D.F. is presented with a T-shaped object and asked to insert it into a T-shaped hole, she is able to accomplish the task on about half the trials (Goodale, et al., 1994a). On most of the

remaining trials, the orientation that she chose was approximately  $90^{\circ}$  from the correct orientation. Although the dorsal system has enough information to be able to align a piece of the object (e.g. the bar of the T) relative to one of the two slits that make up the T-shaped hole, it is the ventral system that provides the correct matching of the corresponding components.

An interesting interaction between the dorsal and ventral systems (or lack thereof) has also been observed in A.T. by Jeannerod (Jeannerod, et al., 1994). The most significant deficit in the grasping task is seen when the object being grasped is a neutral one - in other words, an object that is not part of A.T.'s everyday experience. When the object is one that is more familiar, the deficit is not as pronounced. In these cases, A. T. typically preshaped as one would normally.

In addition to the simple how/what distinction of visual processing, Jeannerod has gone further to partition the dorsal system into two distinct pathways (Jeannerod, 1994). One of these pathways is responsible for the representation of object shape and size, which is then used in the programming of the grasp. The other pathway captures the location of the object in space, providing information that can be used in selecting a goal position for the arm. This distinction explains D.F.'s ability to reach accurately towards an object while not being able to shape her hand appropriately.

### **Interaction of the How/What Pathways**

The experiments performed with D.F. and A.T. have demonstrated that both the how and what pathways can individually compute grasp parameters for certain classes of situations. But how do these two pathways interact during grasping in a normal primate? One hint of this interaction is provided by experiments performed with a hemi-field blindsight patient (Toni, 1994).

Blindsight is a condition in which a lesion has occurred within the primary visual cortex, effectively removing it from the visual processing circuit (King & Cowey, 1992). Because it is the primary visual cortex that feeds many of the other visual processing areas, the visual

deficit is very significant -- to the extent that subjects report that they are unable to see anything (Weiskrantz, Warrington, Sanders, & Marshall, 1974). However, it has been observed that subjects are still able to perform some visual tasks (most involving high-contrast or movement situations), although the patients are not consciously aware that they are using vision to perform the tasks (and in many cases have to be asked to "guess" at the answer). These observations are explained by the fact that the retina has an evolutionarily older pathway to the superior colliculus. The superior colliculus, in turn, provides some very limited visual information to the higher centers via the pulvinar (Goodale, 1993).

In the experiment of Toni, the hemi-field blindsight patient (only one hemisphere of the visual cortex had been damaged) is presented with a series of objects which must be grasped (bright objects against a dark background). With his "good eye" covered, the patient reported that he could not see the object presented to him. However, when asked to reach out and grasp the object, he was able to perform the task with reasonable accuracy. Moreover, just as with normals, the maximum grasp aperture (distance between the tip of the thumb and of the index finger) that was achieved during the reaching movement correlated with the size of the object. This implies that some visual information, via the superior colliculus, is available for aiding in the selection of grasp parameters.

In contrast, when the patient was asked to pantomime or verbally report the size of the object, the indicated size did not correlate with the actual size of the object. Furthermore, when asked to verbally report the size of the object during a reaching movement to the object, not only was the reported size not correct, the preshape was also incorrect. There was, however, a very strong correlation between the reported object size and the size of hand aperture during preshape.

What we can conclude from this set of experiments is that:

1. The visual information via the superior colliculus is not available to the what system.
2. This information is available to the how system.

3. When the what system is queried for an analysis of the visual scene and does not have any useful inputs, not only does it randomly select an answer, but it imposes this answer upon the how system.

### **Adaptation of the Visual/Motor Mapping**

In a recent experiment, Gentilucci examined the potential plasticity in the mapping from visual input into grasp parameters in human subjects (Gentilucci, Daprati, Toni, Chieffi, & Saetti, 1995). The subject is seated in front of a horizontally placed mirror (Figure 4.1). An object that is to be grasped by the subject is placed beneath the mirror (this object is referred to as the *real object*). Due to the placement of the mirror, the object and the subject's hand and arm are not visible. A second object (the *virtual object*) is placed such that its image in the mirror is superimposed on the location of the real object. The result is that the subject believes that she is looking at the real object, but that during grasping is just not able to see her hand or arm.

During the first block of trials, the virtual and real objects are identical. During each trial, the subject responds to a go signal by reaching out and grasping the real object using a precision grasp; the kinematics of movement are measured during this process by tracking the 3-dimensional position of the wrist, and the tips of the index finger and thumb. As has been observed in earlier experiments (MacKenzie & Iberall, 1994; Paulignan, Jeannerod, MacKenzie, & Marteniuk, 1991), the maximum aperture achieved during the preshape phase of movement (the distance between the tip of the thumb and of the index finger) relates directly to the size of the object being grasped.

For the second block of trials, the virtual object is the same as in the first block of trials, but the size of the real object is changed by a small amount (+/- 10-12% in diameter). After completion of the second block of trials (12), Gentilucci observes the following:

1. The subject is not consciously aware that there is a difference between what is being seen and the object being grasped.

2. During the course of the second block of trials, the subject adapts the movement such that at the end of the block, the kinematics are as if the subject was both looking at and grasping the real object (i.e., the maximum aperture corresponds best to the size of the real object, not the virtual object).

3. When presented with the virtual object and asked to pantomime its actual size, the size indicated by the subject correlates with that of the real object, not that of the virtual object.

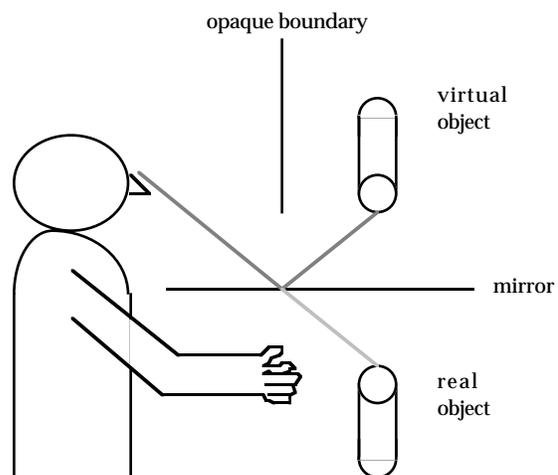


Figure 4.1: Experimental setup for the Gentilucci grasp adaptation task. The position of the virtual object is such that its image in the mirror corresponds to the location of the real object. An opaque boundary prevents direct viewing of the virtual object.

What observation #2 tells us is that there is a relatively quick adaptation of the mapping from the visual representation of the object to the parameters of the grasp that is subsequently executed. So, where is this adaptation taking place? The pantomiming task (observation #3) involves conscious access to the size of the object (note that D.F. also has difficulty with pantomiming the size of objects (Goodale, et al., 1994b) ). Because this task is also affected by the adaptation that takes place during the second block of trials, we can conclude that at the very least, changes that are occurring involve the inferotemporal system (the what pathway) or a system that makes use of the inferotemporal object representation.

### High-Level Model of Visual Processing

A high-level model summarizing the key experimental observations presented thus far is depicted in Figure 4.2. The primary visual pathway from the retina forks at the visual cortex

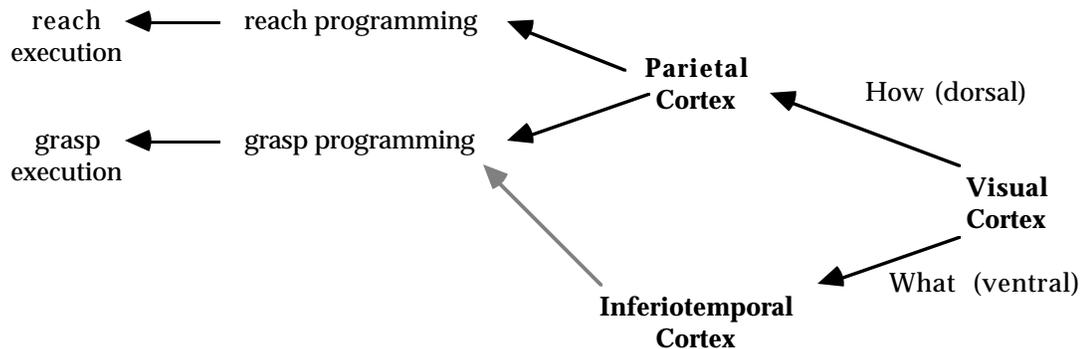


Figure 4.2: A high-level model of cortical visual processing and its support of the reach/grasp motor program.

into two functionally distinct pathways. The *what pathway*, leading to the inferotemporal cortex (IT), extracts the identity of the attended object, which in turn is associated with the object's properties (e.g., name, color and size).

The *how pathway* (Parietal Cortex) computes a parametric representation of the object being attended. In other words, motor-related descriptors are extracted that code information including the shape of the object, its size, and its location and orientation in space (Goodale, 1993). By *motor-related descriptors*, we mean that any object that is attended can be represented to some approximation with the descriptors, whether the object is a familiar one or not. In addition, two objects that share a common shape and size (e.g., a stick and a pen) will be coded in essentially the same way. In contrast, because these two objects are actually very different entities (in this case in color, texture and function), IT will utilize dissimilar representations.

Both the dorsal and ventral systems contribute to the selection of the grasp to be executed, as well as the grasp's associated parameters (in the figure these decisions are made in the *grasp programming* module). Here, *grasp programming* is distinguished from *grasp execution* in that it is a static representation of the grasp that is about to be executed. The grasp execution module is responsible for expanding this description into a temporal sequence of movements.

Because a single object is coded in two different ways in the dorsal and ventral systems, they contribute complementary information to the grasp programming process. The dorsal pathway provides a set of grasps that are specific to the shape and size of the object, which will often work for many situations. The ventral system provides context-dependent biasing of the grasp type (for selection purposes), and may also tune the grasp description (type of grasp and the associated parameters) for very specific situations that rely on context not carried by the shape and size of the object.

In the case of a parietal lesion (as with A.T.), the ventral system is the primary source of information for making the decision about the grasp (disregarding the sub-cortical contribution). When the presented object is a familiar one, a representation of this object exists within IT, from which an approximate grasp and accompanying parameters can be derived. However, when the object is one that is unfamiliar, its presentation does not activate IT, leaving the grasp programming module with no information. In human patients (specifically A.T.) this forces the adoption of other strategies for grasp programming.

### **Cortical Mechanisms for Object and Grasp Representation**

In this section, we turn from human behavioral data to neurophysiological experiments performed in monkey. From this data, a more detailed model is derived of the cortical regions involved in seeing, reaching, and grasping. We first focus on the brain regions involved in the visual processing (Parietal Cortex and Inferotemporal Cortex), then turn to representation of grasps (F5), and finally discuss the mapping of object to grasp (AIP). These discussions are then summarized with an updated model.

#### **The Dorsal System in Monkey**

As has already been discussed, the dorsal system is made up of two functionally distinct sub-pathways. One stream carries object-centered information (shape and size), while the

other represents the object's location and orientation in space. In monkey, distinct brain regions have been identified which are responsible for each of these classes of information.

The location of target objects is represented in the *Ventral Intraparietal area* (VIP) (Colby, Duhamel, & Goldberg, 1993). The coordinate frame appears to be a body-centered one, in that neurons are not modulated by the monkey's head or eye position. Locations in space are represented using a broadly-tuned population code. In this coding scheme, an individual neuron fires maximally for a small region of space (termed the *preferred region*). This activity drops off as the target position moves away from the preferred region. Although the individual neuron is active for a rather large region of space, the activity pattern across a large set of neurons can code positions rather accurately.

A neighboring region, the *Posterior Intraparietal area* (PIP), has been identified by Sakata as coding object-centered information (Sakata, 1994). Here, non-overlapping populations of cells have been shown to fire in response to different shapes that are presented to the monkey (common shapes used in the experiments included cylinders, plates, spheres and tori). Within a single population of cells, some showed specificity towards different parameter ranges (e.g., a *cylinder cell* might be responsive only to cylinders of width 10-20 mm). More generally, sub-populations of shape cells code the relevant parameters using a population-like coding scheme. Further details of this object representation scheme will be discussed in Chapter 5.

Note that Sakata does not claim (nor is it claimed here) that the monkey brain represents objects in terms of geometric primitives. Rather, these primitives act as a convenient partitioning of the objects into classes for testing of the monkey and for discussing what has been observed.

Neurons in PIP that code for the 3-D orientation of objects have also been identified (Sakata, 1994). A population code is used here as well -- with an individual neuron firing maximally for *preferred orientations*.

### **The Monkey Ventral System**

(Logothetis & Pauls, 1995; Logothetis, Pauls, & Poggio, 1995; Tanaka, 1993) have examined cell responses in IT as a function of the object configuration that is presented to the monkey. In general, individual cells seem to be responsive to complex combinations of object properties. The cell responses were generally independent of the object orientation in the coronal plane, but did show specificity towards certain orientations in depth. In addition, many cells were not affected by the retinal size or position of the object's image.

A number of related experiments have looked at the population response to different types of visual stimuli. One key finding is that individual objects (or faces) are represented by sparse populations of neurons (Rolls & Tovee, 1995; Young & Yamane, 1992). This fact, combined with the fact that individual cells respond to complex combinations of features implies that two distinct objects are very likely to be coded by non-overlapping (or only slightly overlapping) populations of neurons.

Finally, several authors have also noted that cells within IT are capable of rapidly becoming responsive (on the order of minutes) to novel objects that are relevant to performing the task at hand (Miyashita, Date, & Okuno, 1993; Rolls, 1995).

### **High-Level Representation of the Grasping Program**

The inferior premotor cortex (composed of F4 and F5) plays an important role in the high-level representation of both the reach and grasp. We have already reviewed in detail the data related to this region (Chapter 2). Here, we will summarize the key computational points and show how this region interacts with the visual areas already discussed.

Neurons in one subregion of Rizzolatti's area F4 have been hypothesized to code the goal position of reaching movements (Fogassi, Gallese, Fadiga, Luppino, Matelli, & Rizzolatti, 1995). Activity of single cells are sensitive to the goal position of the movement, but not to the starting position of the arm nor to the position of the head or eyes relative to the body. The

positional goals are represented using a population vector-type representation, with individual neurons being active for rather large areas of space. Although F4 receives a very significant projection from VIP, it is not at present clear whether the space coded by F4 can be described by a peripersonal coordinate system (as is the case in VIP) or a joint-centered coordinate system. Temporally, we see activity related to the phase of movement: preparatory (or set), movement, and hold of position.

Area F5 is involved in the control of distal movements, with the dorsal half relating to the control of hand movements (Rizzolatti, et al., 1988). The key properties of individual cells in this region are as follows:

1. Cells tend to be most responsive for specific grasp types (or oppositions) made by the monkey. Cells coding for grasps that are not being executed are sometime inhibited.
2. The largest number of cells are dedicated to grasps involving the thumb opposing the fingers (pad opposition or precision grasp); a moderate number of cells is involved in the coding of side opposition; and a few cells are active during power grasps.
3. Most cells are most responsive to a particular phase of movement. The primary identified phases are: set, extension, flexion, and hold. The largest number of cells are active during the extension and flexion phases of movement.

### **Transformation from a Visual Representation to a Grasp**

Up to this point, we have discussed cells that code information about the properties of the object to be grasped (in PIP and IT), as well as cells that code for the grasp that interacts with the object (in F5). In this section, we implicate AIP as the intermediate area responsible for the transformation of object to grasp.

Anatomically, AIP receives significant projections from PIP (Sakata, 1994), and exchanges heavy recurrent connections with F5 (Jeannerod, Arbib, Rizzolatti, & Sakata, 1995). Neurophysiologically, as we have already seen in detail in Chapter 2, individual cells

exhibit behavior that reflects both object-related and grasp-related properties (Taira, et al., 1990). The key properties of these cells are summarized as follows:

1. Many cells within this area are selective for the type of object that is presented. Most often, cells are actually selective to varying degrees for several objects. Typically, these objects share some common characteristics, and may also be grasped in the same way.

2. Many cells show significant modulation in correlation with the phase of movement. The identifiable phases include: the *key* (set phase), extension (preshape), flexion (enclose), hold, and return. Cells participate in varying degrees during different phases of the movement, but most are highly active during the extension and flexion phases of movement.

3. Once a cell becomes active in the process of executing a reach and grasp, in most cases it will remain active until the task is completed (the time at which the object is released).

4. A few cells respond when simply looking at the object (Sakata refers to these as *visual-related cells*), others are modulated only by the movement to manipulate the object (these *motor-related cells* are even active when the monkey is behaving in the dark). However, most cells fall somewhere between these two extremes.

5. Some cells demonstrate specificity toward the size of the object to be grasped.

6. A small number of cells show modulation based upon the object's position and/or orientation in space.

### **Expanded Model**

An expanded version of the high-level model is presented in Figure 4.3. Comparing with the original model (see Figure 4.2), the strictly functional modules (reach/grasp programming/execution) have now been assigned to particular brain regions.

VIP codes target object position and orientation in space using a population code in a peripersonal coordinate system. This target position is passed to F4, which represents the arm goal position. The distinction that was initially drawn in Figure 4.2 between the *programming*

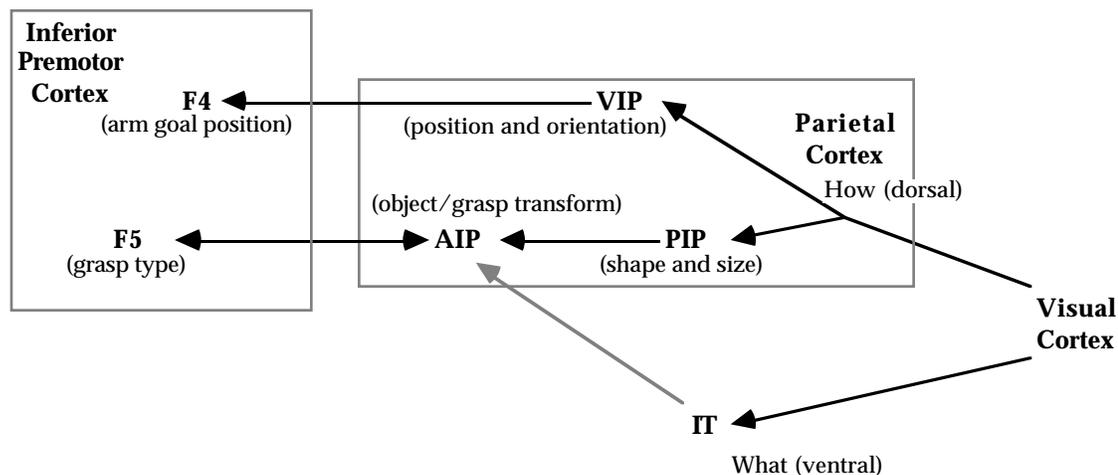


Figure 4.3: The expanded seeing/reaching/grasping model. The primary areas are IT (inferotemporal cortex), VIP (ventral intraparietal), PIP (posterior intraparietal), AIP (anterior intraparietal), and F4 and F5 (of the ventral premotor cortex).

and the *execution* of the reach has not gone away, but instead both have been absorbed into F4. Within F4, it is the *set cells* that are responsible for setting up the initial reach program, and then the *movement-* and *hold-related* cells that are responsible for executing the movement and maintaining the arm position, respectively.

As discussed above, PIP captures the shape and size of the object to be grasped. The details of this coding in the model will be discussed in Chapter 5.

AIP is hypothesized here as the first stage in the *grasp programming* process. As such, it is responsible for integrating object-related information from both the dorsal and ventral systems. The situation, however, is actually quite a bit more complicated in that AIP also receives grasp-related information from F5. Given that there are heavy recurrent projections between F5 and AIP, what are the differences in their behavior and hence their relative roles? The two key differences in behavior are:

1. Cells in F5 do not exhibit object-specific activity patterns (only grasp-specific behavior is observed), whereas we see both object- and grasp-specific activity in AIP.

2. Cells in F5 are active for only a small period of time during the execution of the movement. In AIP, once a cell becomes active (most are active by the beginning of the extension phase of movement, a few do not become active until the flexion phase), it remains active for the entire task, including the holding phase.

One interpretation of these differences (especially #2) is that AIP is serving as an *active memory* for the grasp that is about to be executed or that is being executed by F5. In other words, AIP is responsible for remembering the selected grasp, and F5 is responsible for unfolding this grasp in time during the execution of the movement. But why is it the case that F5 must send recurrent connections back to AIP? This would allow F5 to update the active memory with information about the grasp that is *actually executed*, which might differ in some ways from the grasp that was initially specified. Note here that we have deviated from the traditional neuroscience usage of the term *memory*, which is typically taken to mean that information is simply being held for some delay period so that it can be used later (e.g., Funahashi, Bruce, & Goldman-Rakic, 1993). Rather, we have taken a broader notion, in which the memory is continuously being updated as the result of other computations (in this case, as F5 monitors the progress of the grasp execution). This is similar to Dominey's *dynamic remapping*, in which motor afferent signals were used to update a model of the state of the eye plant as a saccade was being executed (Dominey, 1993). Here, the difference is that ascending sensory inputs can also be used to update the model state.

### **Execution of the Grasp**

In this section, we will first review experimental data that focuses on several cortical regions that are involved in the execution of the grasp. We will then further develop the model to incorporate this data, and finally show a more detailed perspective on the interaction between AIP and F5.

### **Cortical Mechanisms of Low-Level Grasp Control**

Feedback sensory information regarding the hand flows via the thalamus (ventral posterior lateral nucleus) through the primary somatosensory cortex (SI, which is composed of Brodman's areas 3a, 3b, 1, and 2). SI is a topologically organized strip that runs in parallel with the primary motor cortex, posterior to the central sulcus. Most cells in SI are responsive to only one modality: kinesthetic (areas 3a and 2) or tactile (areas 3b and 1) (Pons, Garraghty, & Mishkin, 1992). For each modality type, we see cells that reflect the basic receptor classes: slowly adapting (giving absolute information), and fast adapting (giving derivative information). The receptive field of a single unit in areas 3a and 3b tend to be restricted to small regions (on the hand); in areas 1 and 2, cells are responsive to larger regions of the hand.

The secondary somatosensory cortex (SII), receiving inputs from SI, provides a higher-level perspective of the current state of the hand. Many cells are responsive to both modalities and have receptive fields covering very large areas of the hand. These higher-order cells are capable of coding for very specific global hand configurations (Corkin, 1978). A lesion of SII in human results in a deficit in tactile recognition of objects, although the subjects are capable of describing the individual features of the object. In addition, a recent PET study examined the differences in pointing at an object versus grasping of the object (Grafton, Fagg, Woods, & Arbib, 1995). Subjects were instructed to grasp the object *very gently* and hold the position for a short period of time before releasing. One of the key findings of the study was that SII was more significantly active in the grasping case. Most likely this result is directly related to the fact that the subjects were instructed to very carefully grasp the objects.

The primary motor cortex (referred to as MI, Brodman's area 4, and F1) is one of the main sources of the corticospinal pathway (which supplies motor commands to the muscles) (Kandel, Schwartz, & Jessell, 1991). In the case of the hand, however, an even larger proportion of the corticospinal tract originates from MI. Microstimulation of the hand area of MI produces visible twitches in the fingers of the hand (Picard & Smith, 1992a; Picard & Smith, 1992b). An

individual PTN MI cell typically projects to a set of muscles that are involved in controlling a single joint or several neighboring joints. We shall refer to this set of muscles as a *motor assembly*.

MI receives projections from many regions; here we are concerned with two sources. F5 sends motor commands to the hand via MI (Matelli, 1994). In our model, a group of F5 cells forms a coarse code for the grasp that is being executed. Therefore we imagine an F5 cell as selecting a set of motor assemblies that are relevant to implementing that grasp. We will return to this point in the next Chapter.

Sensory information from SI (and thalamus) can also activate motor cortex cells, implementing one level of reflex loops. Picard & Smith (1992a) examined the modulation of motor cortex activity by cutaneous and proprioceptive information during a precision grasping task. Two monkeys were taught to grasp an object with a lateral pinch (side opposition). For any particular trial, the object was chosen from one of three weights and one of four surface textures (oiled metal, smooth metal, and fine and coarse sandpaper). The monkey was allowed a view of the object before it was grasped. The grasped object was instrumented with a position sensor, as well as force sensors that measured horizontal grasping force and vertical lifting force.

Grip force was examined (during dynamic and static phases of movement) as a function of object type. For a particular surface type, grip force increased with weight of object. In all but one case, larger forces were applied to the more slippery objects. Most cells sensitive to weight changes increased in their activity, but a few showed a reduction in activity. Figure 4.4 compares the receptive field and motor output of a single cell that increases in activity as the object weight is increased. The cell was responsive to cutaneous stimuli on the lateral side of the index finger. However, stimulation of the cell resulted in abduction of the thumb.

### Model of Grasp Control

The component of the model responsible for the execution of the grasp is shown in Figure 4.5. As has already been discussed, the grasp programs in F5 are represented by units that participate in a specific grasp and a particular phase (for purposes of the discussion we can assume that a single cell participates in exactly one grasp and during one phase). For phases in which movement is required (i.e. extension, flexion, hold, and release), the F5 cells drive the hand by activating very specific sets of primary motor cortex (MI) cells. This activation can either specify equilibrium positions for the joints of the hand (Katayama, 1993) or can configure reflex loops in which MI cells respond to inputs from SI (Picard & Smith, 1992b). More details of this implementation will be discussed in the following chapter.

SII cells are responsible for detecting specific combinations of SI activity patterns (sensory hyperfeatures). These hyperfeatures are used in two ways:

1. Signaling of key events in the execution of the grasp, which can be used by F5 to trigger transitions in the phase of the program (e.g. transition from the flexion to the hold phase when contact with the object is detected).
2. Recognition of the object being grasped, which allows for refinement or reconfiguration of the original vision-based program.

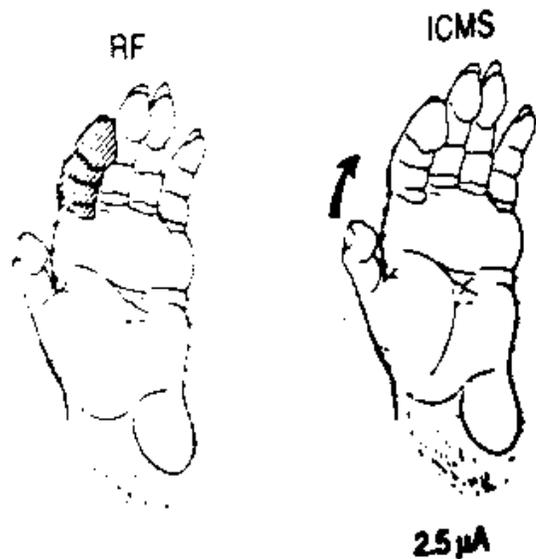


Figure 4.4: Motor cortex cell whose firing correlated with cutaneous stimulation of the lateral side of the index. Direct stimulation of the cell caused abduction of the thumb (Figure 4 of Picard & Smith, 1992a).

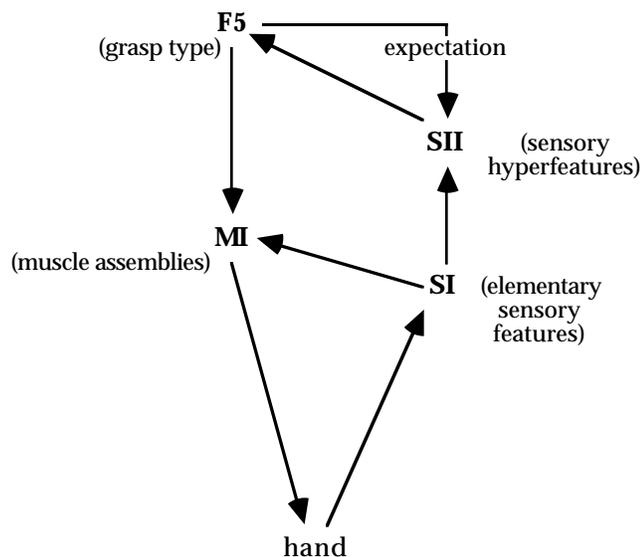


Figure 4.5: The cortical regions involved in the low-level control of the grasping motor program. The areas are F5 (inferior premotor cortex), MI (primary motor cortex), SI (primary somatosensory cortex), and SII (secondary somatosensory cortex).

In addition to the inputs received from SI, SII cells can also receive inputs from F5 units. These inputs express an *expectation* of coming sensory events (for example contact with a particular object), which allows the SII cells to more quickly and more significantly respond to the detection of the event. This expectation is also an important mechanism for filtering out incidental events which are not relevant to the current grasp (or phase) being executed.

Because expectation alone should not necessarily cause an SII cell to become active, a fundamentally different type of input is used. We refer to this input as a *priming signal* (Dakin, 1986). An individual SII cell must receive both a priming signal (from F5) and an *activating signal* (from SI) in order to fire. We will return to this distinction in Chapter 5.

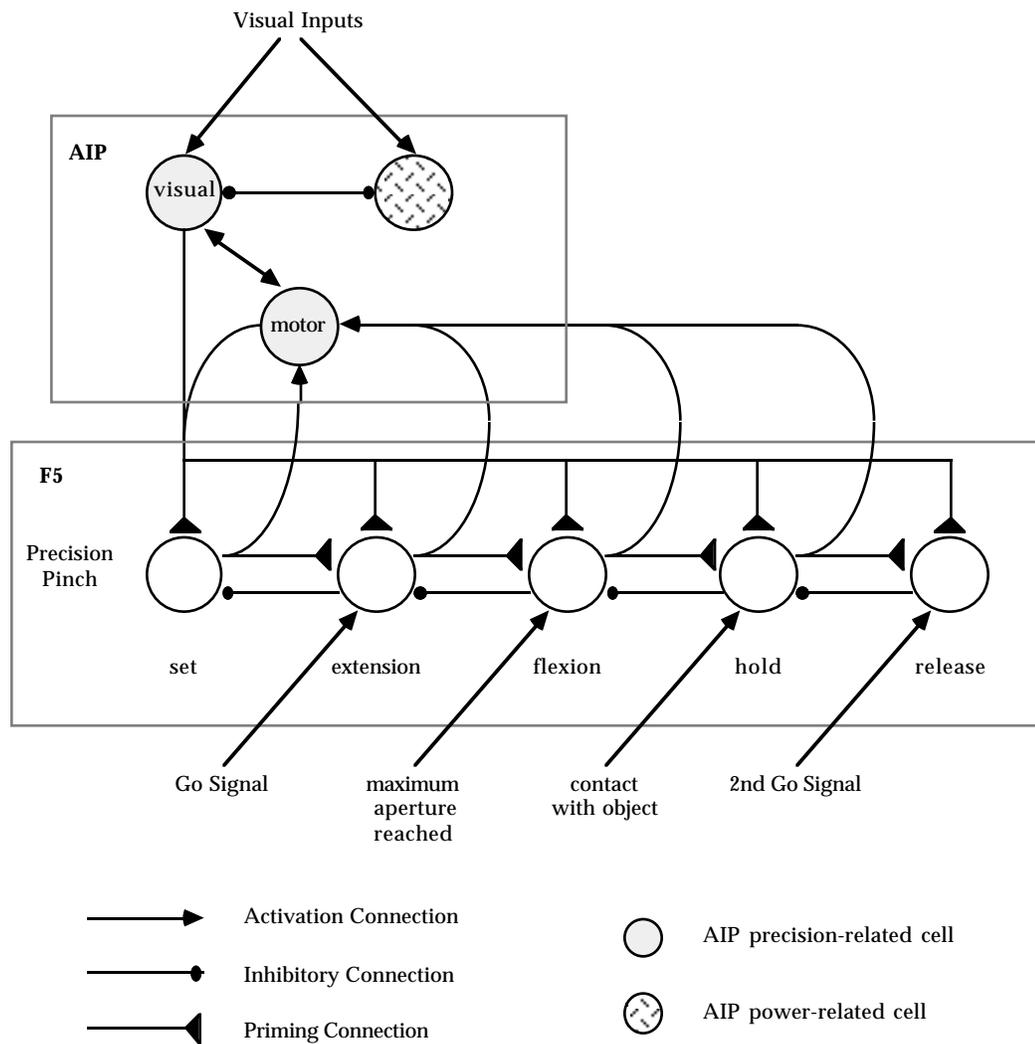


Figure 4.6: Interaction between AIP and F5 populations (circles) during execution of the preshape/grasp/hold program. Activity within F5 cascades from left to right as the program is executed. At each program phase, the state is reported back to the AIP motor-type population.

### Interaction of AIP and F5 During Execution of the Grasp

Figure 4.6 demonstrates at a schematic level the interaction between populations of cells in AIP and in F5, with particular emphasis on the active memory and phasic activation mechanisms. Three AIP cells<sup>1</sup> are shown: a visual-related cell that recognizes objects that

<sup>1</sup> Although we are using the term *cell* here, we are really referring to a population of cells.

require a precision pinch, a motor-related cell of the same type, and a visual-related cell that recognizes objects requiring power grasps. The five F5 units participate in a common program (in this case, a precision grasp), but each cell fires during a different phase of the program. We illustrate the dynamics that take place between AIP and F5 by walking through one full task (in this case the Sakata paradigm).

Initially some set of visual parameters that is specific to a particular object activates the visual-related cell in AIP (the cell that is precision-grasp related). A competitive process takes place within AIP, shunting the activity of cells that correspond to other grasps (in this case with a power-grasp related cell).

This activation *primes* a set of cells within F5 that correspond to the grasp to be used (in this case the precision grip). As a result, the F5 set-related cell becomes active, signaling a preparation for execution of the precision grasp. The set-related activity prepares the sub-program for execution by priming the cell representing the extension phase of the motor program.

When the trigger (Go) signal is given by the experimenter, the extension-related cell begins to fire, causing the preshape of the hand to begin (downward motor commands are not shown). At the same time, the activation of this cell forces the set-related cell to turn off (via the backward inhibitory connection), and the flexion-related cell is primed.

At the time that the fingers reach an extension appropriate for the size of the object to be grasped (as detected by the internal model of hand state or by an SII unit, neither of which are shown in the Figure), the flexion-related unit is activated, marking the beginning of the enclose phase of movement. In response to this activation, the extension-related cell is turned off and the hold-related cell is primed. The next phase transition occurs when a tactile stimulus is detected (by an SII unit); the hold-related neuron activates, shunting the flexion-related cell, and priming the release-related cell. The hold neuron continues to be active until the external

secondary go signal is received. This signal causes the activation of the release neuron and the shunting of the hold neuron.

At each phase of the program, the corresponding F5 unit sends excitatory input to the motor related cell in AIP. The recurrent priming signal from this AIP cell serves as a active memory of the motor program that is currently being executed. This memory ensures that at any point in the execution of the grasping program, the F5 cells that correspond to the precision grasp (in this example) are activated during the next phase of movement, and not cells that participate in some other grasp. We will return to the details of these dynamics in Chapter 5.

As presented, the neural circuit in Figure 4.6 appears as a rather rigid structure. The neural implementation, however (as we will see in Chapter 5), generalizes along several dimensions. First of all, the "populations" of cells (depicted by circles in the Figure) do not necessarily fall into discrete sets of cells. Different grasps (in F5 and AIP) or phases (F5) are able to share cells to varying degrees. Second, we do not imagine the connections implementing the phasic behavior as being hardwired into F5. In the following Chapter, we present a more general structure in which the basal ganglia (BG) are implicated as representing the essence of the phasic connections. We imagine that in future models, the functional connectivity of the different populations could be reorganized so as to create arbitrary assemblages of component motor programs (Arbib, 1990) and/or sequences (Dominey, 1995). Third, although we treat each phase as an equal entity in the model, the extension and flexion phases of movement are probably better interpreted together as a *grasp phase* in the context of motor units that are selectable by BG. The within-phase modulation that implements the preshape and enclose are most likely handled by mechanisms other than the BG.

### **Remaining Computational Issues**

It is very often the case that a set of biological experiments yields an incomplete view of the system being studied. As part of the model-building process, however, the goal is to

construct a complete story as to how the system functions. As such, we are forced to rely on other sources of information to constrain the modeling. Possible sources of additional information come from *computational constraints*, in which we ask the question "given that our model must perform some task or exhibit some behavior, and that it has certain information available to it, how does it transform this information into a form such that it can successfully perform its task"?

Here, we outline several key computational issues that have not yet been addressed experimentally, and in the next section show how the model evolves in light of these constraints.

### **Object/Grasp Mapping is Many-to-Many**

Any object with which a human or monkey will interact has many different ways that it can be grasped. However, to this date, the experimental paradigms that have looked at the mapping from object to grasp have only focused on situations in which a single grasp is appropriate for any particular object that is presented to the monkey. How can this more general transformation take place (i.e., from an object to a set of grasps)? And when we have a set of grasps that are appropriate, what information is used to select one to execute?

In the model, we address these questions by expanding our view of the roles of AIP and F5. However, before we do this, it is appropriate to first introduce a somewhat new perspective on what it is that AIP might be computing. As we have noted earlier, cell activity within AIP can exhibit specificity towards both grasps and objects. One way to interpret this partitioning is that AIP is computing the set of *affordances* for the object being attended. Here, we use the term *affordance* to mean that the system is highlighting the features of the object that are specifically relevant for physically interacting with that object. Note that we are deviating from Gibson's use of the term (Gibson, 1950) in that we do not assume that the environment is providing the affordances, but rather they are being computed by the two visual pathways.

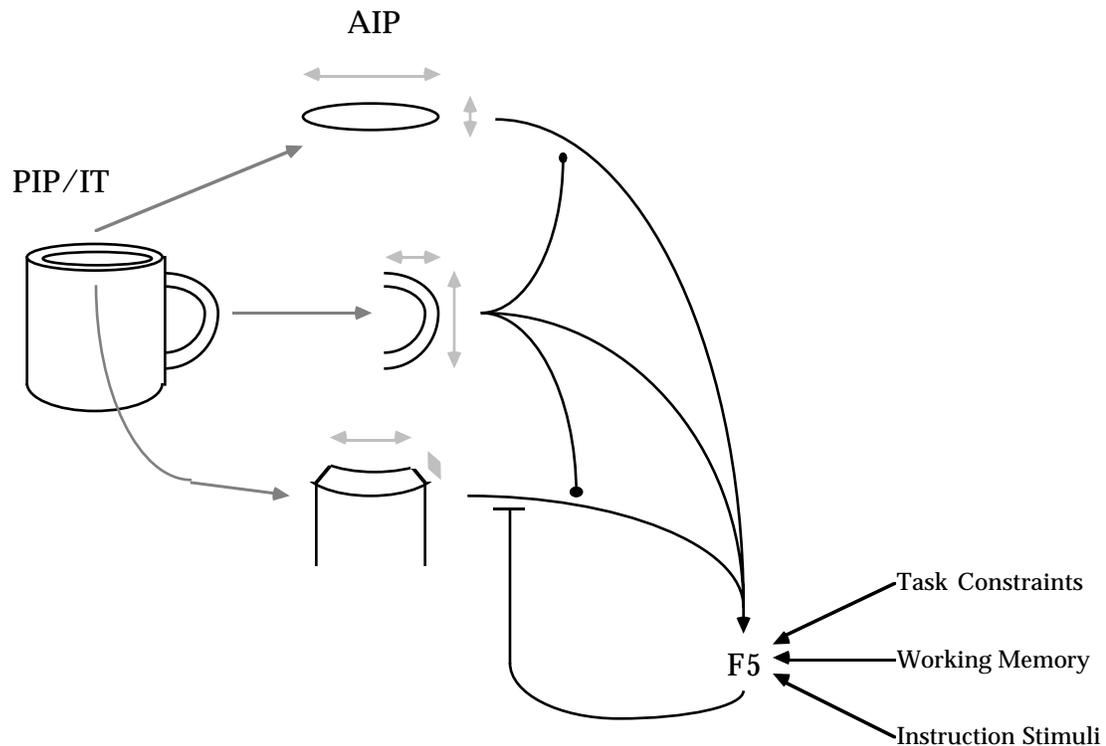


Figure 4.7: AIP computes the set of affordances for an attended object. These affordances highlight the features of the object that are relevant to physical interaction with it.

Figure 4.7 illustrates an example of computing the set of affordances for a mug. In this case, three different features are being highlighted. The circular rim of the mug affords a pad-opposition involving all four fingers (as in a grasp from the top of the mug). The handle of the mug affords a grasp that combines a palm opposition (e.g., index and middle against the palm) with a side opposition. Finally, a section of the rim of the mug affords a grasp in which the pads of the index and middle fingers oppose the thumb. Note that not only are the individual features highlighted, but the relevant parameters (e.g., the diameter of the rim) are also made explicit. Although this example has assumed one object feature is mapped to a single grasp, this is not generally the case. A cylinder, depending upon its size, can map to a precision pinch, a power grasp, or a lateral pinch.

When the mug is initially presented to AIP, based on inputs from both the dorsal and ventral streams, it computes the set of affordances. The corresponding set of grasps are passed to F5. As a function of task or abstract information, F5 selects one of the specified grasps. This decision is broadcast back to AIP, which shunts the other affordances, leaving only the affordance that corresponds to the selected grasp. During the execution of the grasp, the affordance represented by AIP is continually reinforced by inputs from the active grasp program in F5.

### **Modulation of the Reach Program by the Grasp Program**

Introducing the possibility of a single object being mapped to multiple grasps brings an additional level of complexity--the wrist must be positioned such that the selected grasp can be established correctly (Arbib, Iberall, & Lyons, 1987; MacKenzie & Iberall, 1994). The implication here is that not only must the object position be taken into account (as represented in VIP), but once the F5/AIP system selects the grasp to execute, some form of modulatory signal must be combined with this representation in order to compute the final goal position of the wrist. One anatomical candidate for this modulatory signal is the projection from AIP to F4 (Matelli, 1994).

### **Parameter Coding**

Experimental work on grasp coding in monkey cortex to date has focused primarily on the representational differences between grasp modes or objects of radically different shapes. However, in human behavioral experiments we can see subtle changes in the preshaping movement with small modifications to the size of the object (with the same grasp mode) (MacKenzie & Iberall, 1994; Paulignan, et al., 1991). How is it that these small differences in the object configuration and in the corresponding motor program are represented in PIP, AIP, and F5?

Sakata has observed that some AIP and PIP cells demonstrate sensitivity to the overall size of the object (Murata, Gallese, Kaseda, Kunimoto, & Sakata, 1993; Sakata, 1994). The fact that he observes cells that are both size-selective and size-independent indicates that within a population of cells that code for a particular affordance, a sub-population of these cells are responsible for capturing size of the object.

An important next question is: what is meant here by size? If we take as an example a cylinder, two parameters that describe the cylinder are diameter and length. In the experiments performed by Sakata these two parameters are scaled together, and never independently. If we could independently vary these parameters, how would this be reflected in PIP and AIP? Given that we have interpreted the cell behavior in PIP as simply representing the object and its parameters, we would expect to see these parameters reflected equally (approximately the same number of cells involved in each). However, if we interpret AIP as capturing only the information that is relevant to the physical interaction with the object, then a balanced representation may not be the case. Considering the cylinder, if the subject is performing a side opposition, then the diameter of the cylinder will be much more relevant to the execution of the movement than its length. Note that the length of the cylinder does not go completely unused--in the more general problem of which grasps are afforded by the object, this parameter could play a key role.

## **The Full Region-Level Model**

### **Further Biological Background**

The supplementary motor area (SMA) has been implicated in the planning and execution of complex movements. In human PET studies, we see a larger degree of activity in SMA during execution of a complex sequence of finger movements than for simple flexion of an individual finger (Roland, Larsen, Lassen, & Skinhøj, 1980). A number of studies have also demonstrated

an effect in SMA related to mental imagery of complex tasks (Grafton, Arbib, Fadiga, & Rizzolatti, 1996).

In monkey, a unilateral lesion of the SMA has been shown to disrupt the monkey's ability to allocate his hands to different subtasks of a bimanual task (Brinkman, 1984). In addition, the SMA appears to be involved in the temporal organization of complex movements (Tanji, 1994). Based on cytoarchitectonic and microstimulation evidence, Luppino et. al argue that there are actually two distinct areas within what has been traditionally called SMA (Luppino, Matelli, Camarda, & Rizzolatti, 1993; Luppino, Matelli, & Camarda, 1991; Luppino, Matelli, & Rizzolatti, 1990). These two regions are referred to as SMA-proper (F3; the caudal region), and pre-SMA (F6). F3 is somatotopically organized and has heavy projections to the limb regions of F1, as well as direct projections to the limb-related portions of the spinal cord. F6 does not project to the spinal cord, and has only moderate projections to areas F3 and F2 (the dorsal premotor cortex) (Luppino, et al., 1990). However, there is a very heavy projection to area F5. Inputs into area F6 include VIP, and area 46 (principal sulcal region of the prefrontal cortex). Recording studies by Rizzolatti have observed neurons that become active when an object that the monkey is about to grasp moves from being out of reach into the peripersonal space of the monkey (Rizzolatti, 1990). The interpretation of this neural response is that this class of neuron is responsible for generating a *go signal* when it is appropriate for the monkey to begin a reaching movement.

Area 46 has been implicated as a working memory in tasks requiring information to be held during a delay period (Kandel, et al., 1991; Quintana & Fuster, 1993). This memory can participate in the learning of tasks involving complex sequences of movements (Dominey, 1995). Anatomically, this region projects to F6, and also exchanges connections with area F5 (Luppino, et al., 1990; Matelli, 1994). In human, (Decety, Perani, Jeannerod, Bettinardi, Tadary, Woods, et al., 1994) have recently discovered that area 46 is involved when a subject is asked to

imagine herself grasping an object. In addition to this region, they observe activity in area 44 (possible F5 homologue), as well as a site along the intra-parietal sulcus (Grafton, et al., 1996).

Dorsal premotor cortex (F2) is thought to be responsible for the association of arbitrary stimuli (an IS) with the preparation of motor programs (Evarts, Shinoda, & Wise, 1984; Kurata & Wise, 1988; Mitz, Godshalk, & Wise, 1991; Wise & Mauritz, 1985). In a task in which a monkey must respond to the display of a pattern with a particular movement of a joystick, some neurons in F2 respond to the sensory-specific qualities of the input. However, many units respond in a way that is more related to the *motor set* that must be prepared in response to the stimulus. When a muscimol lesion in this region is induced, the monkey loses the ability to correctly make the arbitrary association (Kurata & Hoffman, 1994).

### **The Complete Model**

The complete region-level view of the model is shown in Figure 4.8. This figure brings together our earlier models for visual processing and grasp execution, and incorporates the biological data discussed in the previous section. In the following Chapter, we describe the neural-level implementation of these circuits, focusing on PIP, AIP, F5, MI, SI, and SII.

The projection from AIP to F4 carries wrist position modulatory information that is to be combined with the object location information flowing from VIP. This modulatory information is specific to the single affordance/grasp pair that is selected by the AIP/F5 system.

Regions 46 and F6 are responsible for representing the high-level program. F6 first prepares the ventral premotor regions for execution of the coming reach and grasp by priming both F4 and F5. It then detects the *go signal* given by the experimenter and initiates execution of the program (generates the go signal shown in Figure 4.6). Initiation of the *release phase* of movement is also handled in this way.

Area 46, working in conjunction with F6, is responsible for supplying any task-dependent biases for the selection of the grasp in F5. This selection can be based upon the task

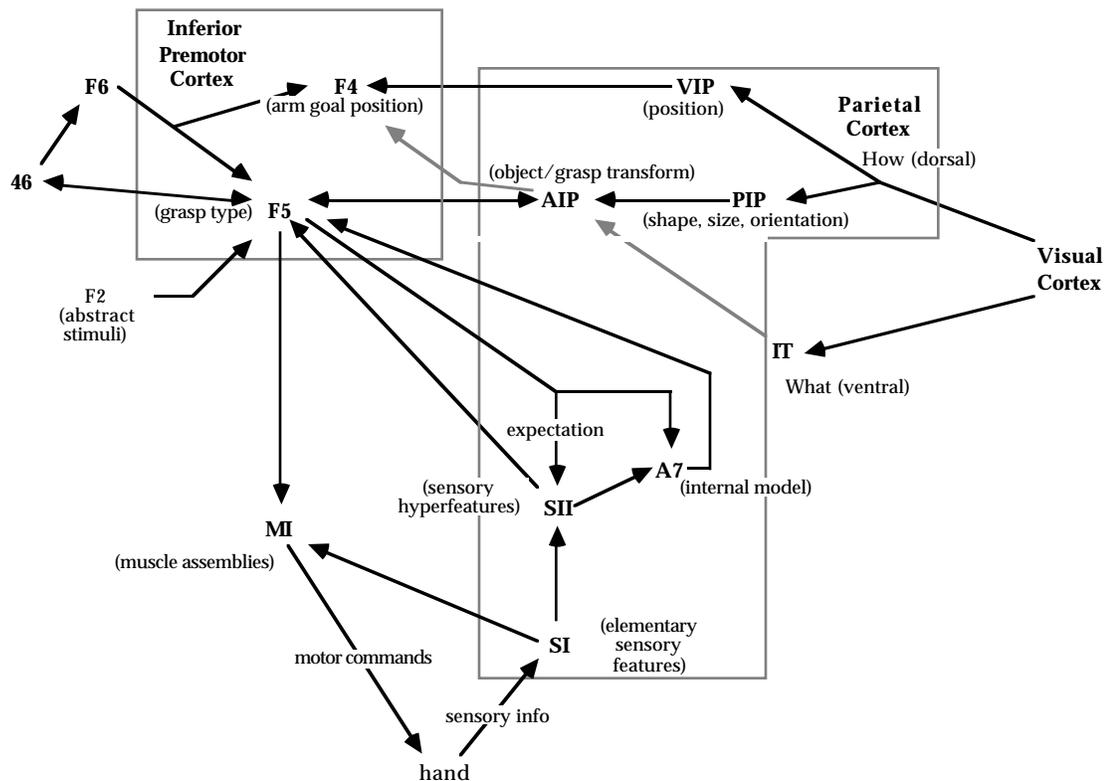


Figure 4.8: The complete model. The two key additions to the model are the projection from AIP to F4, and the introduction of regions 46, F6, F2, and area 7.

requirements (such as what is going to be done after the grasp), or based upon a working memory of a recently executed grasp. The biasing can be done at the level of the class of grasp (e.g. power versus precision), or can also include the parameters of the grasp (e.g. width of the aperture).

### Progression of Activity During Task

In this section, we chart the flow of information through the model as a variant of the Sakata task is performed. The steps are similar to those presented at the beginning of this chapter.

Step 1: The trial begins with F6 preparing to receive the *ready signal*.

Step 2: The ready signal is presented and the target object becomes visible. F6 responds by priming F4 and F5. Area 46 extracts any task-specific information that is available and computes a grasp bias which is passed to F5.

The visual information flows from the visual cortex to the inferotemporal and parietal cortices. IT identifies the object (if it is a familiar one). VIP computes the object's location in peripersonal space. PIP extracts the parametric description of the object (shape and size). The resulting activity patterns in PIP and IT activate a set of affordances in AIP.

The active affordances in AIP excite the set of corresponding grasps in F5 (only the set cells in F5 are activated). Within F5, the active grasps compete through a winner-take-all mechanism (Didday, 1976; Fagg & Arbib, 1992), which incorporates any biases that might be received from area 46. The final remaining grasp is reported back to AIP, where the other affordances are shunted. The selected affordance sends wrist modulatory information to F4, where it is combined with the location of the object from VIP.

Step 3: The flash of the LED (the go signal) is detected by F6 and then passed on to F4 and F5. Within F5, the set-related activity gives way to activation of the extension-related cells (we assume that the F4 execution mechanism is similar to that of F5). These cells initiate the preshape movement by activating the appropriate set of motor assemblies within MI. In addition, event detectors in SII are primed that are responsible for sensing the point at which the maximum aperture is reached. As execution of the extension progresses, the internal model of hand state is updated from the descending motor commands. In parallel with SII (but without the sensory delay), the internal model estimates the point at which the hand has reached maximum aperture (we implicate area 7 as implementing the internal model).

When this event is detected, the flexion-related cells in F5 are activated. In turn, the extension cells are shunted, the flexion movement is initiated within MI, and the SII cells responsible for detecting contact with the object are primed. The MI motor assemblies that are

activated involve not only an absolute movement (to the expected width of the object), but also involve reflex loops that will react (at a joint level) to sensed forces.

Step 4: Contact with the object is established by all fingers involved in the grasp. This event is detected by SII. In F5, we see a phase transition from flexion to hold. The hold cells in F5 activate the MI motor assemblies that are involved in maintaining a stable grasp. In addition, these cells prepare for the release event by priming the F5 release cells.

If there is a miscalibration between the expected and actual global hand states (e.g. in the case of an object with a width that is different than what was assessed by the visual system), then this is detected in the model by SII. The result is that a different set of SII units will become active than the set that was primed by the F5 flexion units. In F5, this means that the activated hold cells will correspond to the object that is actually grasped, and not to the set that would normally become active when there is not a perturbation. This change in grasp program is propagated to AIP, where we see a shift in the representation of the affordance.

Step 5: The reception of the secondary trigger stimulus is detected by F6, which forwards the trigger on to F4 and F5. F5 activates the motor assemblies that cause the hand to open just enough to release the target object. This event is detected by SII. F6 then shunts its priming signal to F5 and F4.

## Summary

In this chapter, we have outlined the primary cortical regions that are involved in the seeing/reaching/grasping process and have assembled a high-level model of these processes. The key components of the model and their contributions are:

1. The dorsal visual stream (parietal cortex) extracts parametric information about the object being attended. This information provides the details of the object configuration that are used to finely tune the grasp plan. The representation is general in that both familiar and unfamiliar objects are represented equally.

2. The ventral system (inferotemporal cortex) extracts information that is used for conscious processing of the object. In the grasping task, this region contributes by biasing the grasp selection process towards grasps that are most appropriate for this specific object. The representation is a localist one, in which an individual object is coded by a sparse set of neurons, and different objects are coded by non-overlapping sets of neurons.

3. The anterior intraparietal area (AIP) is hypothesized as playing a dual role in the seeing/reaching/grasping process. First, it is responsible for integrating visual information from both the dorsal and ventral streams, and computing the set of *affordances* exhibited by the object. Second, as one of these affordances is selected and execution of the grasp begins, AIP behaves as an *active memory* of the one selected affordance and updates this memory to correspond to the grasp that is actually executed.

4. F5 is hypothesized as first being responsible for integrating task constraints with the set of grasps that are afforded by the attended object in order to select a single grasp. After selection of a single grasp, F5 unfolds this represented grasp in time to perform the execution. During execution, F5 monitors the progress and any changes to the program (due to unexpected events) are reported back to AIP so that the active memory can be updated.

In the following chapter, we further develop the neural-level details of the model.

## Chapter 5: Low Level Details of the Grasping Model

---

In this chapter, the region-level model presented in the previous chapter is expanded to a neural level of detail. We first introduce a computational structure, referred to as a  $\mu$ -schema, that serves as an intermediary between the schema level of analysis and the neural level of implementation. Making use of the  $\mu$ -schema mechanism, we turn to the implementation of the model, and discuss 1) how information is represented within each region, and 2) how individual computing units are connected in micro-circuits to implement particular aspects of the neural dynamics.

In addition, the basal ganglia are implicated as a mechanism that is responsible for the general management of schemas. In particular, the basal ganglia are involved in this model in three distinct ways:

1. Arbitration between competing schemas in F5 and AIP.
  2. Managing the phasing that takes place within F5.
  3. Providing a mechanism through which the selection of a grasp may be biased as a function of either task requirements or abstract stimuli.
- 

### $\mu$ -schemas: An Intermediate Computational Mechanism

Up to this point in our modeling discussion, we have thought of the various components of our motor program as abstract entities (schemas) that exist within a layer (or set of layers), but

we have not yet significantly addressed the issue of how these entities might be implemented at the neural level. The schema level of analysis is useful because we have been able to treat different programs, such as those for power grasp and precision pinch, as discrete and encapsulated entities. This allows us to assume essentially independent implementation and execution of these entities. Combined with the concepts of schema priming, cooperation/competition, and activation, this level of analysis significantly simplifies the implementation and our understanding of the resulting system.

In moving down to the neural level of analysis, we would like to maintain some notions analogous to the high-level computational concepts that we used at the schema level. However, two factors conspire against this. First of all, the discreteness assumption begins to break down. We have already seen in the analysis of the Sakata data how an individual AIP cell might participate in several distinct grasps. In fact, the level of participation is graded: for some grasps, the cell responds more strongly than others. The second difficulty is that since we have assumed that an appropriate level of the neural model is the leaky-integrator (Arbib, 1989), we necessarily lose the specialization of the various inputs that we had at the schema level (i.e., priming, competition/cooperation, and activation).

In order to address these difficulties, we introduce the notion of a  $\mu$ -schema (pronounced "micro-schema"), which is a computational unit that acts as an intermediary between schemas and neurons. Physically, an individual  $\mu$ -schema is composed of a set number of leaky-integrator neurons, each with a specialized function. A schema is then constructed by gathering together a set of  $\mu$ -schemas. However, it is possible for two schemas to share a single  $\mu$ -schema (we would refer to these schemas as *overlapping*). In addition, a  $\mu$ -schema's participation in a schema can be graded. In other words, its contribution to a schema's output may range continuously from significant to not at all.

A  $\mu$ -schema differs from the concept of a sub-schema (already provided by schema theory) in that sub-schemas are themselves schemas.  $\mu$ -schemas, on the other hand, are considered

atomic elements and cannot be decomposed further without losing the properties of the  $\mu$ -schema.

In the remainder of this section, we first show the  $\mu$ -schema implementation using a group of neurons, referred to as a *column*, and then show how columns within the different regions connect together to implement the neural programs described in Chapter 5.

### The Columnar Structure

The columnar structure of the  $\mu$ -schema is designed such that the schema theoretic concepts of priming, mutual cooperation/competition, and activation can still be used in the design and analysis of our neural models. In addition, the column is intended as a generic computing unit; every brain region in our model is made up of a set of these columns (as will be discussed, some configurational differences exist from one region to another).

The column consists of a set of four neural units (Figure 5.1). The primary function of the column is performed by the *signal unit*. This unit fires in response to the presentation of a specific set of *trigger inputs*. Trigger inputs can be, for example, *go signals* (as is the case for F5) or some combination of object features (as in AIP). Activation of the signal unit can cause the *output unit* to fire, which results in the generation of an output from the column. When this happens, we refer to the column as being *active*. The column output, depending upon the brain region, is connected either to other columns (in the same or other layers) or to actuators.

The simple feature detection that is performed by the signal unit is subject to two levels of constraints, without which the column cannot become active. The first constraint is that the column must be *primed*. Priming signals (detected by the *priming unit*) are typically received from high-level regions. For example, when F6 begins to prepare a grasp program, it will prime all the columns within F5. In this architecture, priming signals are considered to be non-specific; the priming of F5 by F6 prepares the system to perform a grasp, but does not necessarily specify the details of the grasp.

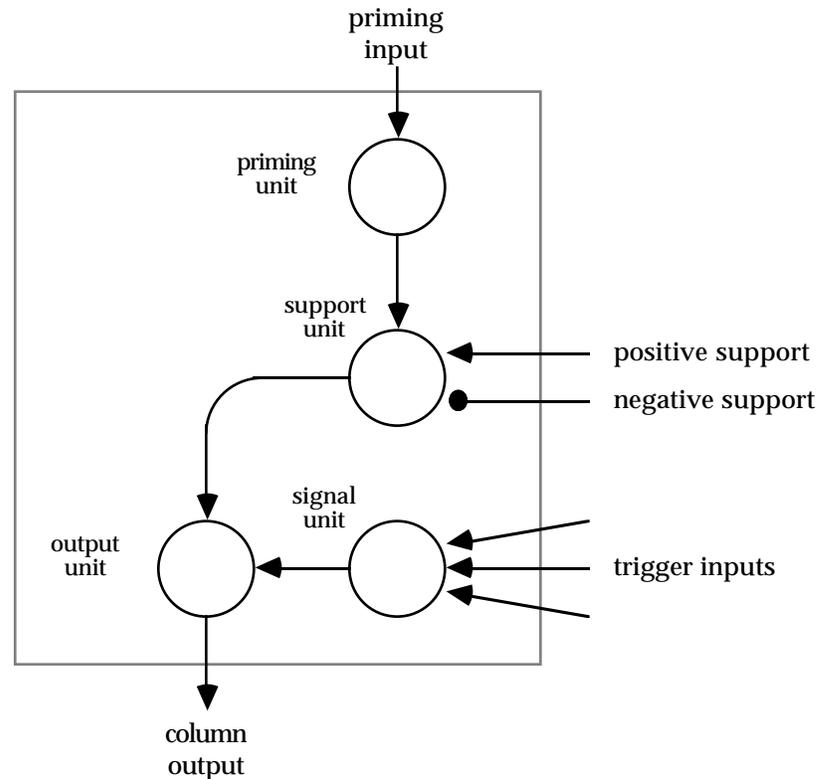


Figure 5.1: Column architecture.

The second level of constraints, received through the *positive/negative support* inputs, are used in two different ways. First, in order to enforce the coactivation of columns that participate in the same schema (e.g., a power grasp), positive connections are established from each column output to the support input of the other columns. Likewise, when two columns are involved in different schemas (e.g., a power and a precision grasp), mutually inhibitory connections are established. Second, when a higher-level region (e.g., area 46) selects a very specific schema within another region (a grasp in F5 as a function of some task constraints), it will establish positive connections to the support input of the columns participating in the schema.

The constraints are considered to be satisfied only when the column is both primed and receiving (at the very least) non-negative support. When this occurs, the column is prepared to become active when the correct trigger input is received.

The cortical column model proposed by (Alexandre, Guyot, Haton, & Burnod, 1991) is a connectionist-level model that attempts to capture the basic computations that are performed by small assemblies of real cortical neurons. This model has a number of interesting similarities to the  $\mu$ -schema structure presented here. First, the state of the fundamental computing unit (the column) is multi-dimensional; the outputs that are generated by the column are affected by the state of the two internal neurons (although the individual neurons can have only one of three states). Second, the inputs are segregated into two classes, the contribution of which depends upon the entry point into the column. In contrast, a self-organization-based learning algorithm has been proposed that relies on Hebbian-like mechanisms to adjust the column's connectivity with other structures.

### Column Dynamics

The equations that describe the dynamics of the column are given below. For computational efficiency reasons, only the output unit is implemented as a leaky-integrator neuron (Arbib, 1989). All other units in the column are implemented in a simple feed-forward manner.

The state of the *priming unit* is computed according to:

$$prime\_mem = prime\_inputs \cdot prime\_weights \quad (5.1)$$

$$prime = \text{Esat}(prime\_mem, p\_base) \quad (5.2)$$

where:

*prime\_mem* is the membrane potential of the priming unit.

*prime* is the firing rate of the priming unit.

*prime\_inputs* is a vector of inputs from other columns.

$prime\_weights$  is a vector of corresponding connection strengths.

$p\_base$  is the minimum firing rate when a non-zero priming input is received.

$$Esat(x, base) = \begin{cases} 1 + base & 1 \leq x \\ x + base & 0 < x < 1 \\ 0 & x < 0 \end{cases} \quad (5.3)$$

The *support unit* state is computed as follows:

$$support\_mem = prime + support\_inputs \cdot support\_weights \quad (5.4)$$

$$support = Esat(support\_mem - s\_threshold, s\_base) \quad (5.5)$$

where:

$support\_mem$  is the membrane potential of the support unit

$support$  is the firing rate of the support unit

$support\_inputs$  is a vector of inputs from other columns.

$support\_weights$  is a vector of corresponding connection strengths.

$s\_threshold$  is the unit's threshold.

$s\_base$  is the minimum firing rate when positive support is received.

$s\_threshold$  and  $p\_base$  are chosen such that the support unit can only fire if 1) a priming input is received and 2) the input from the support connections is at least non-negative.

The state of the *signal unit* is computed accordingly:

$$signal\_mem = signal\_inputs \cdot signal\_weights \quad (5.6)$$

$$signal = NSLramp(signal\_mem - sig\_threshold, sig\_base) \quad (5.7)$$

where:

$signal\_mem$  is the membrane potential of the signal unit.

*signal* is the firing rate of the signal unit.

*signal\_inputs* is a vector of inputs from sensory inputs or other columns.

*signal\_weights* is a vector of corresponding connection strengths.

*sig\_threshold* is the threshold of the signal unit.

*sig\_base* is the minimum firing rate when a non-zero signal is received.

$$\text{NSLramp}(x, \text{base}) = \begin{cases} x + \text{base} & 0 < x \\ 0 & \text{otherwise} \end{cases} \quad (5.8)$$

The output state of the column is computed according to:

$$\tau \frac{d \text{output\_mem}}{dt} = -\text{output\_mem} + \text{support} + \text{signal} \quad (5.9)$$

$$\text{output} = \text{thresexp}(\text{output\_mem} - o\_thresh, o\_gain) \quad (5.10)$$

where:

*output\_mem* is the membrane potential of the column's output unit.

*output* is the firing rate of the output unit. It is this value that is connected back into other columns.

*o\_thresh* is the output unit threshold.

$$\text{thresexp}(x, \text{gain}) = \begin{cases} x^{\text{gain}} & 0 < x \\ 0 & \text{otherwise} \end{cases} \quad (5.11)$$

*o\_gain* determines the steepness of the activation curve. In the model presented here, the two values used (depending upon the region) are 0.45 and 1.

*o\_thresh*, *sig\_base*, and *s\_base* are set such that both the support unit and the signal units must both be firing in order for the output unit to fire.

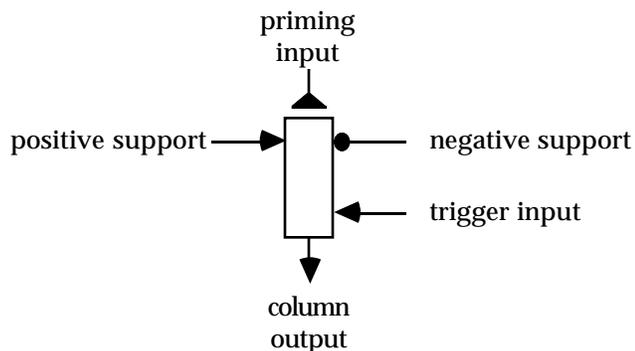


Figure 5.2: Convention used for drawing columns.

## Connectivity of Regions

In this section, we define the representation of information that is used in each region and how the columns within the different regions are connected together. In what follows, columns are drawn as rectangles, as shown in Figure 5.2. Priming inputs are drawn as entering the top of the rectangle. Positive and negative support inputs enter from the sides of the upper half of the rectangle; signal inputs from the lower sides. Outputs from the column exit from the bottom.

### Object Representations in PIP and IT

The different object classes are represented in PIP by separate populations of neurons (Figure 5.3). Within each population, we use a unit that codes for the general recognition of the shape itself, as well as sub-populations of units to code for the object's parameters. The parametric information is coded using a *population code* (Georgopoulos, Ashe, Smyrnis, & Taira, 1992; Georgopoulos, Kalaska, Caminiti, & Massey, 1982). Each of the neurons responds maximally to a particular parameter value (referred to as a *preferred value*). The activity of a parametric neuron is related to the difference between its preferred value and the actual value being coded (in the case of this model, the activity level is related through a Gaussian function with a specified standard deviation). Note that as it is used here, the term *population code* is not as

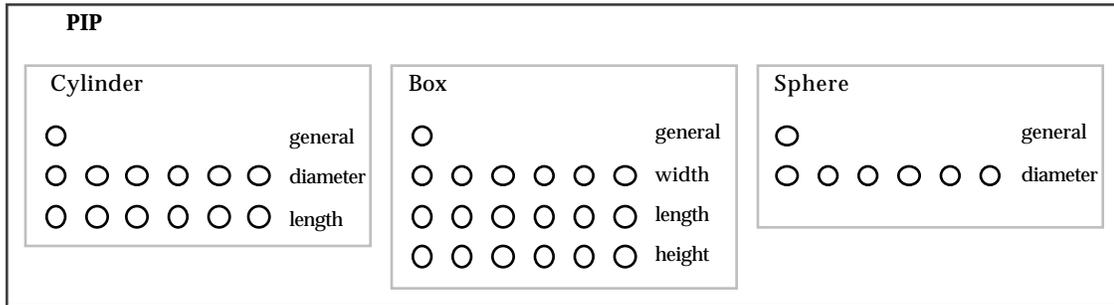


Figure 5.3: Illustration of PIP coding. Every object class is represented by a population of units. Sub-populations of units code for the object parameters.

strict as is used by Georgopoulos. Here, we relax the cosine tuning function requirement (replacing it with a Gaussian), and do not necessarily interpret the coded parameters in a peripersonal coordinate system.

Note also that each fundamental object class maintains its own set of parameter sub-populations. As a result, the diameter of a cylinder and the diameter of a sphere are coded using non-overlapping sets of cells. However, this is not to say that the space of all objects can be partitioned into discrete sets. Rather, we imagine these sets as defining a space of objects; objects that mix properties of two classes will have PIP representation that recruits from both sub-populations.

As is noted in Chapter 5, individual cells in IT appears to capture complex combinations of object features (such as shape and color of the object). This combined with the observation that object representations appear to be rather sparse (Logothetis & Pauls, 1995), implies that relatively different objects are coded by disjoint groups of neurons. In the model, we further simplify the coding by dropping the view-direction dependence (for rotations out of the plane) (Logothetis, et al., 1995), and assign a single unit to each object with which the model is "familiar."

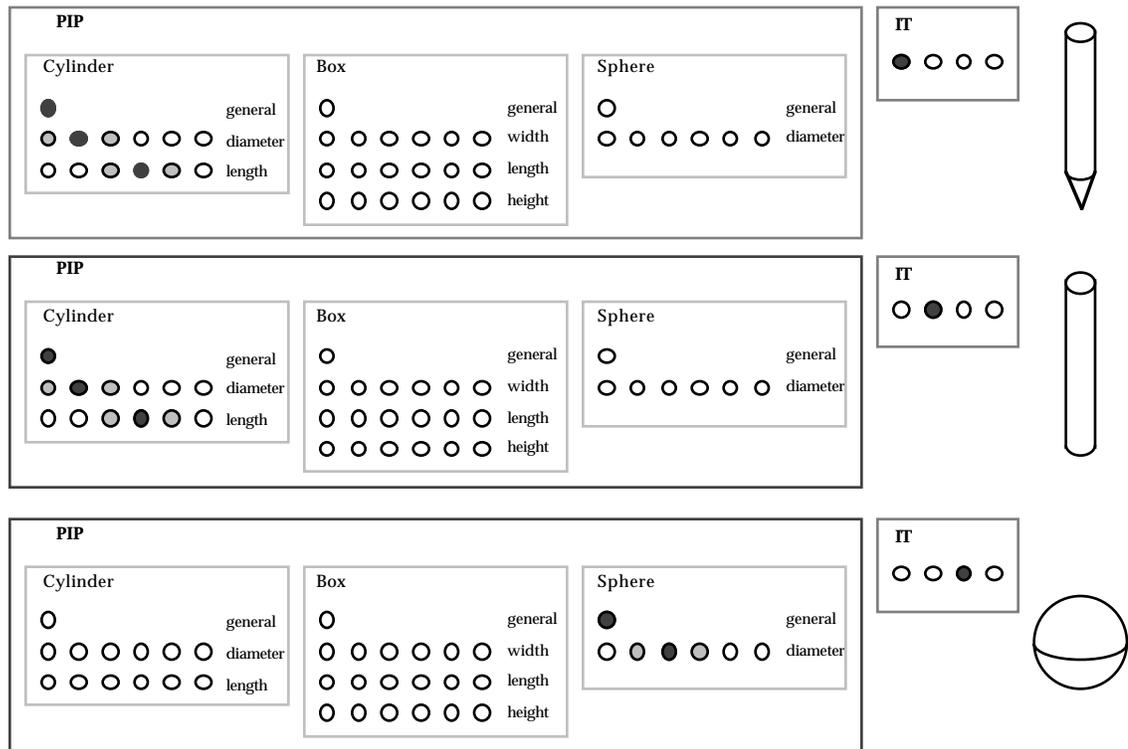


Figure 5.4: Example coding of three objects (pencil, stick, and a ball) in PIP and IT.

These two distinct representations are illustrated in Figure 5.4 for three different (familiar) objects. Note that because the pencil and stick are roughly the same shape and size, they are coded by PIP using identical activity patterns. However, the respective codes are orthogonal in IT. This simplification in the model captures the notion that these objects are *conceptually different* from one another. In the more general case, we would expect that IT would represent conceptual classes of objects in an overlapping manner (e.g., the set of all different mugs). Also, note in the Figure that the medium sphere is coded in a manner that is completely orthogonal (in both PIP and IT) to the stick and pencil.

## AIP

In this section, we will first discuss how individual AIP units are described behaviorally. Then, we will show how these behavioral descriptions give rise to projections from PIP and IT, as well as connections that implement cooperation and competition between different AIP cells.

### Classes of Neurons in AIP

AIP columns in the model are described according to the behavioral parameters defined in Table 5.1. As we will show below, these parameters determine how individual AIP columns are connected to each other and to columns in other regions. In the following sections, we will outline part of this connection process (for more details, see Appendix A).

### Projections from PIP

An affordance derived from PIP maps a very specific object configuration to one possible grasp for that object. Figure 5.5 illustrates two examples of such affordances. Column B (of AIP) is a general precision grasp unit. It is responsible for detecting a cylinder that is *both* narrow and short. The neurons in PIP that code for both of these parameter values project to the signal input of column B. When column B activates in response to the pair of inputs, it generates an output. The resulting signal provides positive support to column A, which is responsible for coding the aperture width for the precision grasp. In this case, the aperture width is derived from the diameter of the cylinder.

Likewise, column C detects wide and short cylinders, while column D captures the appropriate aperture width. Note that the behavioral specification (a precision grip) for columns B and C could be identical. The differentiation into the two different grasps (narrow precision versus wide precision) occurs because a single object configuration does not activate all appropriate AIP columns. Rather, only a randomly selected subset of matching (i.e., appropriate) columns will receive input from a single object configuration. It is important to

**visual:** [0,1]. Describes the degree to which this is a visual-related cell (Taira, et al., 1990). This parameter determines the probability and strength of connections from PIP and IT.

**motor:** [0,1]. Describes the degree to which this is a motor-related cell. This parameter controls the probability and strength of connections from F5

**phase:** {set, extension, flexion}. The phase (AIP) during which the cell begins to fire.

**grasp type:** {precision, power, side, and finger prehension}. The class of grasp in which this unit will participate.

**aperture:** either *none* or from the range [0,100mm]. Determines the preferred aperture for the column (if there is a preferred aperture).

Table 5.1: Description of AIP cell behavior.

note, however, that it is still possible for multiple object configurations to activate the same column.

The necessary conditions for connecting an object configuration to a non-specific column (no aperture specificity) are (referring to the behavioral description of the column):

1. An affordance exists that maps the object configuration to a grasp of the column's grasp type.
2. The column is a visual-type unit (the degree of **visual** response is non-zero).
3. The column's phasic response is of type *set*.

These conditions are also necessary for establishing a connection from the PIP cell that determines the grasp aperture to an AIP column that codes the aperture width. In both of the above examples, the opposition vector is perpendicular to the axis of the cylinder, so it is the cylinder's diameter that is mapped to the aperture width. If it were the case that the opposition vector were parallel to the cylinder's axis, then the length would determine the grasp aperture.

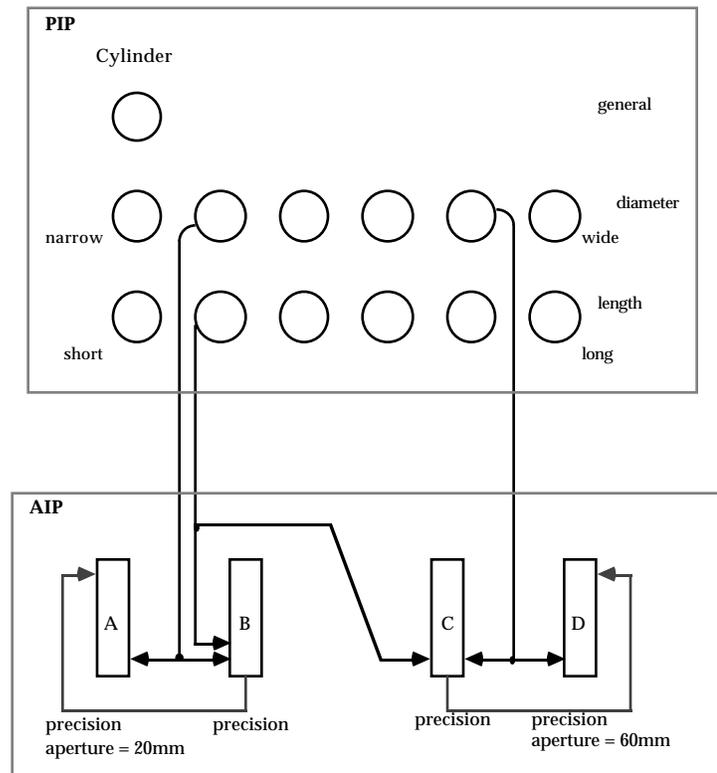


Figure 5.5: Affordance mapping from PIP to AIP. Two affordances are illustrated: narrow and wide cylinders that both map to precision grasps.

### Associations from IT

The mapping from object identity in IT to AIP is somewhat simpler than the mapping from PIP, since unique objects (and hence object configurations) already give rise to non-overlapping activity patterns in IT (Figure 5.6). As a result, the identification of the object maps directly to both the grasp type (e.g. column B) and the aperture of grasp (column A). In practice, this projection is important for biasing particular grasp configurations as a function of the specific context presented by the object. However, in the case of A.T., where no information is available from the dorsal stream (PIP), the projection from IT can provide (in some cases) the necessary grasp type and parameters.

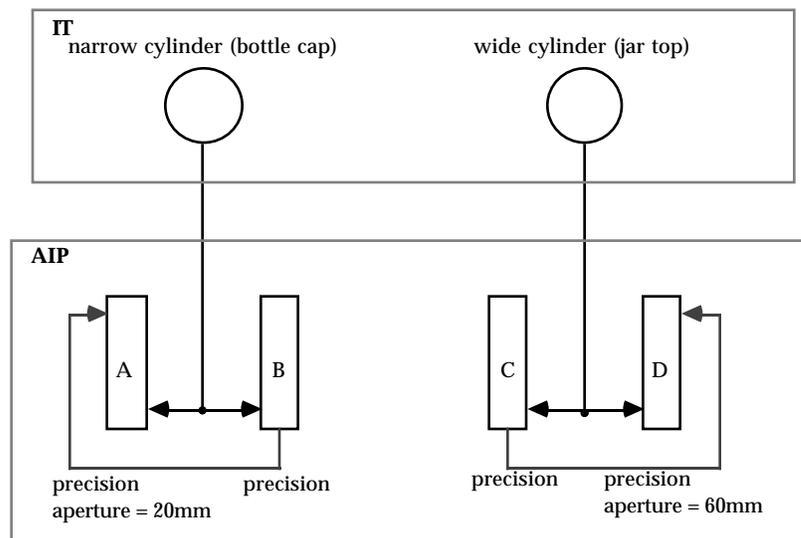


Figure 5.6: Affordance mapping from IT to AIP. The bottle cap activates a precision grasp with a narrow aperture; the jar top maps to a precision grasp with a wide aperture.

### Schema Cooperation/Competition Via the Basal Ganglia

The resolution process between different affordances is handled by local interactions in AIP and loops through the Basal Ganglia (as we will see, F5 ultimately is responsible for resolution between the different grasps). Columns that are members of the same schema ensure that they are coactive by exchanging positive supporting connections (e.g., columns B and C in Figure 5.7). Note that columns B and A participate in slightly different schema instances (aperture of 20mm as opposed to 25mm, respectively). However, because they are similar enough, a small positive supporting connection is established. This connectivity between columns of similar schema instances supports the formation of a population code for the representation of the parameters (in this case, the aperture width).

Competition between different schemas is implemented through an inhibitory loop through the Basal Ganglia (BG). All AIP cells project to this region of the BG, which returns an inhibitory signal to all columns in AIP (not all connections are shown in the figure). This mechanism is similar to the maximum selector circuit of (Didday, 1976). The indirect

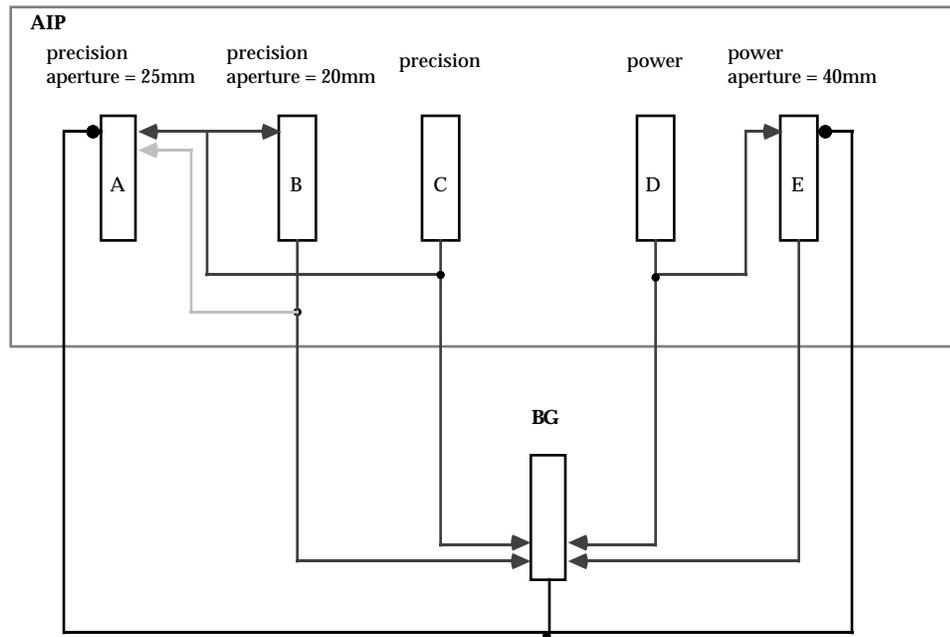


Figure 5.7: Cooperation and competition in AIP. Cooperating columns in AIP exchange positive supporting connections. Columns participating in different schemas inhibit one-another via a loop through the Basal Ganglia.

inhibitory signal due to the activation of column C upon columns A and B is counteracted by the direct excitatory signals.

## F5

In this section, we detail at the column level the interactions that give rise to grasp schema selection, competition/cooperation, and sequencing within F5.

### Representation of Grasp

As with AIP, each column in F5 is assigned a behavioral description. An F5 behavioral description consists of a small set of *grasp descriptions*, each of which are specified as described in Table 5.2. By allowing the behavior of a single column to be determined by multiple grasp descriptions, we are able to address the observations made by Rizzolatti that individual F5 cells can participate by varying degrees in different grasps (Rizzolatti, 1995). However, in the

**grasp type:** {precision, power, side, and finger prehension}. The class of grasp in which this column will participate.

**aperture:** either *none* or from the range [0,100mm]. Determines the preferred aperture for the column (if there is a preferred aperture).

**grasp response:** [0,1]. Degree of participation of this column when executing the specified grasp type.

**phase response** [{set, extension, flexion, hold, release}]: [0,1]. A vector that determines the relative response level (from phase to phase) of the column during execution of the specified grasp. Typically only a small subset of the vector elements are non-zero (1-2).

Table 5.2: F5 grasp descriptor. Description of F5 cell behavior in the model is done by specifying one or more grasp descriptors.

discussion that follows, the scheme of connectivity will make reference to an F5 behavioral description as if it consisted of a single grasp description.

Note that at this time, no data has been reported on F5 aperture size encoding. This aspect is necessarily included so that the model is able to preshape and grasp appropriately for objects of various sizes. The choice of the population encoding scheme (Georgopoulos, et al., 1992) has been made as a generalization of observations made in AIP and PIP (Murata, et al., 1993; Sakata, 1994). We will return to this point in the following Chapter.

### **Schema Priming by F6 and Cooperation/Competition Via the BG**

As discussed in Chapter 5, the setup and execution of the grasping program is initiated by cells in F6. This is implemented in the model by global priming connections from specific F6 cells (which are active for the duration of the reaching/grasping/holding task) to all columns in F4 and F5 (Figure 5.8). The priming signal serves to prepare all columns in the F5 layer (and hence all grasping schemas) for possible participation in the coming motor program, and does not specify or bias particular grasps. Biases for specific grasp schemas are received from AIP and area 46, and as will be discussed below, provide excitatory supporting inputs into subsets of F5 columns.

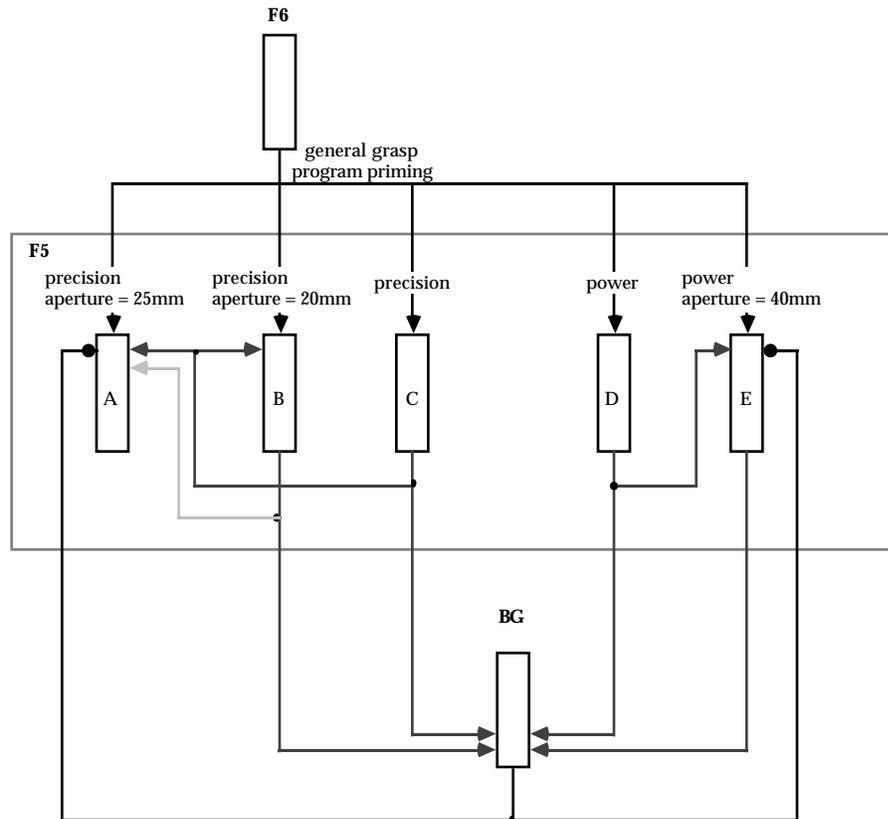


Figure 5.8: Priming of the general grasping schema, and selection of a specific grasp to execute. The selection process involves a cooperation/competition mechanism. Cooperation is implemented through local excitatory support between columns; competition is implemented via an inhibitory loop through the Basal Ganglia (BG).

Once biasing information is made available from AIP and area 46, one responsibility of F5 is to select a single grasp for execution. This is accomplished through a cooperation/competition mechanism similar to that which is found in AIP. Competition is implemented through an inhibitory loop involving BG (providing inhibitory connections to the support inputs of the F5 columns). Excitatory cross-connections between columns that participate in the same schema (also impinging on the support inputs of the columns) encourage these columns to be coactive. As in AIP, columns participating in similar schema instances (e.g., columns A and B) may also support one-another. Unlike AIP, however, excitatory supporting

connections are limited to columns that participate in overlapping phases of the grasping program.

### **Task Specific Biasing of the Selected Grasp**

The task that the system is preparing to execute after grasping the target object can (and often does) affect the selection of the grasp. In the model, high-level task information is stored in area F6. Task-related biases are imposed upon F5 via connections through the BG (Figure 5.9). This biasing is done at the level of different classes of grasp (e.g., precision pinch versus power grasp), and does not include parametric information. Connections from F6 provide positive support to columns in F5 that are involved in the *set* phase of the grasping program. In addition, connections are limited to those columns in F5 that do not specify movement parameters (e.g., aperture). Hence, the model relies on visual input to fill in the parameters. We imagine that in tasks involving movement in the dark (requiring the subject to remember the object to be grasped), area 46 would provide both grasp and parametric information.

Abstract instruction stimuli, as processed by area F2 can also be used to bias the selection of the type of grasp.

### **Interaction between F5 and AIP**

Figure 5.10 illustrates the relationship between columns (in AIP) belonging to a single affordance, and the columns in F5 that represent the corresponding grasp schema. The columns in AIP as a group are responsible for continually providing excitatory supporting inputs to the F5 columns. This implements an active memory that ensures that once selected the entire grasp program (involving all phases of movement) will be executed.

In the Figure, the AIP columns (A-D) differ in the type of signal input that they receive. Column A does not receive any recurrent inputs from F5, and hence is considered a visual-related unit (visual inputs from PIP and IT not shown). Columns B-D receive some amount of recurrent activity from F5, and are thus considered motor-related units (or at least have a motor-related

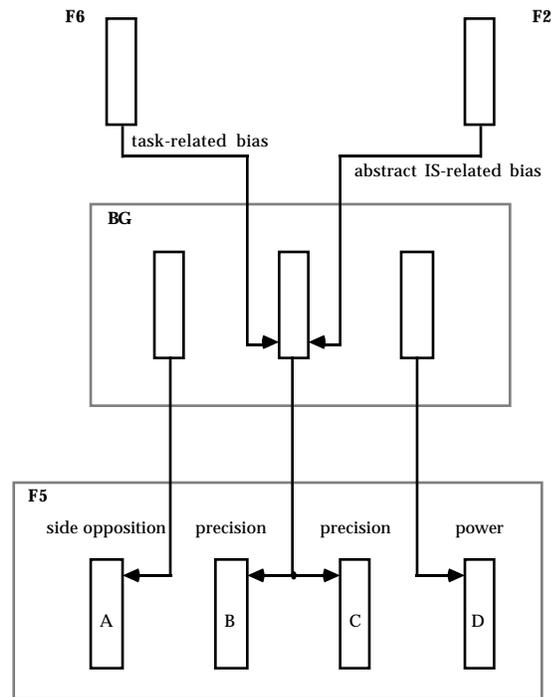


Figure 5.9: F6 and F2 biasing of the grasp schema selection process.

component). However, note that each of these three columns receive inputs from different subsets of the F5 columns. The subset is determined by the phase entry of the AIP column's behavioral description. Specifically, column B (which is a set-related unit) receives projections from the F5 set, extension, flexion, and hold units. As a result of these connections, this AIP cell will begin to fire (in response to recurrent motor signals) during the set phase of the task.

In contrast, column C (an extension-related unit) only receives recurrent activity from the F5 extension, flexion, and hold units, and thus will only become active (disregarding any visual components) once execution of the movement begins. The F5 flexion and hold units are the only phases to project to column D, which is a flexion-related unit.

Under normal circumstances, the recurrent loop from F5 back to AIP serves to update the AIP state in a feed-forward manner (i.e., AIP receives this information before sensory feedback from the descending motor command can return). However, when a perturbation occurs, this is

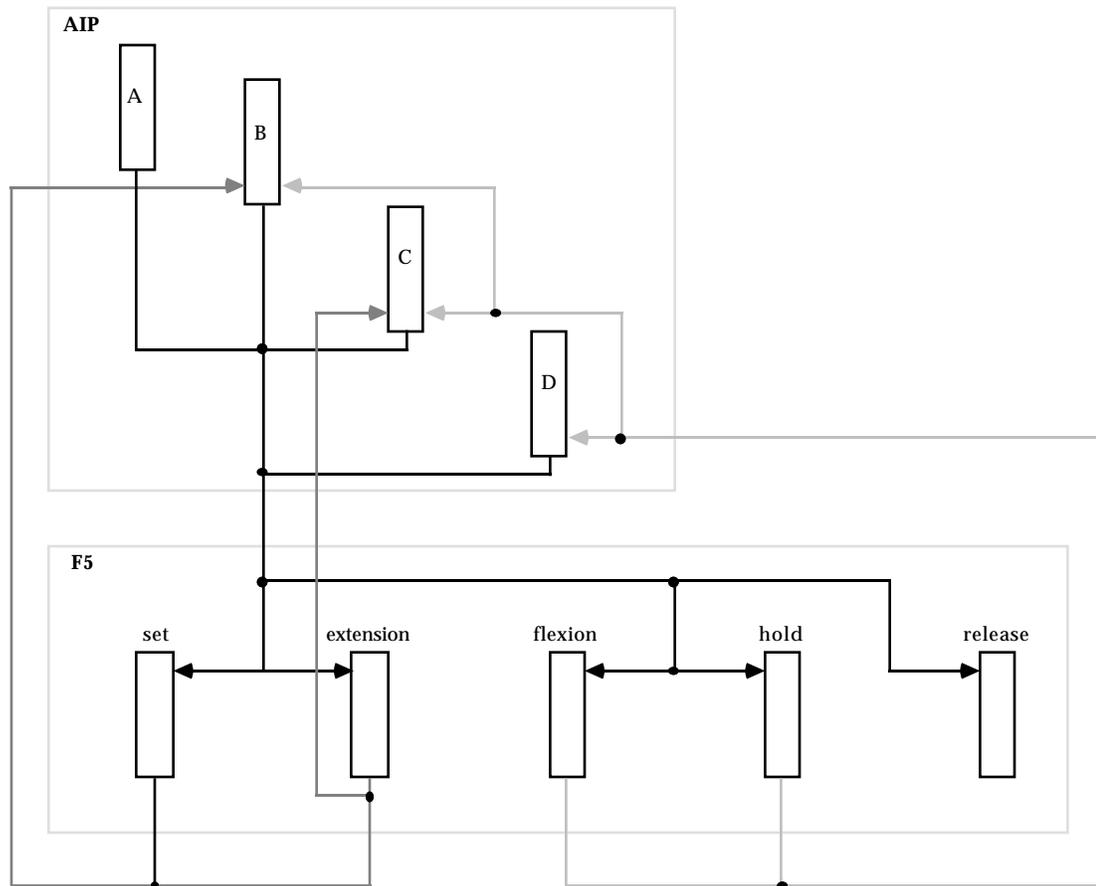


Figure 5.10: Interaction between AIP and F5 during execution of the reaching/grasping/holding task. Column A is visual-related; columns B-D are motor-related.

detected by F5 after some delay. This new state information is then passed up to AIP (and hence in this case, AIP receives feedback information).

The necessary condition that a group of AIP and F5 columns be connected in the manner illustrated in Figure 5.10 is that there must be a match in the grasp that they specify. This implies that not only must the grasp types be the same, but the parameters of the grasp must be similar, if they are specified. For example, AIP columns that do not specify a grasp aperture will only be interconnected with F5 columns that also do not specify an aperture. Likewise, AIP columns that encode an aperture of X will only be connected to F5 columns of an aperture similar

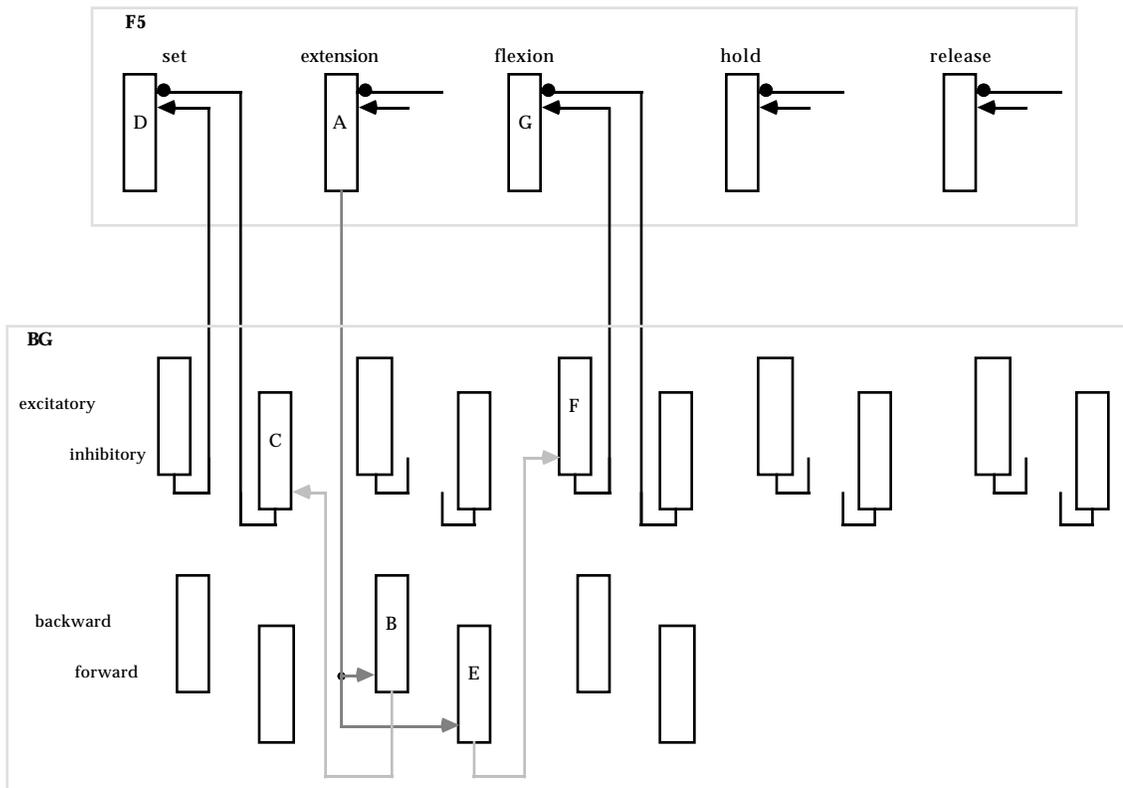


Figure 5.11: F5 interaction with BG implementing the cascade of activity in F5 as execution of the motor program progresses. F5 reports to BG its current phase; BG in turn inhibits the previous phase and prepares the next phase of movement for execution.

to X. The degree of similarity will determine the likelihood and the strength of the interconnections.

### Phasing Via BG

In the previous chapter, we demonstrated the functional connectivity that is necessary to implement the cascade of activity that flows through a set of F5 units as execution of the reach/grasp/hold motor program progresses. In this section, we show the details of the implementation of this cascade.

A subregion of BG (separate from the subregions discussed earlier that are responsible for competition and grasp biasing) is used in the model to implement phase-specific connectivity in

a manner that is independent of the grasp that is actually being executed (hence we are separating the component schemas from the mechanism that implements the *coordinated control program*). At each stage of the motor program, F5 reports its current phase to the Basal Ganglia. In turn, the BG positively supports all F5 columns that are involved in the next phase of movement. In addition, F5 columns that code for the previous phase receive negative support, causing them to be shunted and enforcing a phase transition.

The set of connections required to implement one step in this process are demonstrated in Figure 5.11. F5 column A is an extension unit. When it becomes active, it sends an excitatory signal to columns B and E (the backward/forward columns, respectively) in the BG. Column B projects to the *inhibitory* column (C) that corresponds to the previous (set) phase of the motor program. This column, in turn inhibits column D, the set-related unit in F5. Consequently, column D is shunted and the phase transition from set to extension is complete. In parallel with this process, the active column E projects to column F (the excitatory unit of the flexion phase), which then provides positive support to F5 column G (flexion unit). The result is that column G is now prepared for activation.

In the above description, there is no functional distinction between the backward/forward columns in BG. The situation, however, becomes more complicated when F5 columns are allowed to participate in multiple phases. Figure 5.12 illustrates one example in which an F5 column (A) is active during both the extension and flexion phases of movement. Because this column first becomes active during the extension phase, it receives positive support from the BG excitatory unit H. When the column becomes active in response to a trigger stimulus (in this case the go signal; connection not shown) it activates the BG backward column B. As in the above case, column B then activates unit C, which in turn shunts F5 column D.

The last phase in which F5 column A participates is the flexion phase. It is therefore responsible for preparing for the hold phase of the program. It accomplishes this by activating the BG forward unit E (which corresponds to the flexion phase). Unit E activates the BG

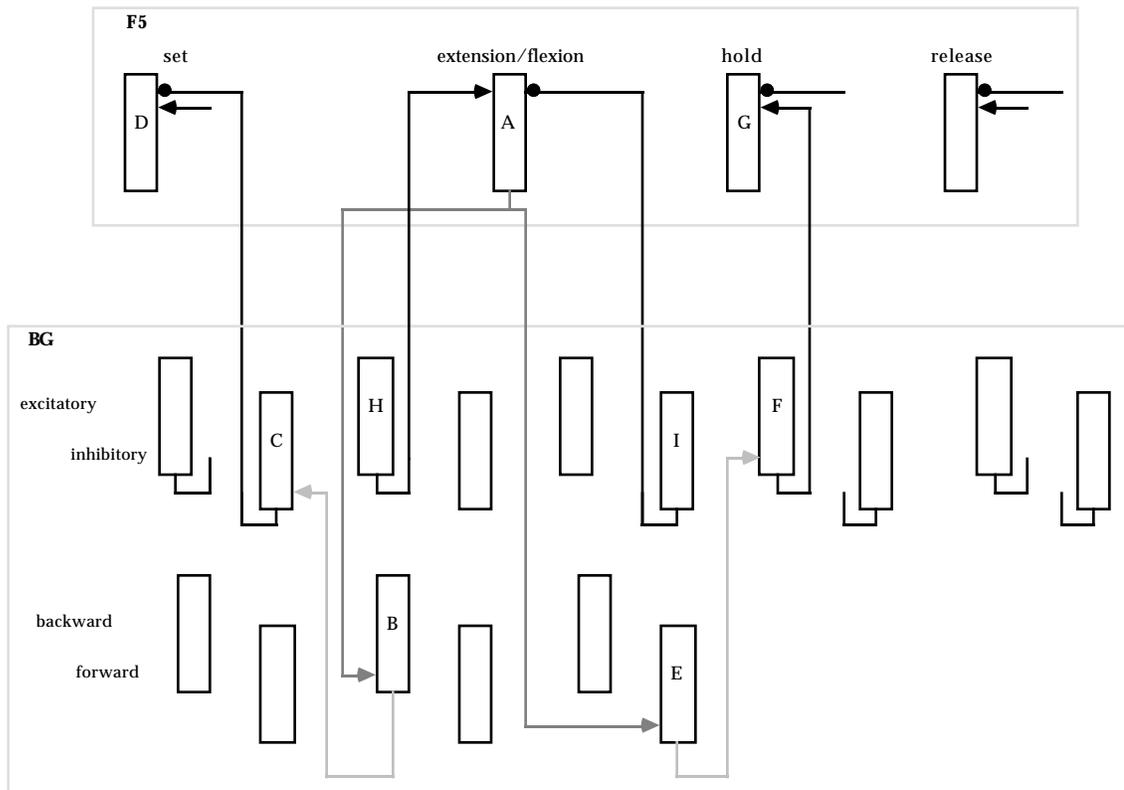


Figure 5.12: BG phasing when an F5 unit is active during both the extension and flexion phases of the reach/grasp/hold program.

excitatory unit F, which then positively supports the F5 hold column (G). Once contact with the target object is established, column G will activate, and (through connections not shown) will indirectly excite BG unit I, which then shunts column A.

It is important to note that all phase transition information is contained within the Basal Ganglia, and that this information is independent of the grasp that is being executed in F5. Specifically, the critical information is stored in the projections from the backward/forward columns to the excitatory/inhibitory columns (we refer to these as the *phasic projections*). We imagine that in the more general case, where the sequence of phases may be different depending upon the task that is being performed, that it would be possible to rearrange these phasic projections. The arrangement of these projections would be specifiable by high-level regions (possibly including F6 and area 46).

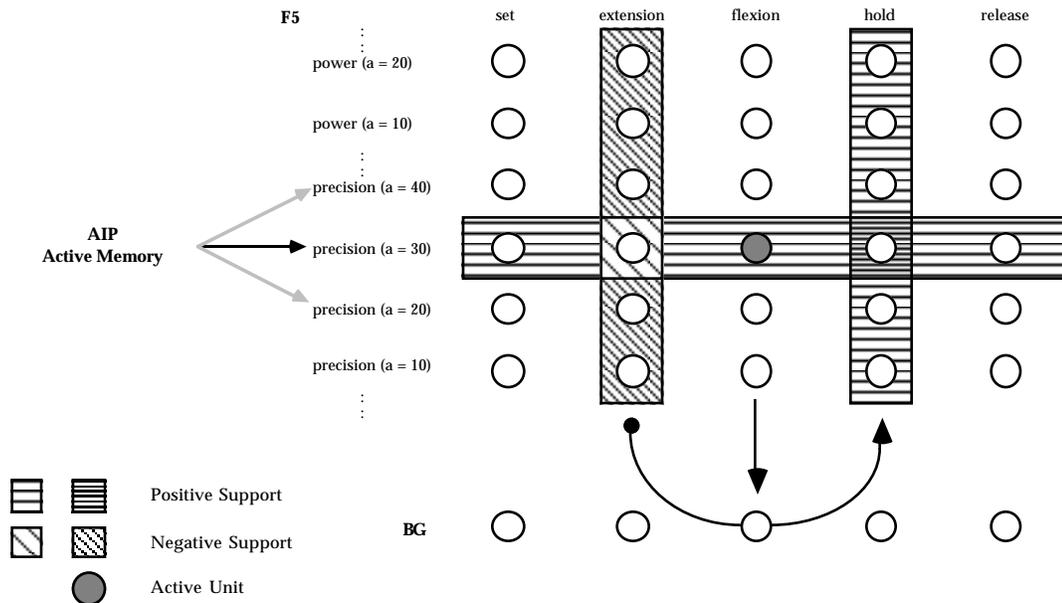


Figure 5.13: Abstract representation of the interaction between the AIP active memory and the BG phasing loops. AIP provides a detailed specification of the motor program; BG is responsible for executing the sequence of steps in the specified program.

### Interaction of the AIP Active Memory and the BG Phasing Systems

In the previous sections, we have illustrated how F5 independently interacts with the AIP active memory and the BG phasing systems. In this section, we demonstrate how these two loops interact with one-another in order to execute a sequence of steps belonging to a specific grasping motor program.

Figure 5.13 illustrates at an abstract level the interaction of these two loops. The matrix of cells in F5 represent a set of grasp/hold motor programs; each program is horizontally arrayed. For clarity, we are explicitly representing many possible grasp type and aperture combinations. In addition, we are discretely representing every phase of the programs (phases are aligned vertically).

At the instant depicted in the figure (the flexion phase), the network is executing a precision grasp with an aperture of 30 mm. This state is captured by the activation of the

precision/aperture = 30/flexion unit at the center of the F5 matrix. The AIP active memory maintains a record of one of the affordances that corresponds to the precision/aperture=30 grasp. This active memory continually provides positive support to all F5 units in the precision/aperture=30 grasp program (as depicted by the shaded horizontal box). In addition, due to the population coding scheme, this active memory also supports (to a smaller degree) two other precision grasp programs (aperture = 20,40).

At the same time, the Basal Ganglia circuit maintains a representation of the current motor program phase (in this case flexion). It responds by sending an inhibitory projection to all F5 cells that are involved in the extension phase of movement (left shaded column of F5 units). In addition, this BG unit provides positive support to all F5 units that code for the hold phase of movement (right-hand shaded column).

When contact with the target object is finally established, several flexion cells receive the signal input from SII indicating the contact event. Because the F5 precision/aperture=20/hold cell is receiving the highest amount of positive support (and assuming that the object is the one that is expected), the cell very quickly activates and wins the competition with other flexion cells. We then see a phase transition to the flexion phase of the precision/aperture=20 motor program.

In general, the convergence of the detailed motor program, specified by the AIP active memory, and the positive support for the next phase of movement from BG allows for the rapid transition to the next phase of movement when the triggering event is detected. However, note that if an unexpected hold triggering event is detected (for the example described by the figure), the other flexion cells (not corresponding to the current motor program) still receive some degree of positive support from the BG. As a result, not only do we see a phase transition in the F5 matrix, but we also can see a shift in the motor program (to a different row in the matrix). Note that this motor program shift is most likely to happen when the hold triggering event corresponds to a motor program (row in the matrix) that is similar to the currently

executing motor program, because it is already receiving a small amount of positive support from the AIP active memory. We will return to this issue in the following chapter.

## **Grasp Execution**

Layer F5 manages the execution of the reach/grasp/hold task through each phase of the motor program. During each phase of movement, F5 recruits 1) a set of motor assemblies in MI that are responsible for actuation of the individual joints; and 2) a set of SII units that monitor the global state of the hand. In this section, we describe in detail the representation of information in each of the key regions, and the neural circuitry that is involved in the low-level execution of the motor program.

### **SI and SII**

The primary somatosensory cortex (SI) represents both the proprioceptive and tactile state of the hand at the level of individual joints and tactile pads, respectively. The proprioceptive information for each joint (the joint position) is represented using a population coding scheme<sup>1</sup> (Figure 5.14). The tactile information, coded in a separate SI channel, represents the force detected by a set of tactile pads located on the fingers and thumb. Force is coded by the firing rate of these units. For each finger pad, one unit increases its firing rate as the force is increased; a second unit reduces its firing rate with the increasing force.

Columns in SII integrate the local state information represented in SI in order to compute an estimate of the global state of the hand. Individual columns in SII receive inputs from many different proprioceptive arrays (column A), and can also integrate tactile input from selected pads (column B). Because these columns have a high threshold, they only respond to very specific configurations of joint position and finger pad state. Hence, they are referred to as

---

<sup>1</sup> Although a more appropriate representation would use firing-rate to encode the joint position (Kandel, et al., 1991), the population coding scheme reduces the number of connections and the computational complexity of the model implementation.

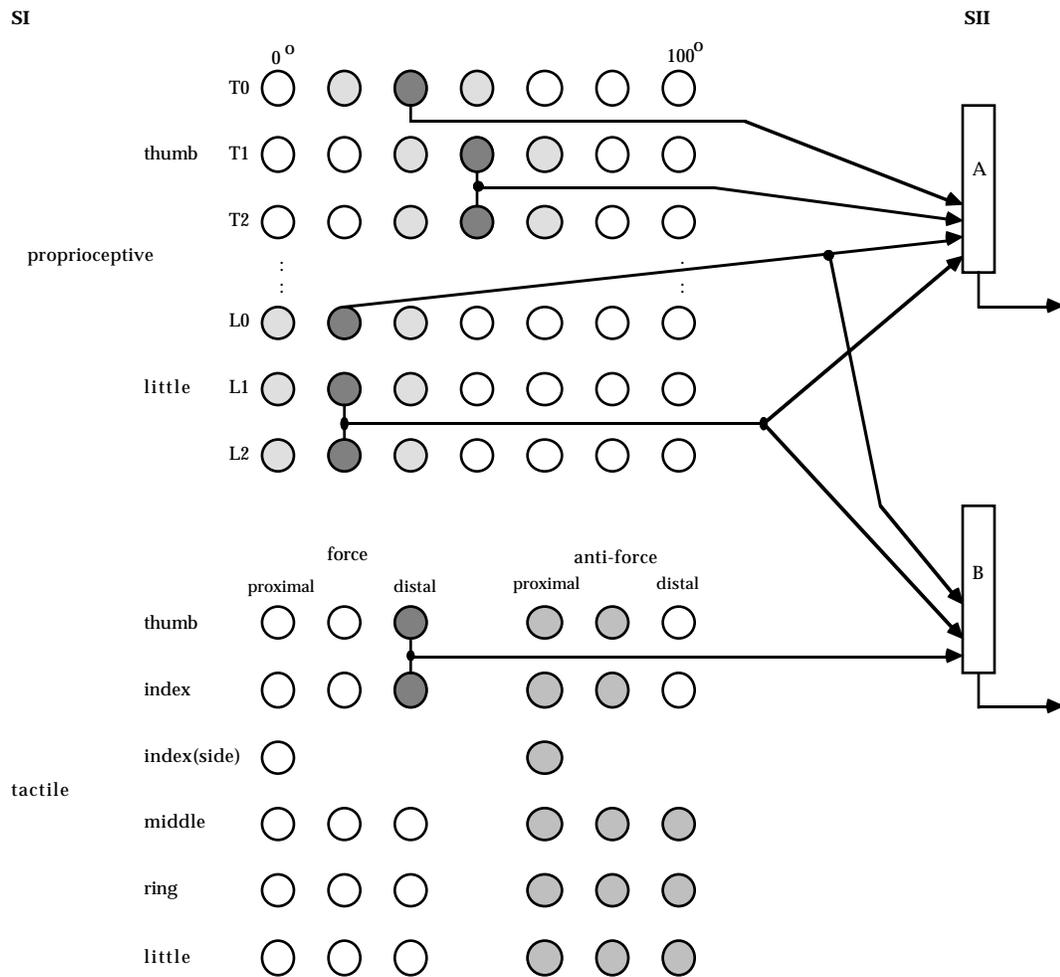


Figure 5.14: Representation of hand state information in the somatosensory cortices (SI and SII). SI carries two channels of information: proprioceptive (or joint position) and tactile. Each individual cell in SI has a small receptive field, and hence only captures a limited view of the hand state. SII columns detect hyperfeatures of SI inputs, thus representing a global view of the hand.

*hyperfeature detectors.* Column B, for example, combines joint position information with tactile inputs from the distal pads of the thumb and index finger. Hence the column not only is capable of detecting a contact event, but it is a contact event with an object of a certain width.

## MI

A column in the primary motor cortex of the model is intended to capture the properties of a set of motor cortex neurons, projecting to both agonist and antagonist muscles that are responsible

for actuating a subset of joints in the hand. In the model, we simplify this interaction by assuming that stimulation of a single MI column specifies a joint equilibrium position for the joints to which it is connected (Giszter, Mussaivaldi, & Bizzi, 1993; Shadmehr, 1993; Shadmehr, Mussaivaldi, & Bizzi, 1993). In addition, the degree of stimulation affects the joint stiffness. We use  $\hat{\theta}_{ij}$  to denote the contribution that MI unit  $i$  has on the equilibrium position of joint  $j$ , and  $S_{ij}$  to denote the stiffness contribution ( $0 \leq S_{ij} \leq 1$ ). When there is no physical connection from unit  $i$  to joint  $j$ ,  $S_{ij} = 0$ . When multiple MI units are active, their contributions are summed, weighted by their activity level:

$$S_j = \sum_i a_i * S_{ij} \quad (5.12)$$

$$\hat{\theta}_j = \begin{cases} \frac{\sum_i a_i * \hat{\theta}_{ij} * S_{ij}}{S_j} & \text{if } S_j > \epsilon \\ \tilde{\theta}_j & \text{otherwise} \end{cases} \quad (5.13)$$

where:

$a_i$  is the activity level of MI column  $i$ .

$\tilde{\theta}_j$  is the default equilibrium position for joint  $j$ .

Finally, the joint actuator is modeled as a revolute spring (derived from (Katayama, 1993)):

$$\tau_j = K(S_j)(\hat{\theta}_j - \theta_j) - B(S_j)\dot{\theta}_j \quad (5.14)$$

where:

$\tau_j$  is the torque applied to joint  $j$ .

$K( )$  and  $B( )$  are linear functions of  $S_j$ .

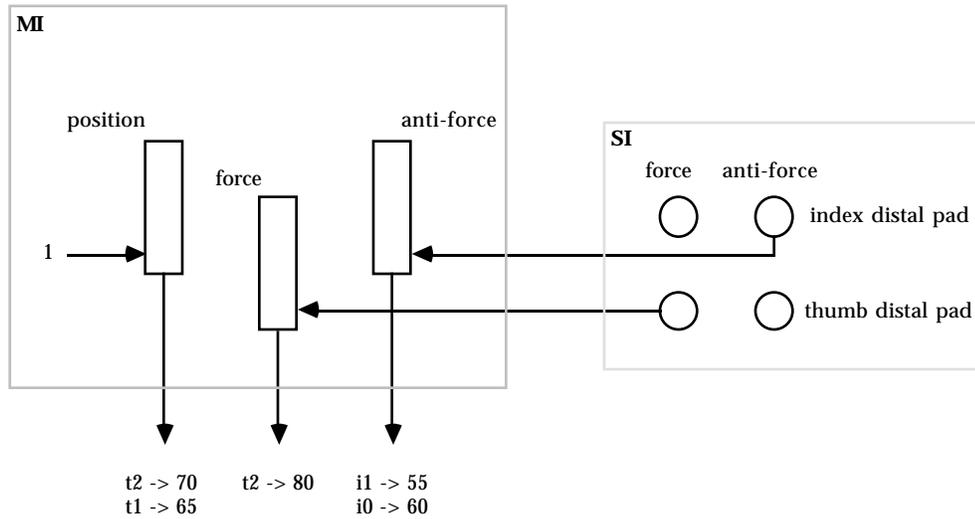


Figure 5.15: Three classes of MI columns. Force responsive columns receive projections from SI units that detect tactile stimulation; anti-force columns are sensitive to the absence of tactile input; position columns are receive a constant input.

### MI Activation

At any given instant during the execution of a grasping movement, the currently active F5 columns are responsible for selecting the set of MI columns that will generate the desired hand movement. This selection is done through excitatory projections from the F5 columns to the support inputs of the MI columns. Additional details of this mapping are given below.

MI columns fall into three distinct classes, depending upon the type of signal inputs that they receive. Inspired by observations of (Picard & Smith, 1992b), one class of columns receive inputs from one or more SI tactile units. These columns are referred to as being *force responsive* (Figure 5.15), because they become active when a force is detected in SI. The complement to this set comprise the *anti-force responsive* columns, which receive inputs from the corresponding SI cells. The final class of MI columns are the *position units*. In the model, these columns receive a constant input. Thus, the only condition for their activation is that they receive positive supporting inputs from the F5 layer.

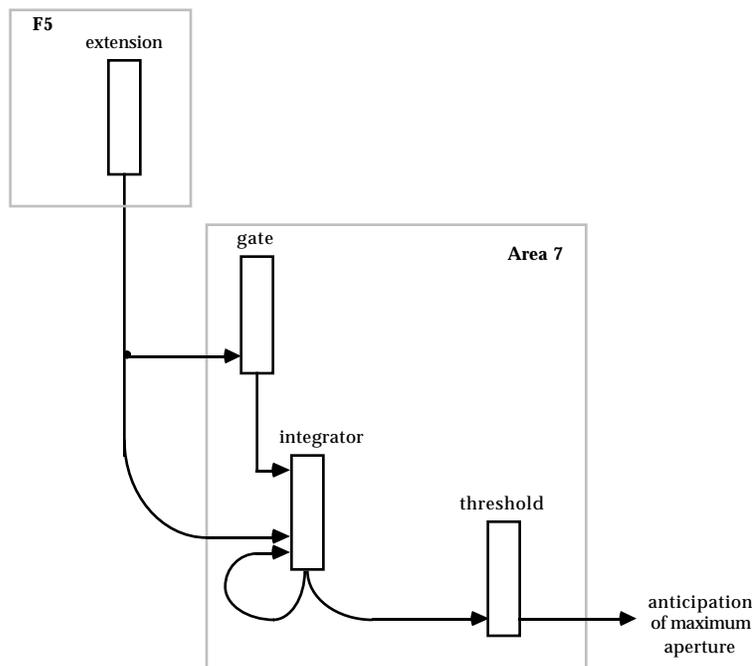


Figure 5.16: Internal model of grasp extension state.

### Internal Timing Model of Current Hand Extension

As described thus far, the model relies on sensory feedback to detect the point of maximum desired aperture during the extension phase of movement so that a transition can be made to the flexion phase of movement. However, during fast movements, the delay of this sensory information relative to the point at which it must be used is significant. In addition, deafferented human subjects are still able to perform preshape and enclose movements, although the movements tend to be slower and more variable than in the afferented case (Gentilucci, Toni, Chieffi, & Pavesi, 1994).

This difficulty is approached in the model through the introduction of an internal model of hand state (Hoff & Arbib, 1993). The internal model makes use of descending motor commands (from F5) in order to estimate the point in time at which the maximum aperture is achieved. This is accomplished by tracking the amount of time that F5 spends in the extension phase of movement. A simple implementation of this timing mechanism is demonstrated in Figure 5.16.

Initiation of the extension phase of movement in F5 activates the *gate column* in the internal model (area 7). Positive support to the *integrator column* initiates the integration of the F5 input. Once this integral passes beyond a specified threshold, the internal model generates an output that can be used to trigger the transition from the extension phase of movement to the flexion phase.

Not shown in the figure is a calibration mechanism that utilizes SII inputs to appropriately set the threshold.

### **Full Execution Circuit**

Depending upon the current phase of the motor program, columns in F5 will recruit different configurations of the MI/SI/SII circuitry. Exclusively set-related columns are not involved in the direct generation of movement, and hence do not activate lower-level regions. Extension-related columns, activated by a triggering signal (Go), instantiate a goal position for the hand joints that corresponds to the maximum aperture of the preshape (Figure 5.17). This is accomplished by sending excitatory signals to the support inputs of a selected set of MI columns in the position pool. The relative strengths of these projections are balanced in such a way that the specified joint equilibrium position is the point of maximum aperture. However, a single F5 column will generally only project to a subset of the hand joints (typically the joints of 1-3 fingers). Hence, a population of F5 columns is necessary to specify the goal state of the entire hand.

Besides driving units in MI, the F5 extension columns also positively support a subset of SII columns that will be involved in the detection of the point of maximum aperture. This detection process only examines the proprioceptive state of the hand (position). In addition, extension columns also activate the integration process in Area 7 that will be used to estimate the point at which maximum aperture will be achieved.

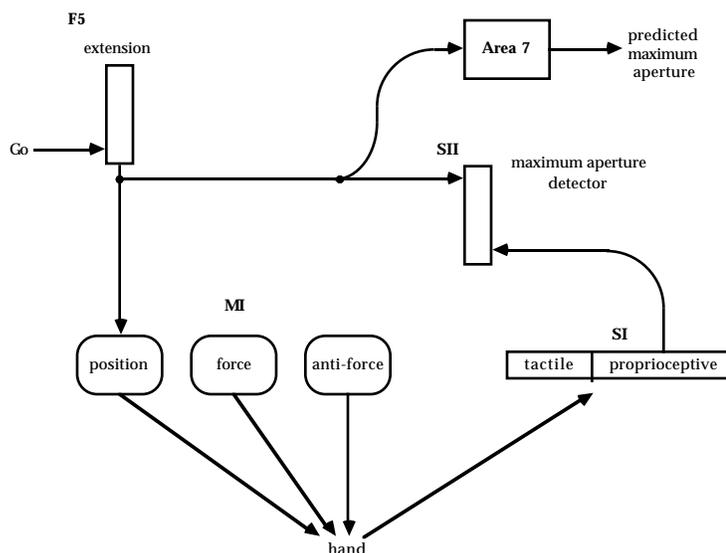


Figure 5.17: The low-level circuit involving F5 extension columns.

F5 flexion columns are activated by the detection of maximum aperture--by either SII or the internal model of area 7 (Figure 5.18). Not only does the flexion column drive the position pool of MI (corresponding to the point of anticipated contact with the target object), but also sets up SI/MI reflex loops (the MI anti-force pool) that will cause the hand to continue to close if contact is not detected at the anticipated position.

The F5 flexion units also provide positive support to SII columns that will detect the combination of hand configuration (proprioceptive state at the point of contact) and the actual contact with the target object (tactile inputs).

The F5 hold columns recruit units from all three MI pools (Figure 5.19). Those joints that are involved in applying a stabilizing force are controlled through a balance of force and anti-force pools. The remaining joints are controlled using the position pool.

Finally, the F5 release columns recruit motor units from the MI force pool to induce release of the object (Figure 20). In addition, SII columns are recruited to detect the point at which the release occurs.

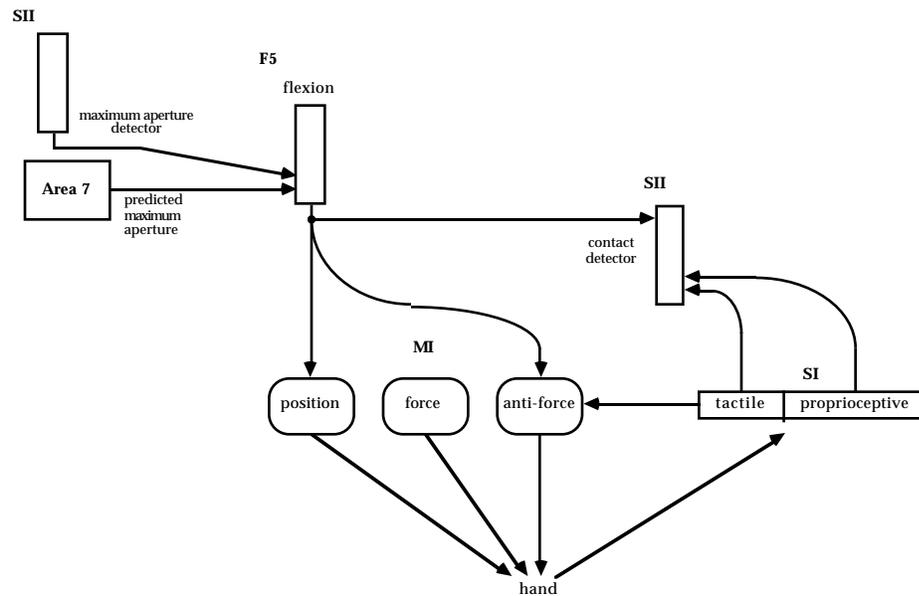


Figure 5.18: Low-level circuit involving flexion columns.

## Summary

This chapter has presented a detailed view of the grasping model. First, we introduced the concept of  $\mu$ -schema as a bridge from the language of schema theory to the neural implementation level. The  $\mu$ -schema maintains the schema-level operators of priming, cooperation/competition, and activation, while still being implemented using a small group of neurons (referred to as a column). In order to construct a schema, a set of  $\mu$ -schemas are bound together through the cooperation operator (implemented as excitatory connections between the columns). Different schemas can then interact in either a cooperative manner (via excitatory cross-connections between the member  $\mu$ -schemas), or in a competitive manner (through inhibitory cross-connections). The level of participation by an individual  $\mu$ -schema within a schema, however, can be graded (as determined by its activity level). In addition, two distinct schemas may share a single  $\mu$ -schema.

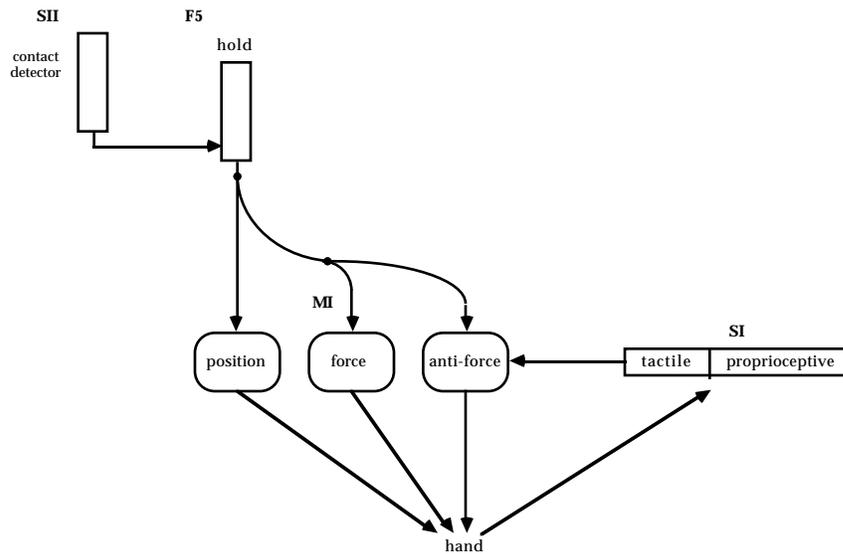


Figure 5.19: Implementation of the hold phase of the motor program.

Second, we demonstrated how information is represented in each layer in the model, and then how micro-circuits of columns involving multiple layers are constructed in order to achieve the desired computations. The primary model components are:

1. PIP represents object shape using non-overlapping pools of neurons. Within these pools, sub-populations of neurons encode the parameters of the object (making use of a population coding scheme).
2. IT represents different familiar objects using non-overlapping representations.
3. PIP and IT associate an object description with a set of affordances in AIP.
4. AIP associates an affordance with a corresponding grasp in F5.
5. During execution of the movement, AIP maintains an *active memory* of that affordance, reinforcing the choice of grasp in F5 in a manner that is independent of the phase of movement. The active memory may be updated by F5 in the face of perturbations.
6. The basal ganglia (BG) provide a mechanism for generating the sequence of phases that are observed in F5. This mechanism provides this functionality in a manner that is independent of the actual grasp that has been chosen.

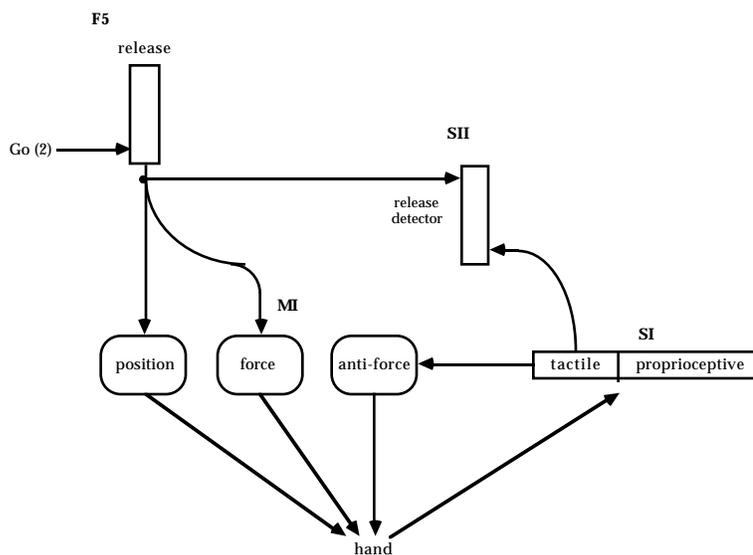


Figure 5.20: Low-level circuit for the release phase of movement.

7. The basal ganglia are also involved in implementing the schema competition mechanisms in F5 and AIP. In addition, task-related biases in the grasp decision process can be introduced by higher level systems (area 46, F6 or F2) via a pathway through the BG.

8. A single MI column specifies equilibrium positions for a subset of the joints of the hand. During specific phases of movement, F5 columns recruit a motor assembly of MI columns.

9. Units in SII provide context-dependent detection of the global hand state. This information is used by F5 to monitor the ongoing grasp so that it may react to unanticipated situations.

10. An internal model of grasp extension state is introduced as a mechanism for initiating the transition from the extension phase of movement to the flexion phase without relying on immediate sensory feedback information.

As we will see in the following Chapter, the functionality hypothesized for several regions (especially F5 and AIP) have several important implications, including:

1. The visual/motor associative process involving AIP affects how objects are encoded at the neural level in AIP.

2. The recurrent connections from F5 to AIP affect the motor-related activity in AIP.
3. The hypothesis that F5 is involved in the selection of a single grasp to execute implies certain patterns of activity when multiple grasps must be considered before a selection is made.

## Chapter 6: Model Simulation Results

---

In this chapter, we demonstrate the capabilities of the model that has been developed in the previous two chapters. First, we examine the key computational mechanisms in the model, focusing on the interaction between AIP and F5 as a grasp is selected and then executed. Second, we draw comparisons between single cell behavior observed in several monkey experiments and similar experiments performed with the model. Finally, a set of novel tasks is presented to the model to explore several important computational questions. Observations made at the single unit and population levels (in the model) serve as predictions for future monkey experiments.

---

### Introduction

In the experiments that follow, the protocol is derived from that used by Sakata (Taira, et al., 1990). We first illustrate the general behavior of the model using this protocol, and then present a series of new questions that are designed to examine more subtle issues of object and grasp encoding at the neural level. These include:

1. *Grasping of two different objects with an identical grasp.* How are the objects and grasp represented in AIP and F5?
2. *Grasping of a single object in one of two ways (as determined by an instruction stimulus).* How are these two different grasps represented, and how is one ultimately selected for execution?

3. *Grasping of a set of objects of varying size.* How do F5 and AIP represent the size of grasp aperture?
4. *Visual/motor perturbation.* How does the system respond when there is a mismatch between the assessed parameters of the object (by the visual system) and the actual parameters?
5. *Object parameter representation in AIP.* Are all parameters of an object represented equally in AIP?

Unless otherwise stated, all experimental results presented in this Chapter are derived from the model.

## Overview of Model Behavior

Figure 6.1 demonstrates the cascade of activity that flows through a set of simulated F5 cells as a precision grasp is executed. The trial starts with the presentation of a narrow cylinder to be grasped ( $t=0$ , event not shown). 700ms later, a *ready signal* is presented, indicating that the movement trial is about to begin. The system responds by preparing to execute a precision grasp (as encoded by a population of F5 set cells, of which one is shown in the figure). This preparation process provides positive support to F5 units that will be involved in the extension phase of movement (the support is provided indirectly through the BG). This state is maintained until the *Go signal* is given, at which point the F5 extension units begin to fire.

The activation of the extension-related units results in the following:

1. Via the basal ganglia, the set-related units are shunted (through negative support).
2. Also through the BG, the next phase of movement is prepared by providing positive support to those units that will be involved in the flexion phase.
3. The SII units that are responsible for detecting the maximum aperture for the current grasp are primed.

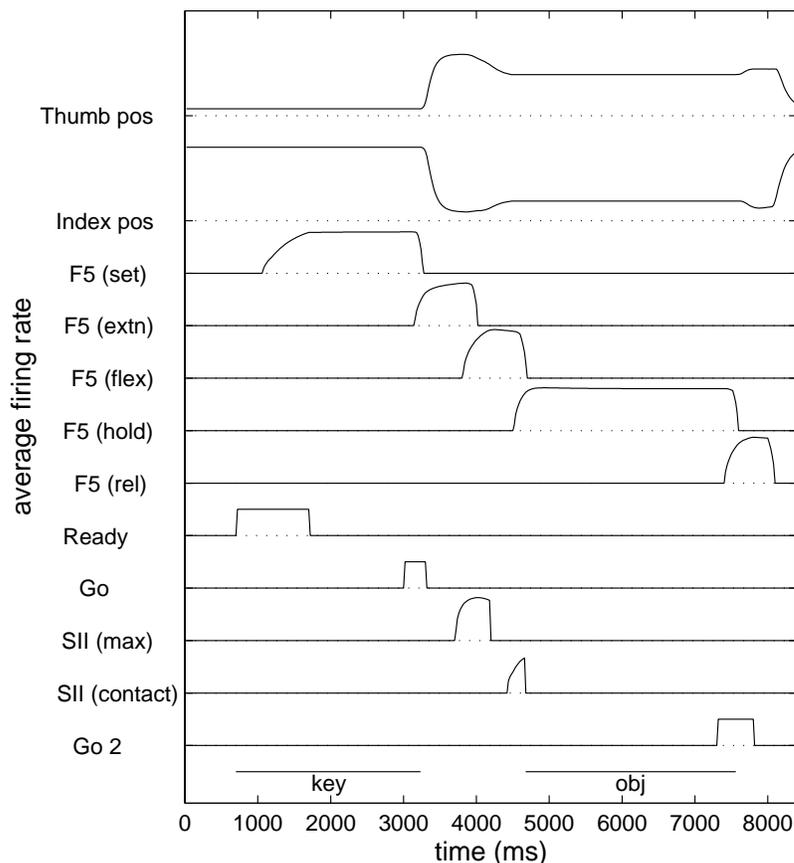


Figure 6.1: F5 activity during execution of a precision grasp. The top two traces show the position of the thumb and index finger, respectively. The next five traces represent the average firing rate of five F5 neurons (set-, extension-, flexion-, hold-, and release-related). The remaining five traces represent the various external (Ready, Go, Go2) and internal (SII) triggering signals.

4. The internal model of current hand state in area 7 is activated.
5. The index finger and thumb begin to open (as shown by the top two traces).

When the maximum aperture is achieved, as detected by either SII (labeled *SII max* in the figure) or by the internal model (not shown), a second phase transition is made. This phase transition is driven by the activation of the flexion-related units. The flexion units shunt the previous phase (extension), prepare for the next (hold) phase, prepare for detection of the contact event (by priming *SII contact*), and flex the thumb and index finger.

The contact event (establishment of the precision grasp in this case) forces a phase transition to the hold phase of the program. The network remains in this state until the secondary go stimulus is provided by the experimenter, at which point the release phase of the motor program is initiated, and the fingers are extended enough such that contact with the object is lost.

For comparison purposes, we also make use of the Sakata *key* and *object* phase notation (see figure). Here, the key phase is defined from the point at which the ready signal is received to the initiation of movement in response to the go signal. The object phase is defined by the point of contact with the object and the point at which the release movement is initiated.

Figure 6.2 demonstrates the relative time course of two AIP cells relative to F5 activity. The trial and displayed F5 unit traces are the same as in the previous figure. The two AIP cells activate after the F5 set unit becomes active, but both begin to fire within the key phase of the task. Once they begin to fire, note that they continue to do so until the task is completed. Note also that both AIP units show peak activity during the extension/flexion phases of the task, as is observed by (Taira, et al., 1990). We explain this property as follows:

Suppose that for a given F5/AIP pair of units that participate together in a grasping program, there is some constant probability that a projection will exist from the F5 unit to the AIP unit. It has been observed by (Rizzolatti, et al., 1988) that the number of F5 neurons which fire during the extension and flexion phases of the program is greater than the number that fire during either the set or hold phases of the program. Hence, on average, a single AIP unit will receive a larger number of recurrent projections from extension/flexion F5 units than any other type. As a result, its level of activity will usually be greater during these phases of the program.

## Comparison to Observed Biological Data

### F5 Phasic Responses

In the previous two figures (7.1, 7.2), the activation of each of the F5 cells was restricted to exactly one of the defined phases of movement. However, in Chapter 2 we observed that it is often the case that a single F5 cell in monkey is active over more than one of these phases. In the model, this property is achieved by connecting the basal ganglia phasing loop to the F5 unit in such a way that it is prepared to activate during one phase but then is not immediately shunted on the following phase. Rather, the circuitry allows for an arbitrary number of intervening phases to exist between these two events (see Figure 5.12 and accompanying discussion for more details). In addition, an F5 unit may exchange positive supporting connections with other F5 units that are active over disparate phases. The result is that from one phase to another, the F5 unit may receive different levels of positive support, causing a modulation of the overall activity of the column.

Both of these effects are demonstrated in Figure 6.3. F5 units A and B are set cells, although unit B continues to fire through the extension phase of movement. Units D and E are only active during the extension and flexion phases; the latter is qualitatively similar to the unit (recorded in monkey) that is shown in Figure 2.23. Unit G fires significantly during the enclose phase of the grasp, and then continues to fire at a moderate level through the hold phase of movement. This mimics the F5 cell behavior shown in Figure 2.24.

### AIP Phasic Responses

Figure 6.4 demonstrates a range of AIP cell behaviors. Like the monkey cell of Figure 2.12, unit A becomes active almost immediately (before the ready signal is given). This feature indicates that unit A receives a significant degree of input from the visual areas. Units B and C only activate once the ready signal has been received. This is the case because these units rely

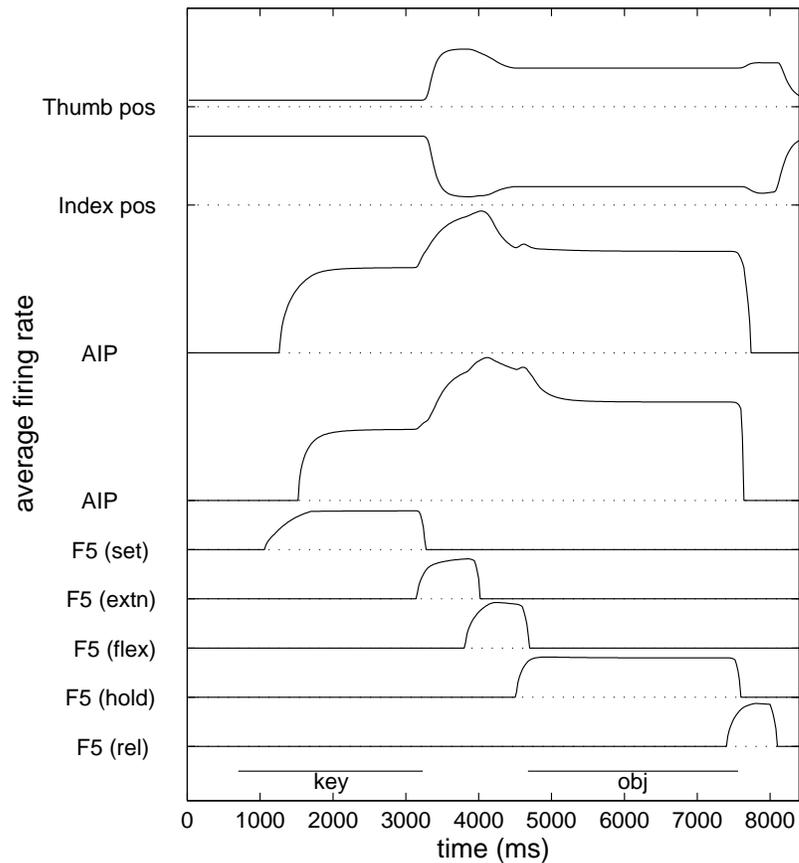


Figure 6.2: F5/AIP interaction during execution of the grasp.

on recurrent activity from the set cells in F5 in order to achieve a level above threshold. In regards to onset, these cells compare to the monkey cell of Figure 2.14. However, in this case, the activity ramps up continuously from onset until movement execution is complete. Units D and E do not receive any significant visual inputs, and thus rely completely on the recurrent F5 activity during the movement and the hold phases of the task (these compare in behavior to Figures 2.15 and 2.13, respectively).

### AIP Visual/Motor Distinction

One classification that Sakata uses to describe AIP cell behaviors is that of visual- versus motor-related activity (Taira, et al., 1990). Neurons are tested in four different conditions: {grasping an object versus fixation of an object} x {performance in the light versus in the dark}.

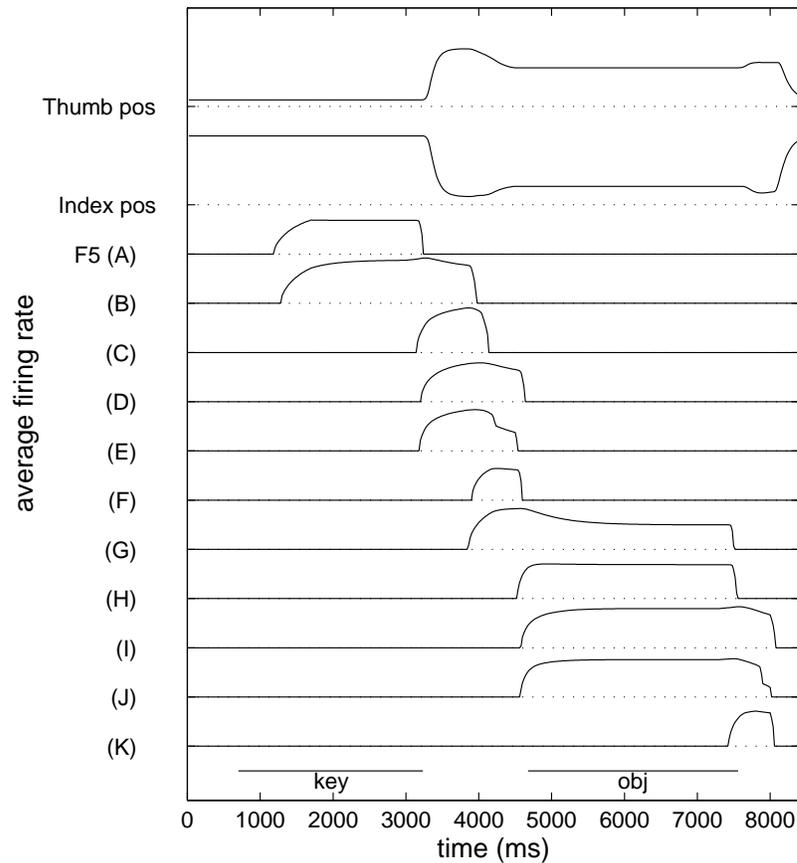


Figure 6.3: A set of F5 cells with a variety of temporal properties.

A neuron is considered to have a motor-related component when it is active during grasping, regardless of whether the task is performed in the light or dark. A neuron is visual-related if it is active for conditions in which the object is visible and is not active for the dark conditions.

In the model, we explain this distinction as a function of the relative strength of the inputs into the AIP columns from the visual areas as compared with F5 layer. Motor-related cells receive relatively large recurrent projections from F5; visual-related cells receive inputs primarily from PIP and IT.

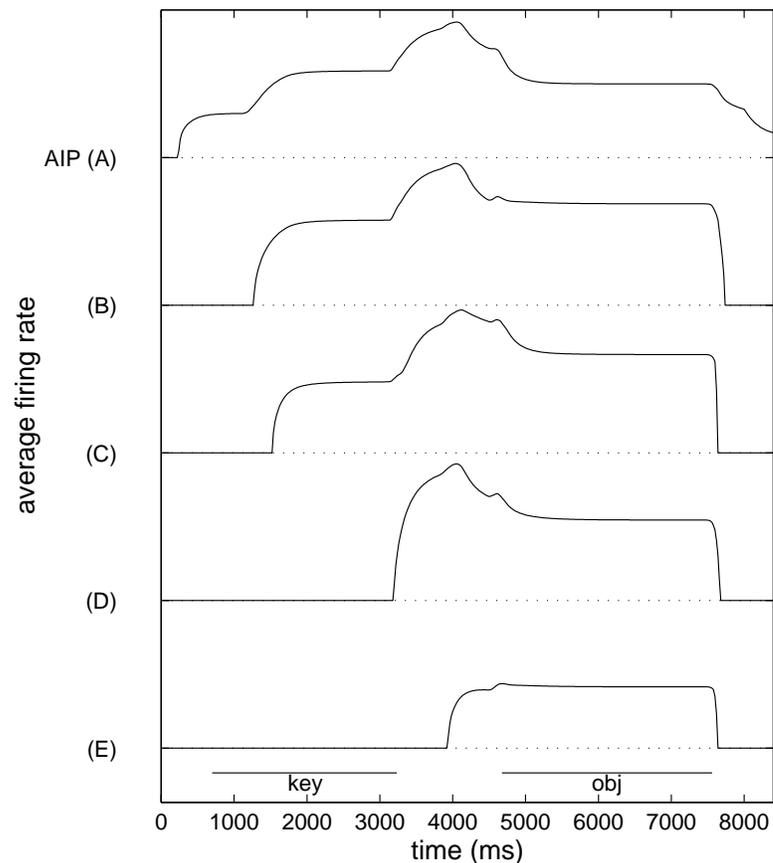


Figure 6.4: Variety of AIP cells.

Different experimental protocols are used to test the model in the three conditions of interest (here, we ignore the dark fixation task). For the movement in the light condition, the protocol is identical to that described above. In the fixation task, the system is presented with the object as in the original protocol, however neither the ready nor the go signals are ever provided. Hence, the entire 8.4 seconds of the task is spent fixating on the object and no movements are generated. For the dark grasping task, the model first performs the task in the light (traces are not shown). During the performance of the task in the light, area 46 maintains a memory of those F5 cells that participate in the grasp. Once the initial grasp is complete, the visual inputs are removed (PIP and IT are cleared), and a second trial is initiated with the

presentation of the ready and go signals. During this second trial, area 46 provides positive support to those F5 cells that were active during the first trial. With this added bias, the F5 cells corresponding to the grasp executed on the previous trial achieve a higher level of activity and are able to shunt F5 units belonging to other grasps.

Note that the area 46 working memory is providing essentially a static description of the grasp that was recently executed. By *static* we mean that the temporal aspects of the grasp are not stored - only a memory of those units that were active at *some time* during the execution. How might this relate to the observation that area 46 is involved in human in the imagination of grasp execution (Decety, et al., 1994)? One possibility is that this region provides the static representation of the grasp to be executed. This information could then provide the context required by the mesial motor areas (F3 and F6) and F5 in order to simulate the grasp execution -- unfolding the static representation in time (Grafton, et al., 1996).

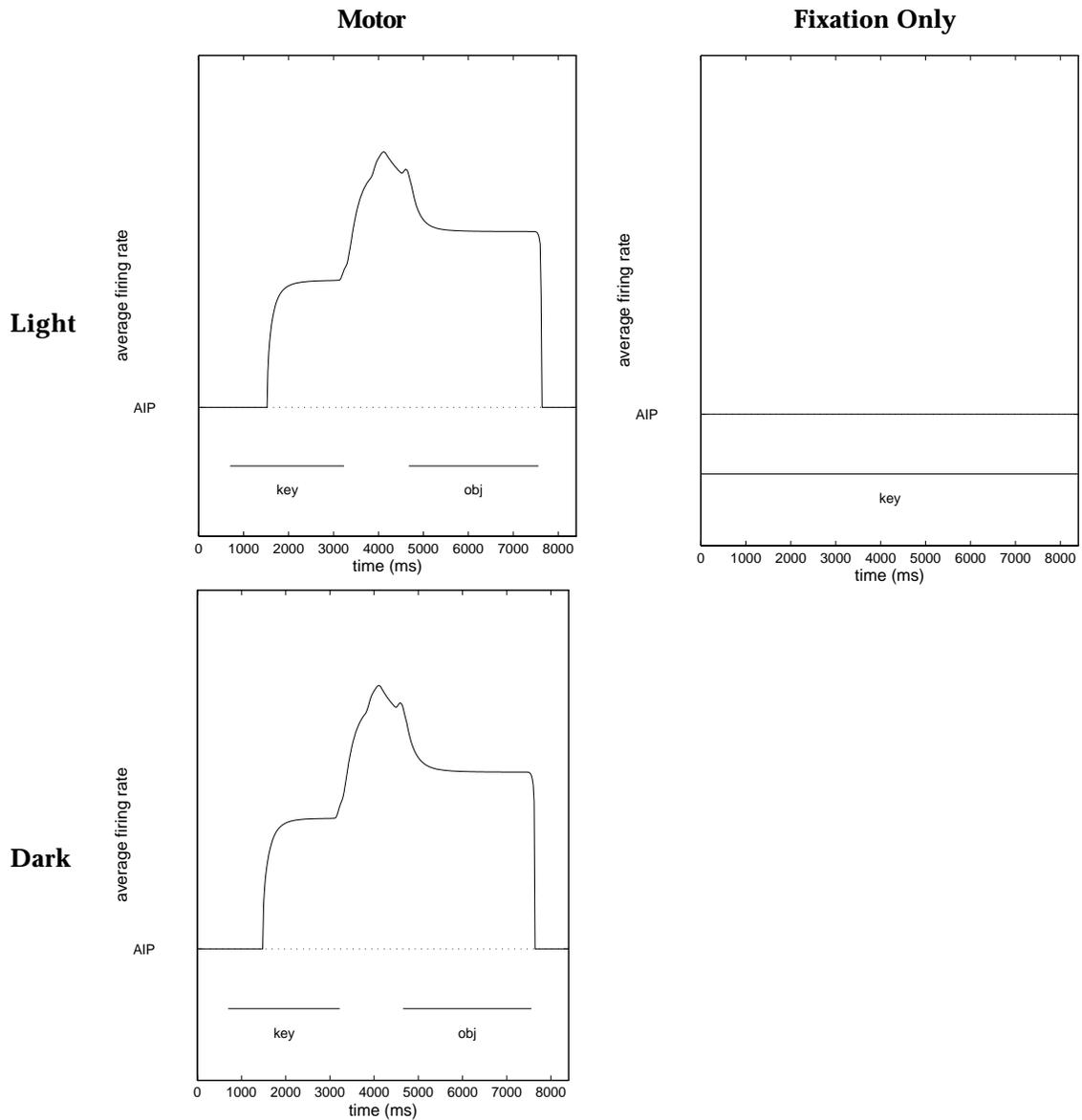


Figure 6.5: A pure motor-related AIP cell.

Figure 6.5 demonstrates one of the model's pure motor cells. During the fixation condition, the unit does not respond to the visible object. In addition, movement in either the light or the dark results in identical activity traces.

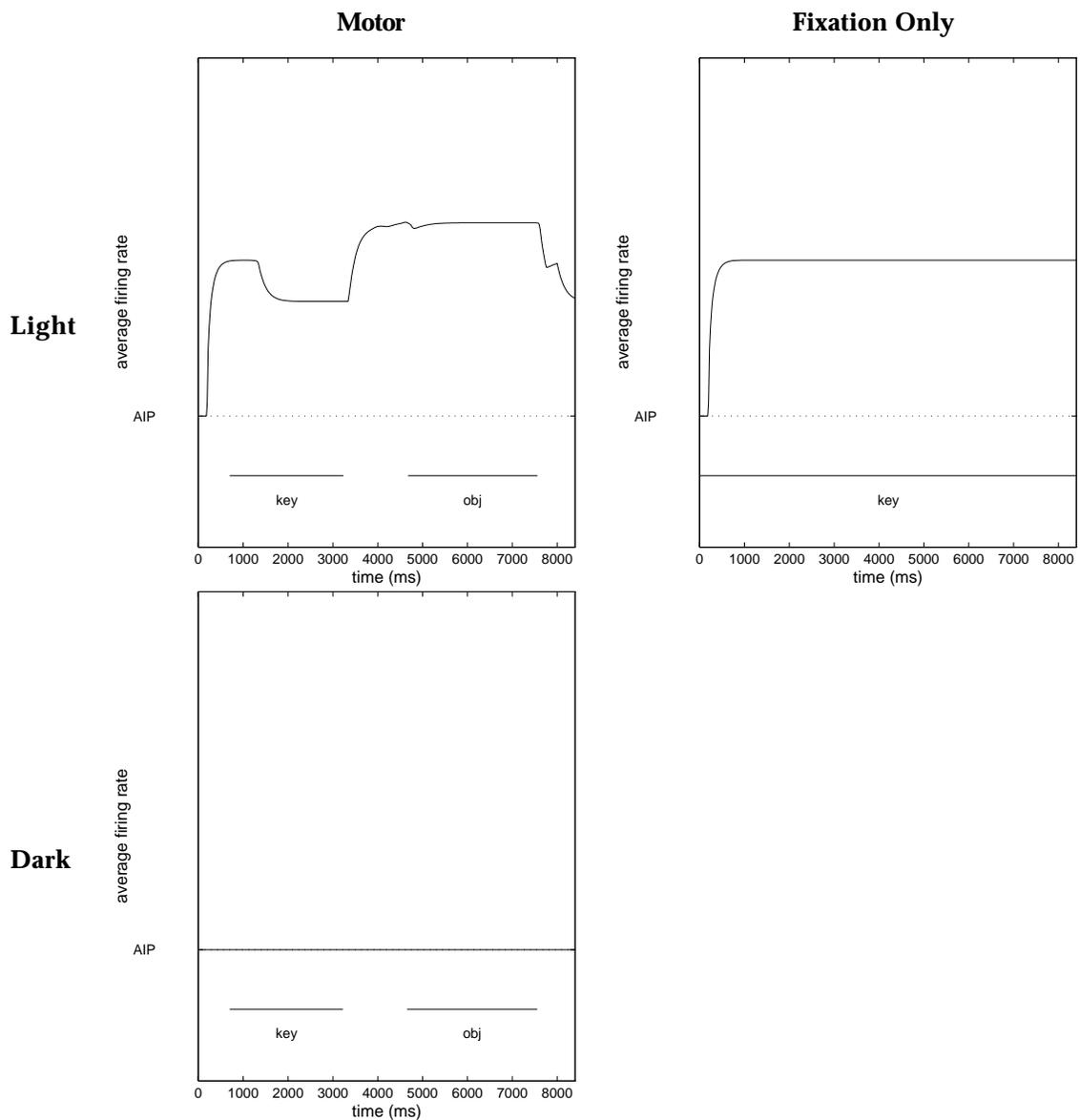


Figure 6.6: A pure visual AIP unit.

A pure visual-related cell is shown in Figure 6.6. Even though this unit does not receive direct projections from any F5 units, its activity level is modulated by the phase of movement. The initial increase in activity is due to the visual input from PIP. Shortly after the ready signal is received, a number of F5 set units begin to activate, which in turn activate motor-

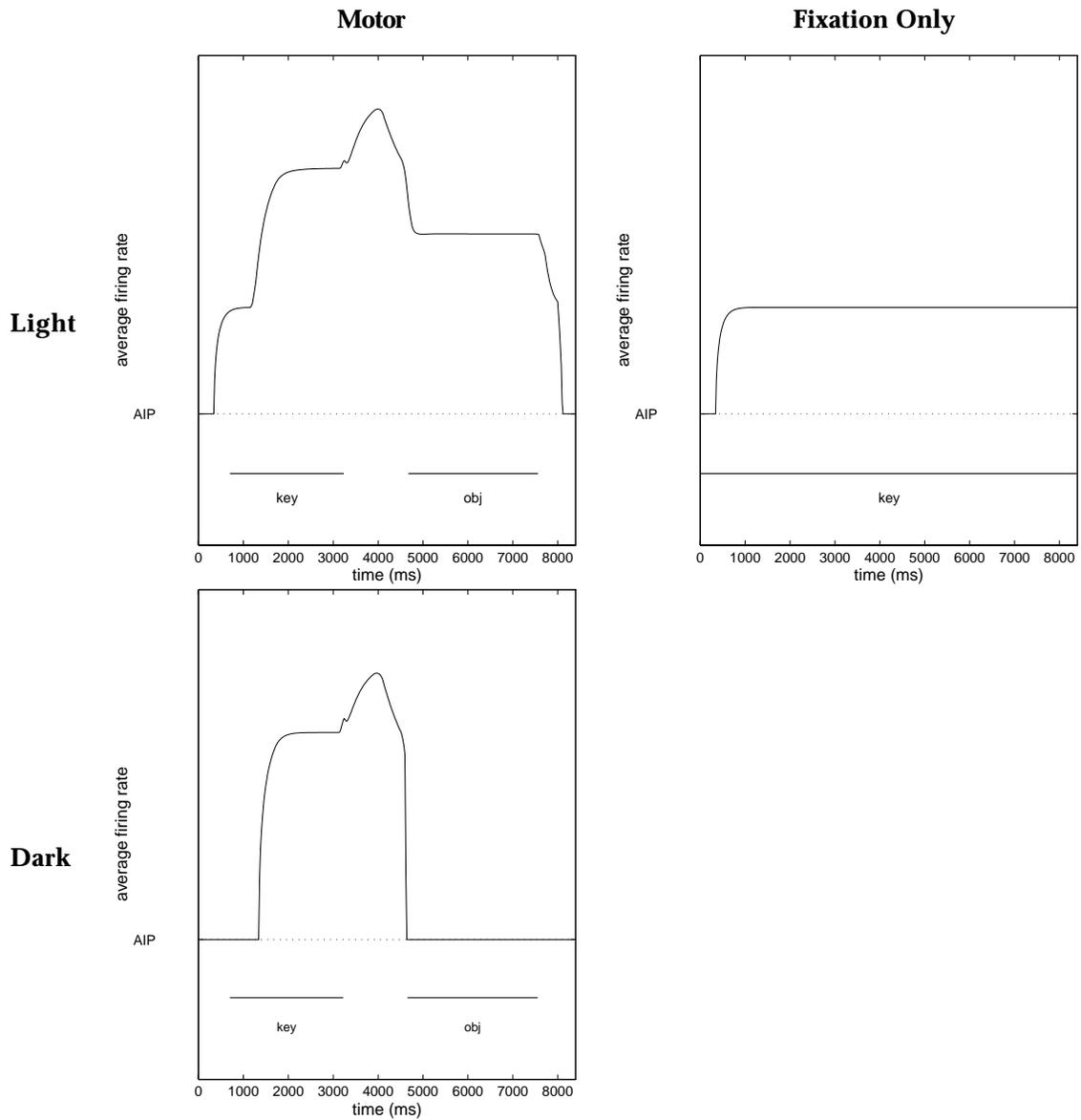


Figure 6.7: A visually-modulated AIP unit.

related units in AIP. This increase in overall activity in AIP results in an increase in the inhibitory signal through the basal ganglia loop, causing a reduction in the activity level of the unit shown in the figure.

The subsequent increase in activity after the key phase is also due to indirect mechanisms: surrounding motor-related units become highly active during movement, and thus provide a significant degree of positive support.

Finally, Figure 6.7 demonstrates the behavior of an AIP unit whose motor-related responses are modified by visual inputs. During the dark condition, the overall activity of the unit is significantly reduced. As a result, the unit is not able to remain above threshold once the network reaches the hold phase of the motor program.

### **Population Analysis**

Examining the behavior of individual cells allows for an understanding of how a cell contributes and is affected by the execution of a motor program. However, it is also desirable to look at how a set of cells work together in order to implement the program execution. The population analysis described here is designed to give a broad view of cell behavior across a large number of cells. The basic scheme is such: we measure the cells "participation" during the execution of a single trial (and possibly over a single phase of the trial). This measure is defined as the cell's average firing rate during the condition and phase of interest (note that much temporal information is lost in this averaging process). Given two (or more) such measures for different tasks or phases, we can compare the relative participation of the cell under these different conditions, and then examine how these relative measures vary over the population.

Pairs of measures can be plotted in a plane (see Figure 6.8) in order to construct a visual representation of these relationships. Panel A of Figure 6.8 demonstrates one such comparison between the normalized cell response during the set and extension phases of movement (during execution of a precision grasp). Each point in the plot corresponds to a single F5 module (the activity of the output unit of the column). By *normalized cell response*, we mean that we first compute the average firing rate of each cell during the tasks and phases of interest:

$$M_i^X = \frac{\int_{t_0^X}^{t_1^X} output_i^X(t) dt}{t_1^X - t_0^X} \quad (6.1)$$

where:

$M_i^X$  is the average firing rate of unit i, under condition X.

$t_0^X$  and  $t_1^X$  define the phase of interest for condition X.

$output_i^X(t)$  is the firing rate of the output unit of column i, at time t, under condition X.

The *normalized response* over a set of conditions {Y} and a population of cells is computed according to:

$$\hat{M}_i^X = \frac{M_i^X}{\text{Max}_{j \in \text{cells}, z \in \{Y\}} \{M_j^z\}} \quad (6.2)$$

### Population Analysis of Phasic Behavior

The relative phasic responses of an entire F5 population during execution of a precision pinch are shown in Figure 6.8. The set and extension phases of the program (A) show relatively little overlap between the two phases, as indicated by the large number of units aligned along the axes (i.e., nearly a 0 component for one of the two measures). Comparing extension and flexion phases of movement, however, we see a significantly different result. Although there are still some neurons that are exclusively involved in one phase of the movement, a large number of units are shared between the two phases, and the degree of sharing varies significantly from one unit to another. Finally, comparing flexion versus the hold phase of movement, we again have a significant separation between the two populations. Also note that the flexion phase of movement utilizes a significantly larger number of cells than does the hold phase.

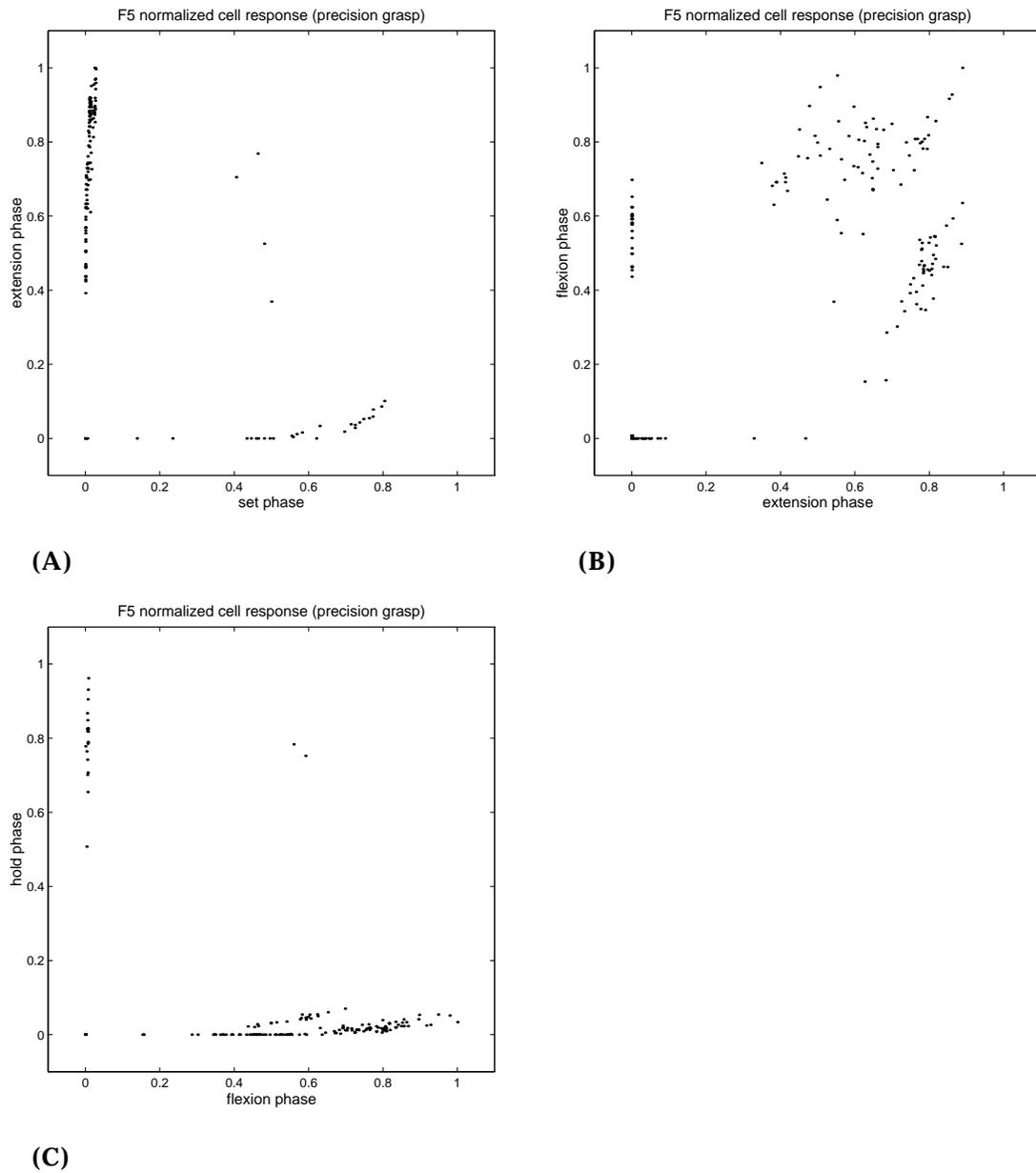


Figure 6.8: Population analysis of F5 phasic behavior during a precision grasp. (A) set versus extension phase; (B) extension versus flexion; and (C) flexion versus hold.

The AIP population, however, demonstrates a significant amount of sharing between all phases (Figure 6.9). This is the case because once many AIP neurons become active, they remain active over the course of the motor program (through the end of the hold phase). The set and

extension phases share a reasonable number of cells, although an equal number only begin to fire as the extension phase begins to execute (A). Note also that the overall response of the cells is greater during the extension phase as compared with the set phase. This is due to the larger amount of recurrent feedback being received from F5 during the movement phase of the program.

Comparing extension phase with flexion, we see a large number of cells that are shared nearly equally (B). However, a few cells demonstrate a very late onset (left-hand column of points). During the hold phase of movement, a number of neurons decrease their activity significantly (the line of points in the lower, right-hand corner of panel C). Otherwise, the two phases of movement share the remainder of the cells equally.

### **Population Analysis of Visual/Motor Properties**

The population analysis technique not only can be used to compare average activity during different phases of the task, but can also be used to compare cell activity under different conditions. This idea is illustrated in the comparison between the lighted fixation and dark movement task (still a precision grasp is being executed in the latter case; Figure 6.10). The most striking effect in the figure is the fact that a large number of cells are only active during dark movement, and are not at all active during fixation. However, a smaller population of cells exhibits both visual and motor responses (and only a small number are exclusively visual-related).

Panel (B) demonstrates the relative activity of the AIP population during the execution of the dark movement task as compared with the light movement task. A significant number of these cells do not show any modulation between the two tasks. However, a few cells are further facilitated during movement in the light. Finally, as a comparison to coding in F5, panel (C) demonstrates that the two grasps that are executed (precision grasp, in the light and dark) are coded using the same set of cells.

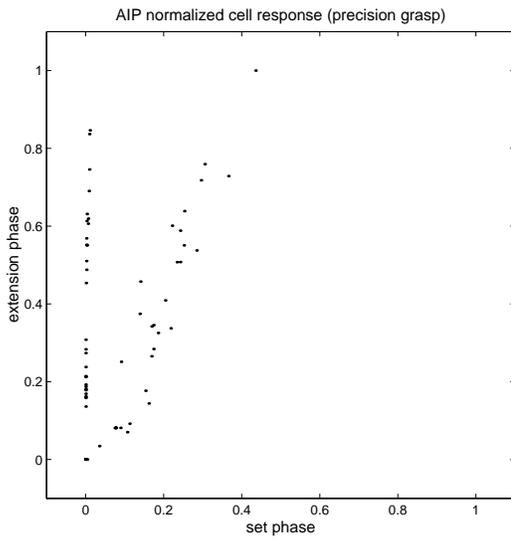
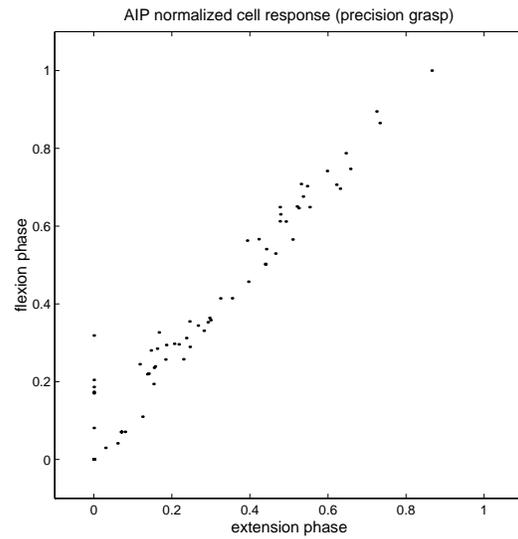
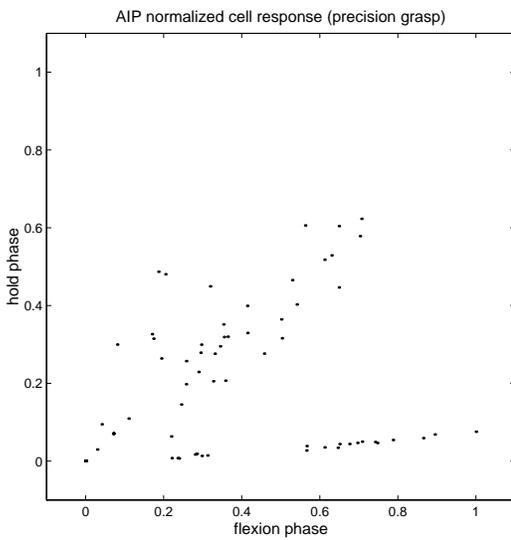
**(A)****(B)****(C)**

Figure 6.9: Normalized AIP responses over the entire population; (A) set versus extension phase; (B) extension versus flexion; and (C) flexion versus hold phase.

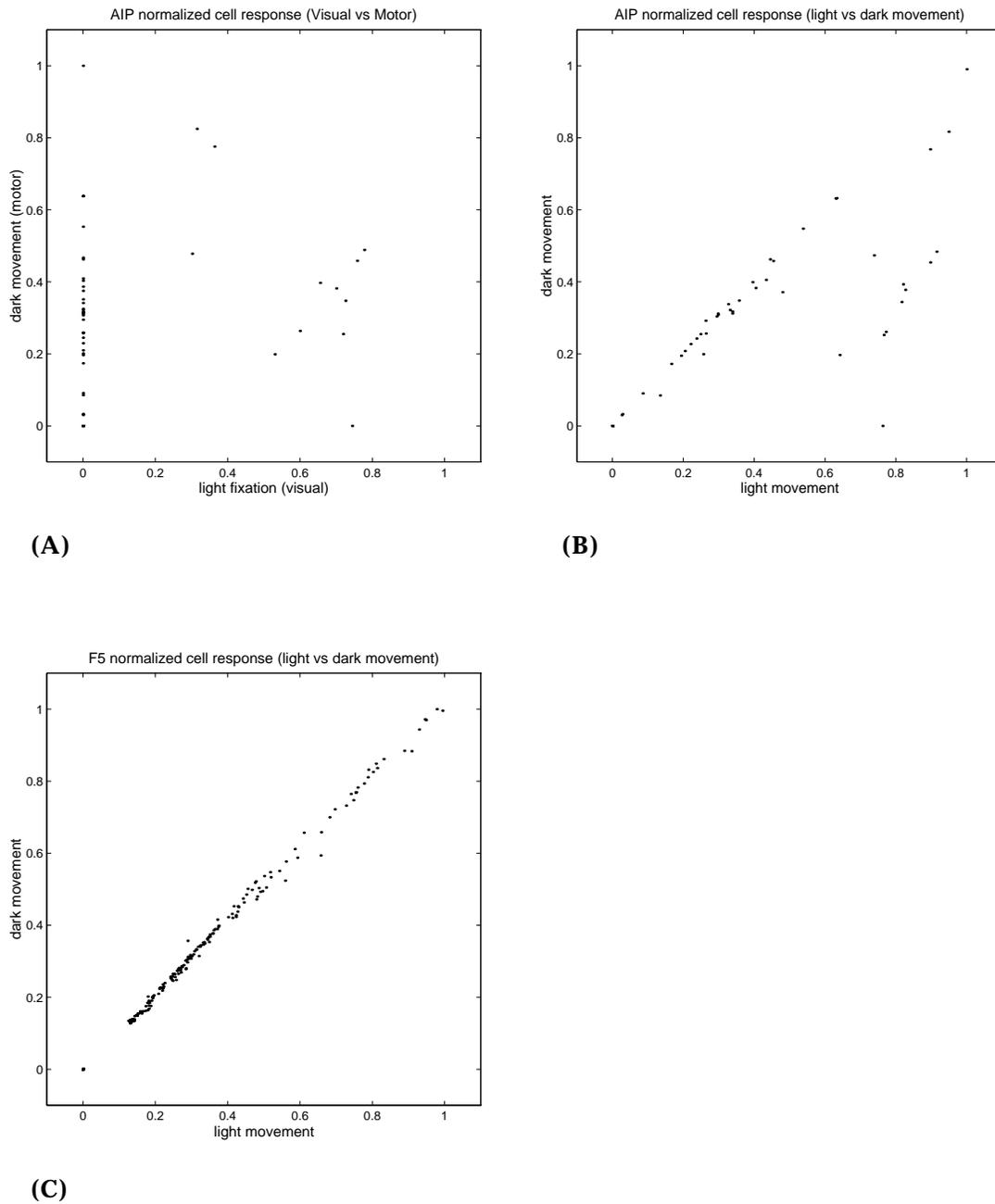


Figure 6.10: Visual/Motor coding in AIP: comparisons between (A) fixation/dark movement, and (B) dark and light movement conditions; and (C) comparison between light and dark movement coding in F5.

## Model Predictions

In this section, we present a series of experiments that are designed to ask further questions about how object and grasp information is utilized and represented in AIP and F5. The behavior of the model in these novel tasks provides a set of specific predictions (at both the single-unit and population levels) as to what we might expect if the same experiments were performed in monkey.

The first task examines one aspect of how objects are encoded in AIP (prediction #1). We then turn to a task in which a single object may be manipulated in more than one way and look at how AIP and F5 respond when a delayed instruction stimulus is used to inform the model as to which grasp to use (prediction #2). Third, we examine at how fine parameters of the grasp (in this case the grasp aperture) might be coded in both F5 and AIP (prediction #3), and how these two regions respond when there is a mismatch between the expected object size and the actual object size (prediction #4). Finally, we ask how object parameters are coded in AIP relative to how they are used in the programming of the grasp (prediction #5).

### Prediction #1: Object Coding in AIP

In this experiment, the model is presented with two different objects that are roughly the same size (Figure 6.11). The objects are chosen such that their representations in PIP are non overlapping (in this case we are using a small cylinder and a narrow plate), but both objects are graspable using the identical precision pinch. We are interested in looking at how AIP encodes the affordances for these two objects and to what degree the object-specific properties are preserved in this region.

Figure 6.12 shows the responses of two visual-related cells under the two experimental conditions. Unit A is active only for the condition in which the system is grasping the plate. Unit B, on the other hand, is somewhat active during the set and movement phases of the cylinder condition, but is most heavily activated when grasping the plate.

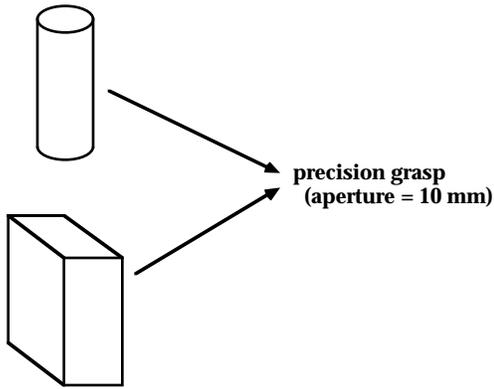


Figure 6.11: Two objects that map to the identical grasp.

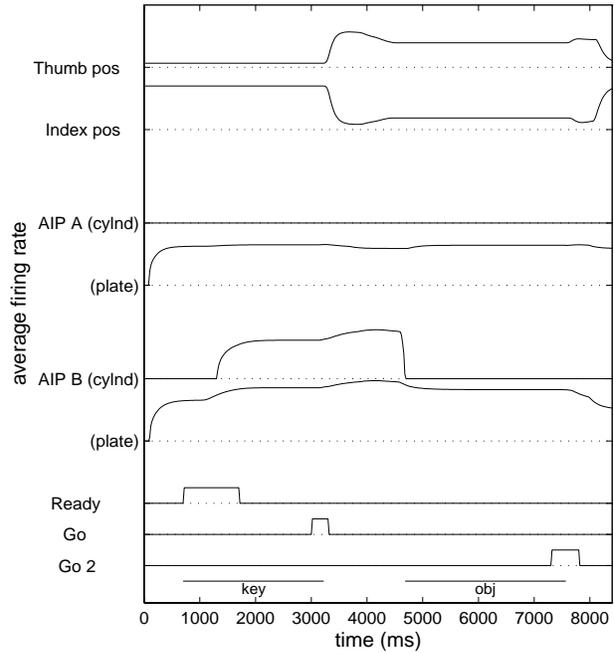


Figure 6.12: Two visual AIP cells that show significant object-specific modulation.

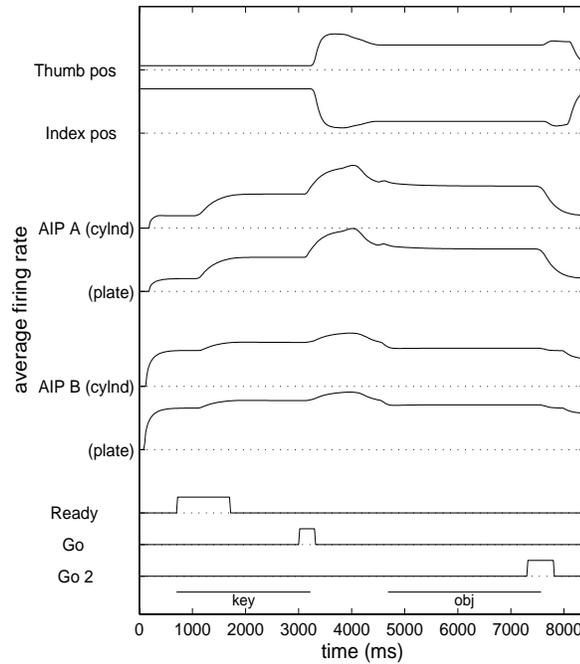


Figure 6.13: Two visual-related (with some motor) AIP cells that demonstrate only slight modulation as a function of the type of object that is presented.

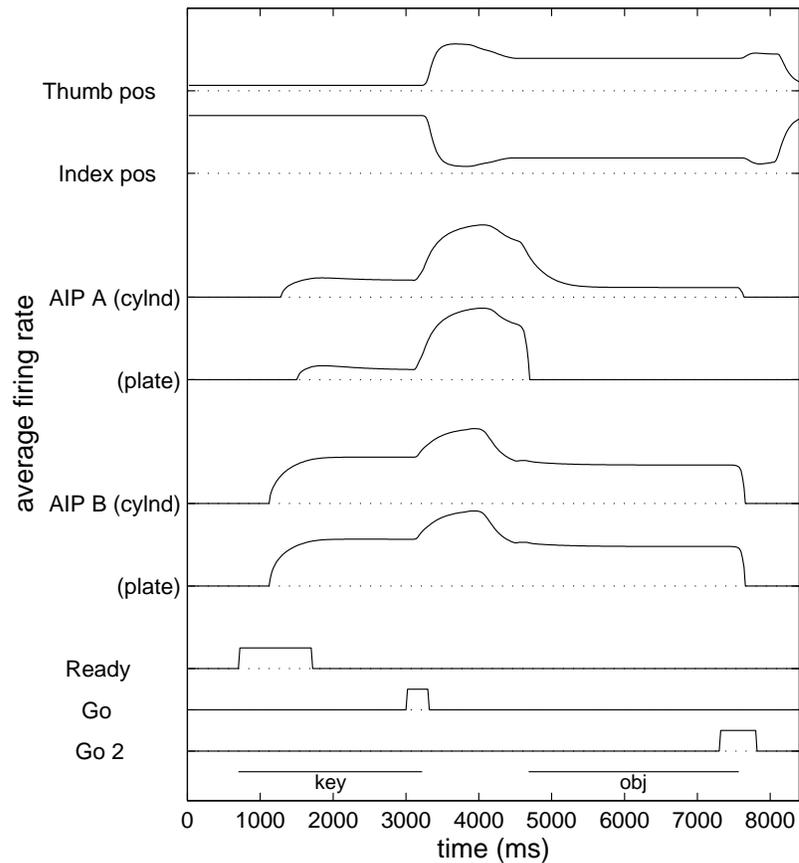


Figure 6.14: Two motor-related (set phase) AIP cells. Only cell A is modulated the type of object that is presented.

Two cells with both visual- and (somewhat) motor-related components are shown in Figure 6.13. In the case of cell A, the primary difference in response level is observable only when the object is first presented to the model (it responds more quickly to the cylinder). Cell B, on the other hand, demonstrates a small overall advantage for the plate condition.

Motor-related cells are less sensitive to the type of object that is presented. Figure 6.14 shows two units that begin to fire after the ready signal is presented. Only cell A demonstrates object-specific modulation, shutting off after the flexion phase of movement is completed. Cell B, along with cells A and B of Figure 6.15 does not change in activity level as a function of the presented object.

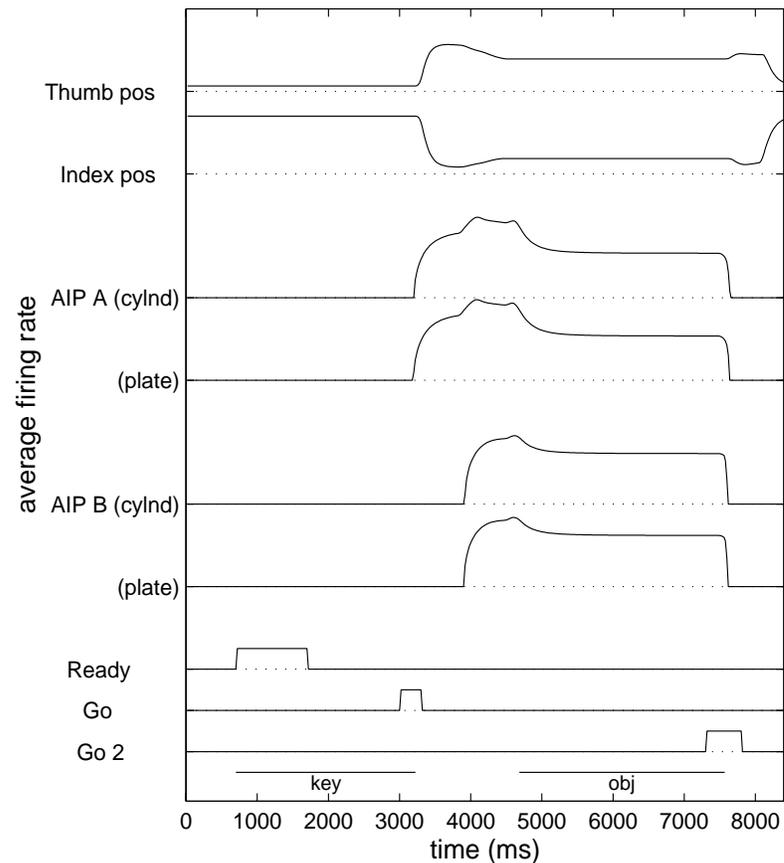


Figure 6.15: Two motor-related AIP cells that are not modulated by the type of object.

### Population Analysis of Object-Related Responses

The visual/motor distinction in AIP becomes more clear at a population level of analysis (Figure 6.16). Under the lighted movement condition, the AIP population demonstrates a range of object responses (A). Although a few cells demonstrate significant differences in response to the two objects, many are equally active for both. As a comparison, almost every F5 cell is equally responsive to both objects (B).

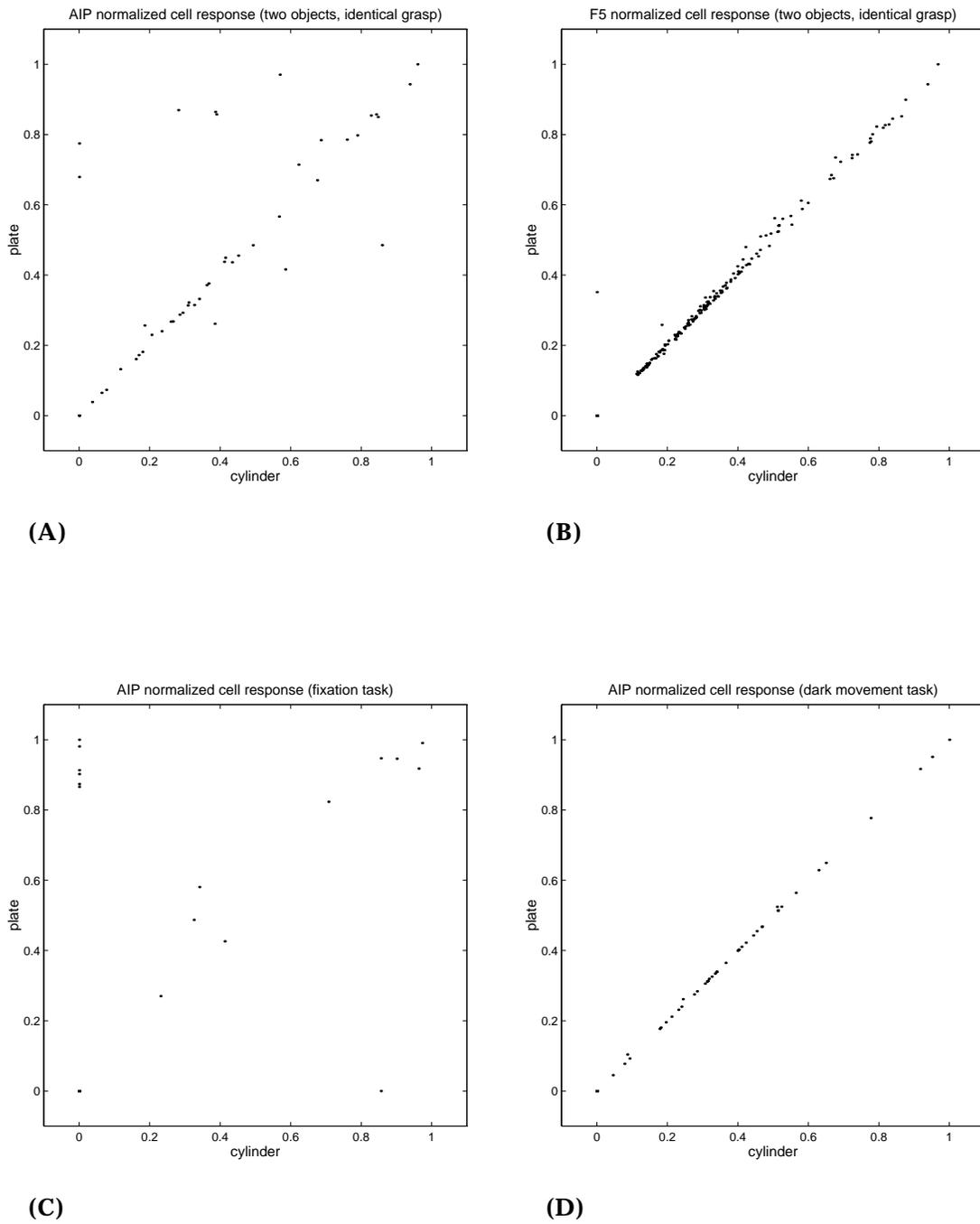
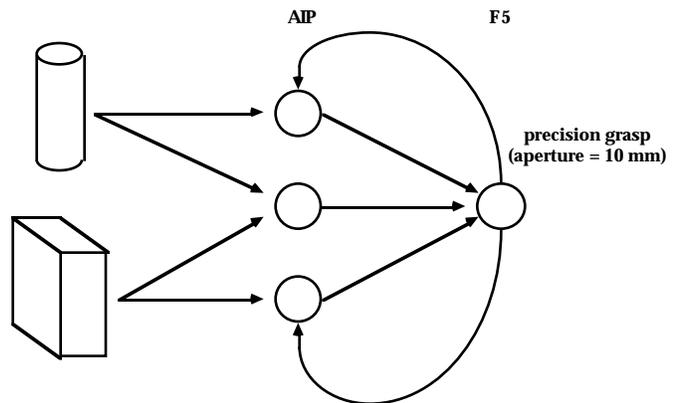


Figure 6.16: Comparison of population responses towards two different objects (but identical grasps). Lighted movement task, AIP (A) and F5 (B) cells; and AIP populations during fixation (C) and dark movement (D) tasks.

In order to examine the visual/motor responses of the AIP cells, we compare the cell responses under the fixation (C) and dark movement (D) tasks. The fixation responses are a direct result of connections from visual areas into



AIP, and show significant differences in behavior between the two objects. However, under the dark movement task, the active AIP cells do not show object specificity.

Figure 6.17: Visual-related AIP receive object-specific inputs; motor-related cells receive recurrent inputs from F5, which do not demonstrate object-specific activity.

In general, units that receive a significant amount of visual input (and thus are classified as visual-related cells) are more likely to display differences in how they represent the two different objects, as compared to the motor-related units. As discussed in Appendix A, when a mapping is established from an object representation in PIP (or IT) to a set of appropriate AIP units, the decision as to whether or not to make a connection to an individual AIP unit is determined probabilistically. In this experiment, because the two objects are represented in PIP using non-overlapping populations of neurons, only a small subset of the possible AIP units will respond to both objects (Figure 6.17).

For the most part, however, the object distinction is lost at the level of F5, which is the source of the motor-related activity. This happens despite the fact that only a subset of possible projections between AIP and F5 are established (also determined probabilistically). The reason is that the motor-related activity flows through the loop formed by AIP and F5. As the activity makes successive cycles through the loop, more and more motor-related cells (from both AIP and F5) are recruited, until all or most of the cells that belong to a single grasp

program become active. As a result, the model predicts very few differences in object coding in AIP cells that are classified as more motor-related.

### **Prediction #2: When an Object Affords Multiple Grasps**

In general, a single object affords many possible grasps, one of which is selected at the time of execution as a function of the current context (which may include task requirements, position of the object in space, and obstacles). How is this one-to-many mapping captured in the F5/AIP circuitry? And might we be able to see evidence of this mapping in a monkey recording study?

In the following experiment, the model is presented with a single object (a small cylinder), and asked to perform one of three tasks. For each task, the model repeatedly grasps the cylinder a fixed number of times. The three different tasks are:

1. Grasp the cylinder using a precision pinch.
2. Grasp the cylinder using a side opposition.
3. As a function of an instruction stimulus (e.g., the color of a light), grasp the cylinder using either a precision pinch or a side opposition.

When the grasp is known ahead of time (tasks 1 and 2), it is assumed that some higher level planning region predisposes the selection of the correct grasp. In the model, it is area F6 that performs this function. However, when the correct grasp is not known a priori (task 3), some other region must be invoked that is capable of associating arbitrary stimuli with the appropriate grasp. As discussed in the previous chapter, we implicate the dorsal premotor cortex (F2) in this association process (Fagg & Arbib, 1992; Mitz, et al., 1991).

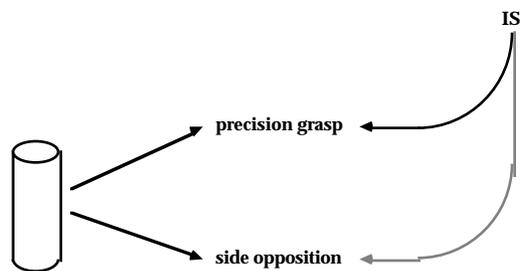


Figure 6.18: A single object mapping to two possible grasps. Before execution, one grasp must be selected based upon the current context (e.g., based upon an Instruction Stimulus).

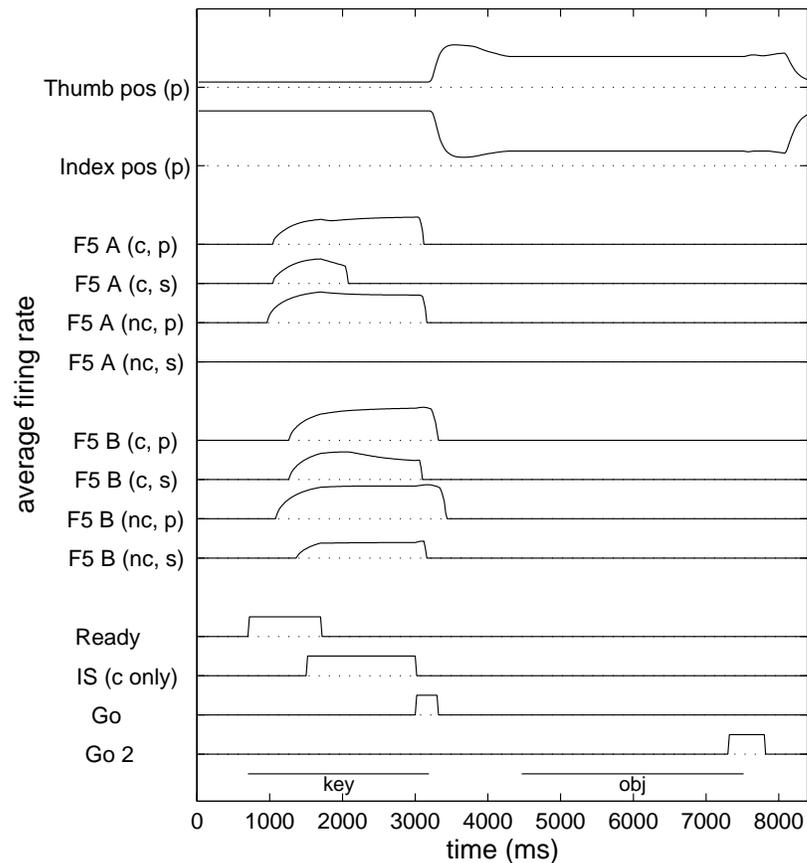


Figure 6.19: Two F5 units in response to the four conditions: (c,p), (c,s), (nc,p), and (nc,s). c = conditional; nc = non-conditional; p = precision grasp; s = side opposition.

The protocol used in this experiment is similar to the one used in the previous set of experiments. The only difference is that for task 3, an instruction stimulus is presented 800ms after the ready signal.

In the following set of figures, we examine in detail the various classes of cell behavior observed in F5 and AIP when presented with these four different conditions. We refer to task 1 as non-conditional/precision, and task 2 as non-conditional/side. Task 3 consists of two separate conditions: conditional/precision and conditional/side.

Figure 6.19 shows the behavior of two set-related cells in F5. For both the precision and side conditional cases, the cell activity is identical up until the IS is presented (as would be

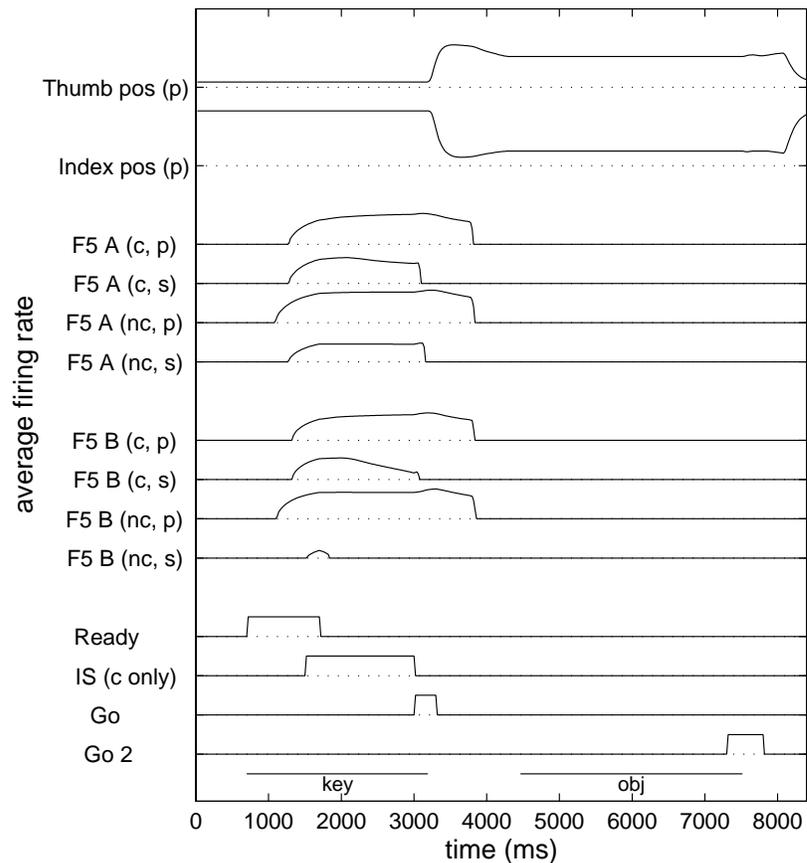


Figure 6.20: A pair of set/extension cells in F5.

expected for all cells). After the IS, we see a divergence in the activity pattern. Because these cells are both involved in encoding the precision grasp, they continue to increase in their activity level when the precision condition is specified, while their activity is shunted for the side condition (note that cell A is shunted almost immediately, while cell B remains active until the go stimulus is received).

Examining the non-conditional traces, we see that cell A is only active for the precision grasp case. During the side opposition case, this cell is significantly inhibited by other F5 units and thus cannot become active. Cell B, on the other hand, is active in both cases, even though the correct grasp is known before the trial begins (at  $t = 0$ ). Note, however, that for the precision grasp case, B is much more active than the side opposition case.

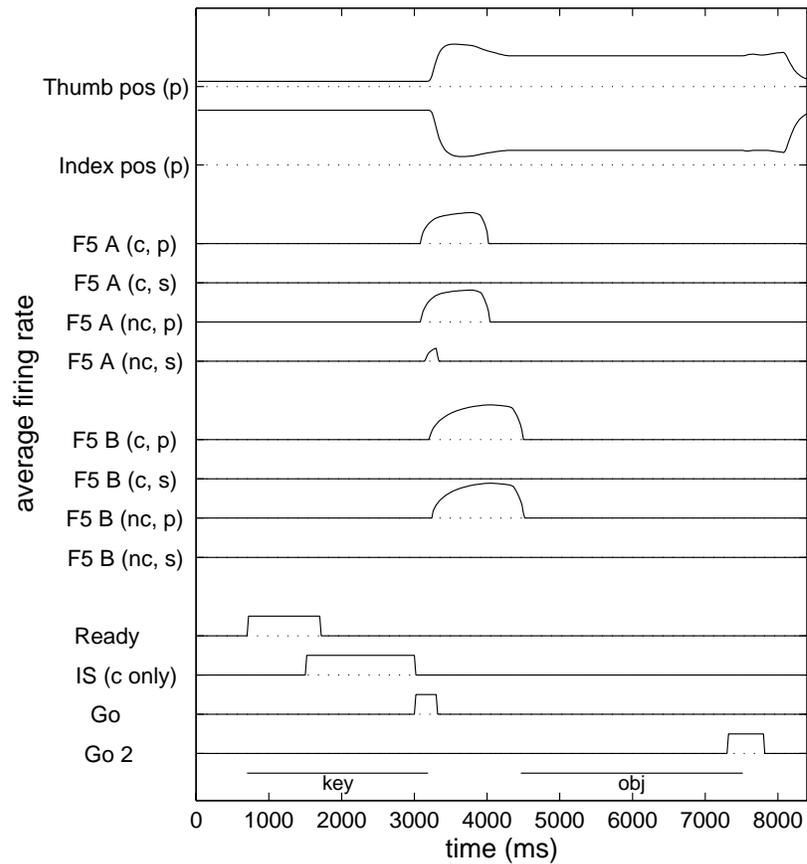


Figure 6.21: A pair of extension-related F5 cells.

Two combination set and extension-related F5 units are illustrated in Figure 6.20. The key behavioral feature of these units is the fact that both cells continue to fire into the extension phase of movement only when a precision grasp is being executed. When the side opposition is executed, these cells are shunted by the time the movement is initiated.

This shunting of competitor grasps just prior to or during the extension phase of movement is a feature that naturally comes out of the structure of the model. Because there are many more extension units in the F5 region than there are set-related units (Rizzolatti, et al., 1988), when phase transition takes place in response to the go signal, we see a rapid increase in the number of active cells in the region. In turn, there is a rapid increase in the level of inhibition through

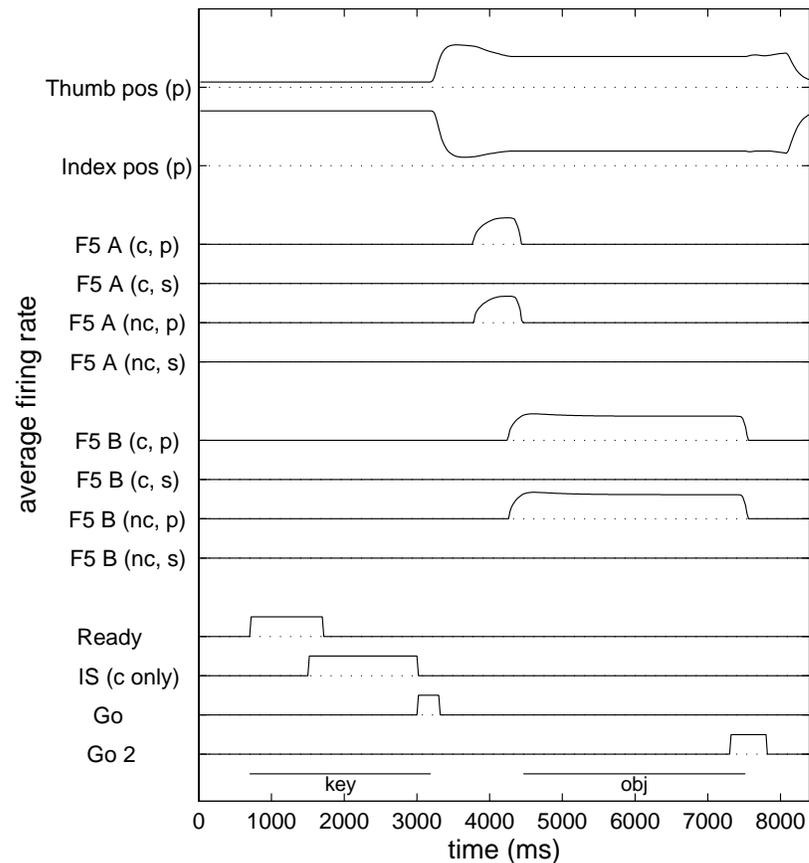


Figure 6.22: F5 flexion and hold units.

the basal ganglia inhibitory loop. The result is a sudden sharpening of the motor program that is about to be executed.

We see this effect in Figure 6.21, which shows a pair of extension-related F5 units. In only one competing case (execution of the side opposition), does cell A become active. However, it only does so for a short period of time. Once the program is sharpened, units participating in the flexion (Figure 6.22, cell A) and hold (cell B) phases of movement are generally compatible with the program being executed.

Because the number of hold-related cells is rather small in the F5 region, the motor program sharpening disappears to some degree during the release phase of movement (Figure 6.23, cell B).

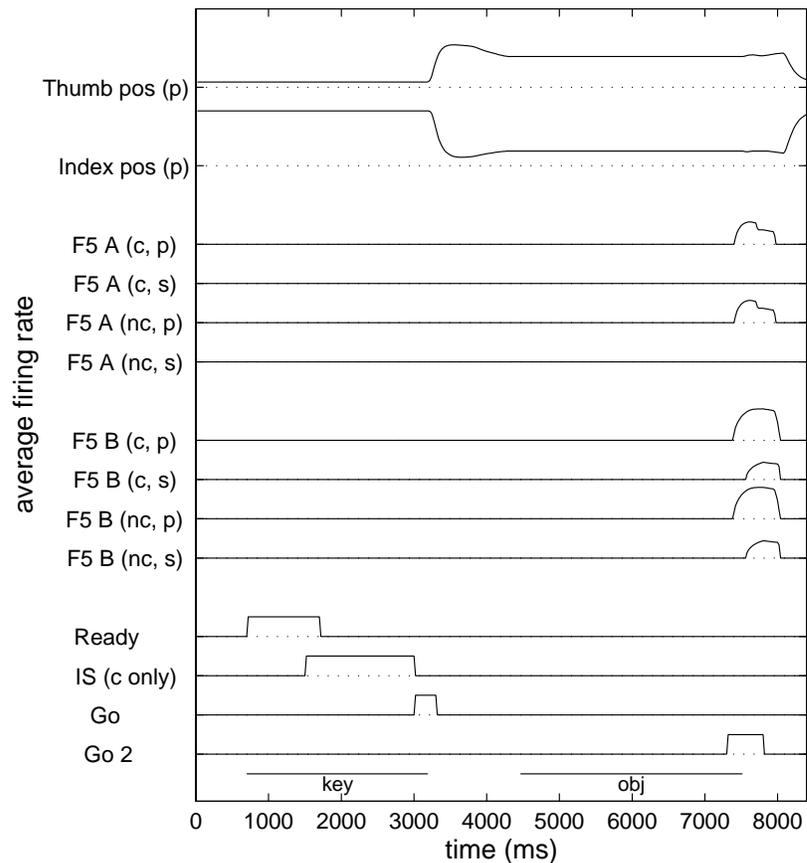


Figure 6.23: A pair of release phase F5 units.

Finally, Figure 6.24 illustrates a pair of set-related F5 cells whose temporal properties are significantly modulated by whether the system is performing a conditional side opposition or a non-conditional side opposition. In the conditional case, both cells delay their activation by at least 400ms. This happens because the cells are below threshold unless they receive both visual input via AIP and positive supporting inputs from either F2 or F6. Neither cell is significantly active for either precision grasp case.

The AIP cells in this conditional/non-conditional task reflect similar behavior as compared with the F5 units. The visual/motor cell shown in Figure 6.25 activates in response to the presentation of the object. In the cases when a side opposition is executed, this cell is

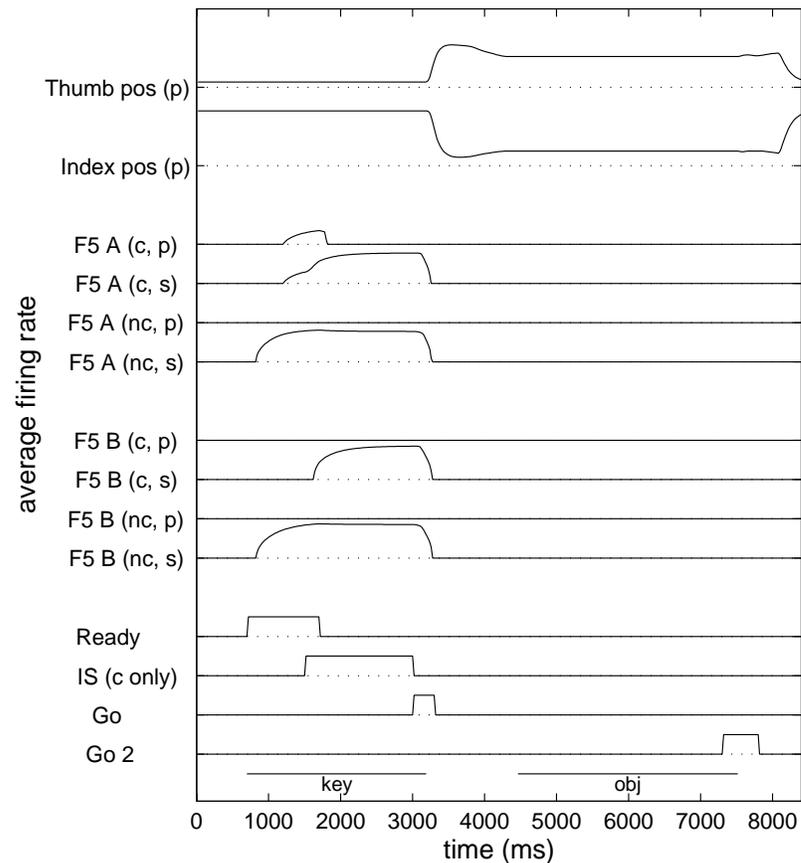


Figure 6.24: A pair of set-related F5 units. Note that in both cases, their temporal behavior is modulated by whether a conditional or a non-conditional task is being performed.

shunted during the set phase. The shunting is a result of competition with other AIP units via the inhibitory basal ganglia loop.

The visual/motor AIP unit illustrated in Figure 6.26 receives a more significant visual projection than the previous cell (Figure 6.25). This is evident in the fact that even in the side opposition cases, the cell is not completely inhibited (although it is significantly depressed).

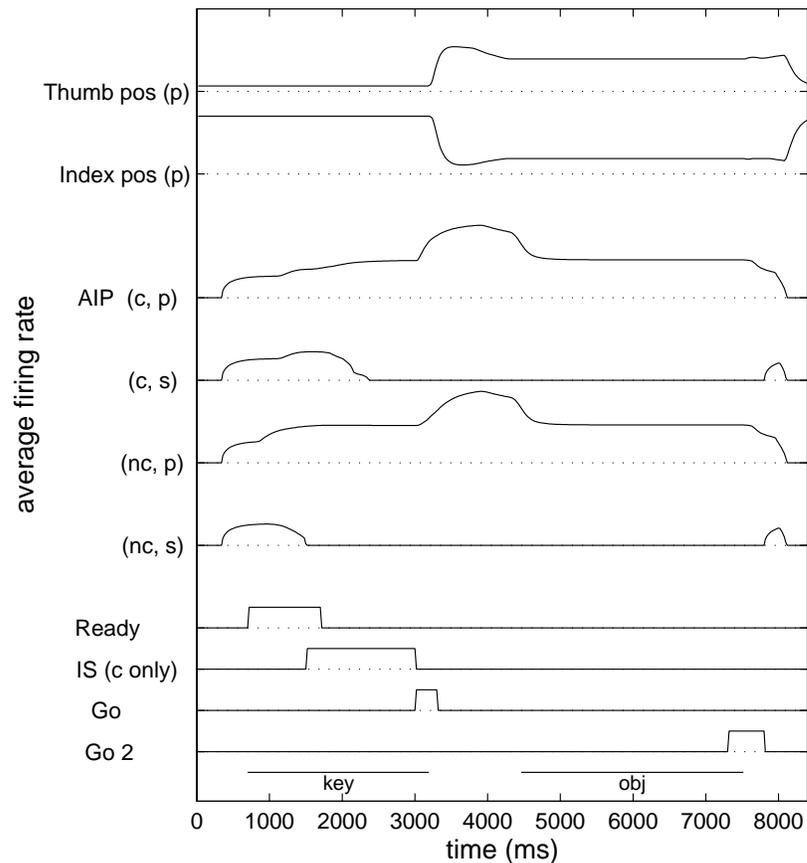


Figure 6.25: A visual/motor AIP cell (more motor-related). The notation is the same as in the previous figures.

Finally, the cells shown in Figures 6.27 and 6.28 are both purely motor-related. The former is a set-related unit, and demonstrates some level activity for all four conditions. In the non-compatible conditions, this cell is shunted by the time movement begins. The latter cell on the other hand (Figure 6.28), is never active in the non-compatible conditions. This is the case because it does not receive recurrent input from F5 set cells. By the time the movement is initiated, any possible F5 inputs into this cell will have been shunted due to the sharpening process that occurs at the set/extension phase boundary.

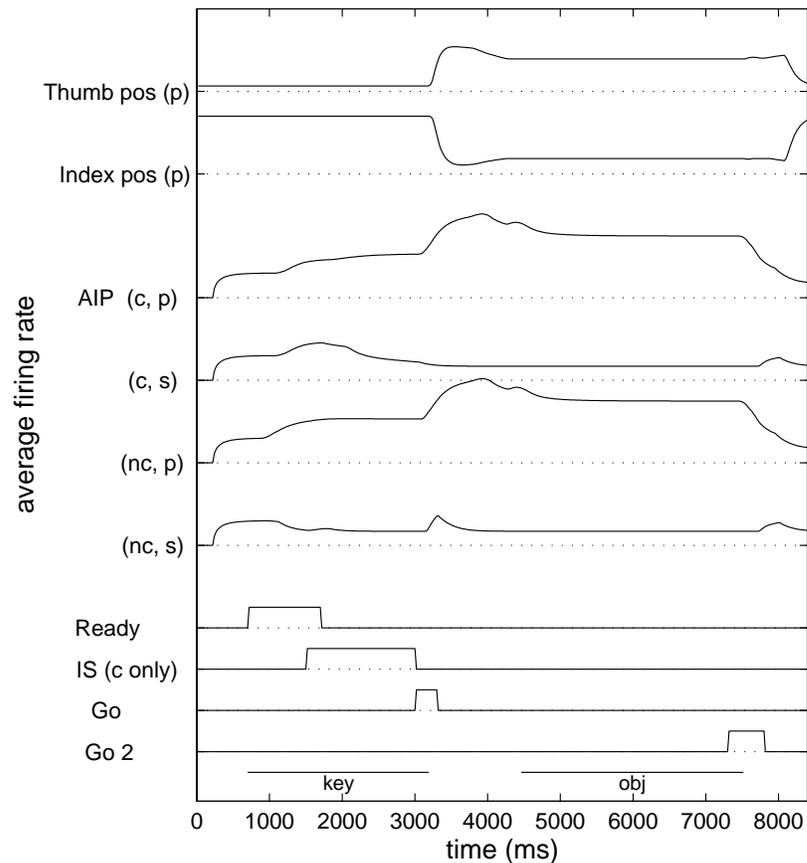


Figure 6.26: A visual/motor AIP cell (primarily visual), which participates in the execution of the precision grasp. During non-compatible (side opposition) conditions this cell is shunted.

### Population Analysis Comparing Conditional and Non-conditional Tasks

The population responses of the cells in F5 and AIP are summarized in Figures 6.29 and 6.30, respectively. In order to explicitly illustrate the sharpening process, normalized cell responses are computed for both the set and movement phases of the task. Many F5 set cells, under the conditional task, are active to varying degrees in both the precision grasp and side opposition cases (6.29.A). However when the grasp is known a priori (non-conditional task), relatively few set cells are active for both cases (panel B). In this task, when the ready signal is received, set cells from both grasps begin to respond. Because of the grasp-specific bias from F6, however, those cells that correspond to the specified grasp are the ones that first achieve threshold.

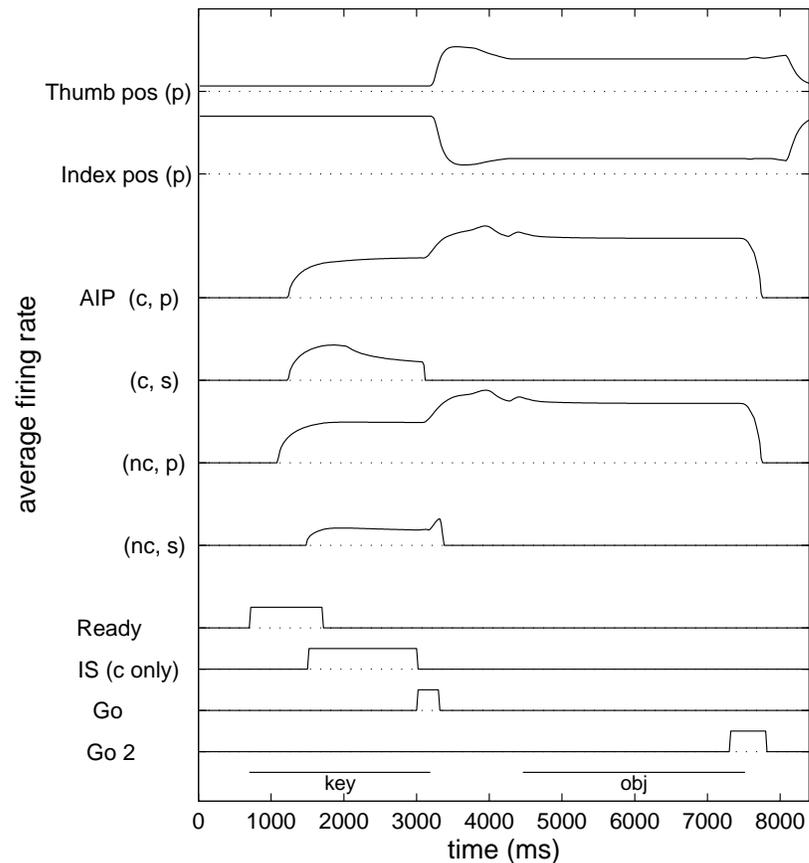


Figure 6.27: A primarily motor-related AIP cell, which is completely inhibited during execution of the side opposition.

Via the inhibitory BG loop, the other set cells (of the non-specified grasp) are then inhibited (in most cases, before they reach threshold).

Comparing the set responses to the movement-related responses (panels C and D), we see that cells participate essentially in one grasp or the other. This demonstrates that as a population, by the time movement is initiated, the grasp decision is completely made.

The AIP cell population reflects a similar behavior as the F5 population when comparing the conditional and non-conditional tasks (Figure 6.30). The primary difference is that a larger number of cells demonstrate reasonable levels of activity for both side and precision grasps. This is the case because many of the cells receive non-discriminatory visual input (even after the grasp is selected), and because the competition between affordances is not as significant as

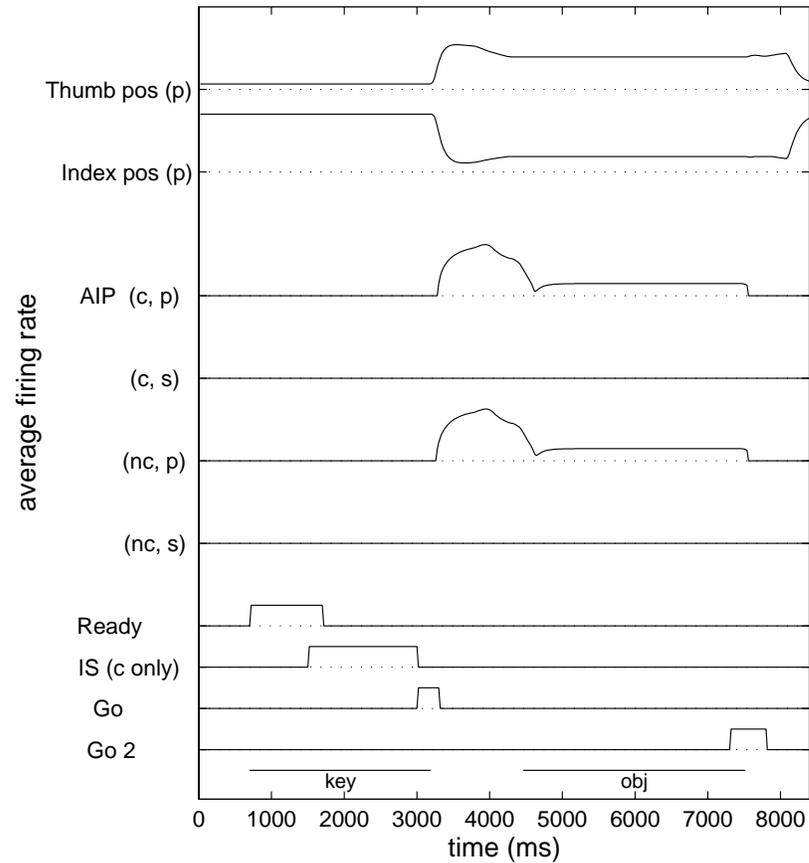


Figure 6.28: A pure motor-related AIP cell, which is active only under the precision-pinch condition.

is the competition that takes place within F5 between the different grasp alternatives. However, note that despite the fact that many cells are active for both grasps, most still show a significant orientation towards one of the two possibilities.

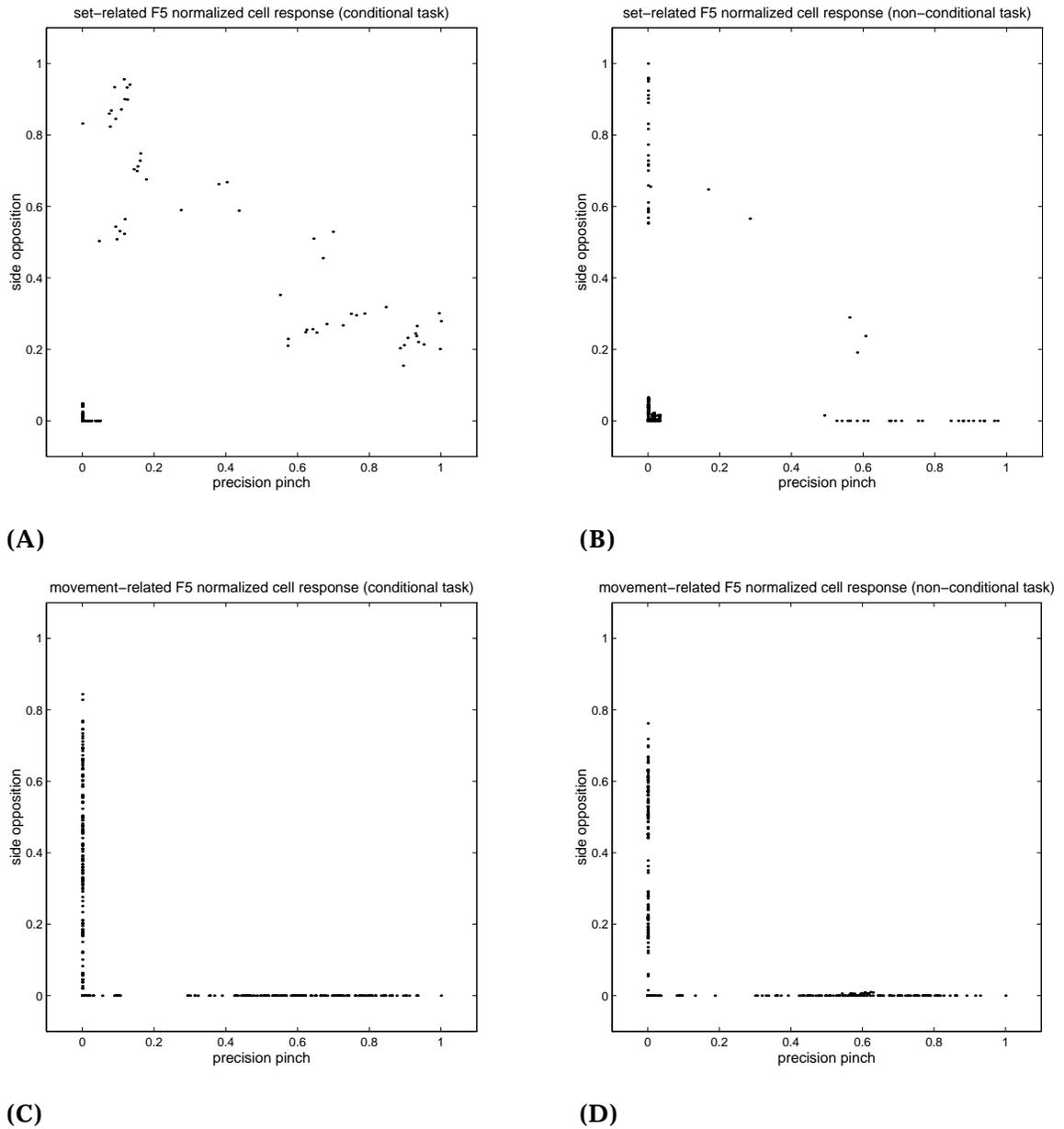


Figure 6.29: Comparison of set-related F5 responses towards two different grasps under conditional (A) and non-conditional (B) tasks; and movement-related responses under conditional (C) and non-conditional (D) tasks.

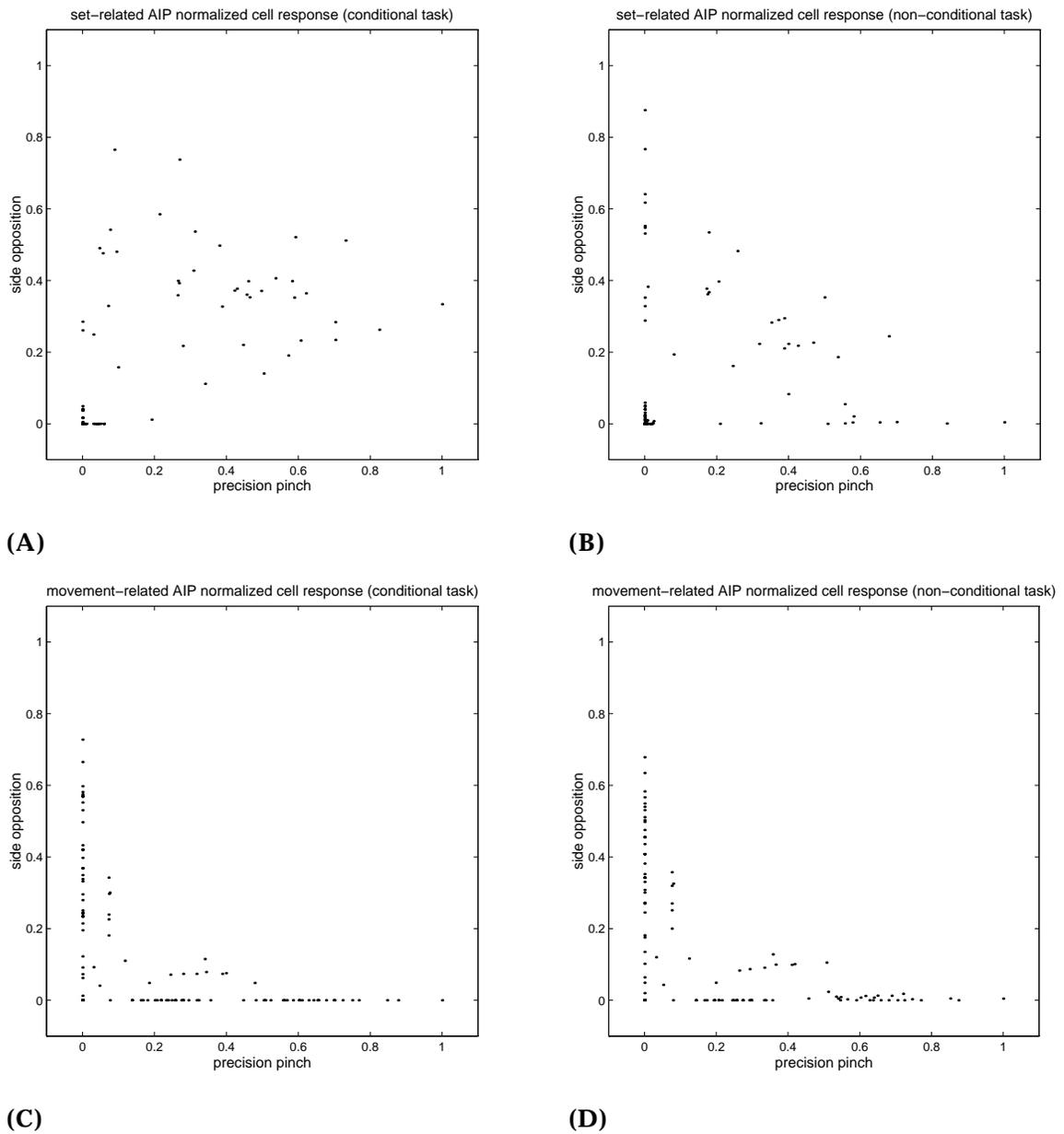


Figure 6.30: Comparison of set-related AIP responses towards two different grasps under conditional (A) and non-conditional (B) tasks; and movement-related responses under conditional (C) and non-conditional (D) tasks.

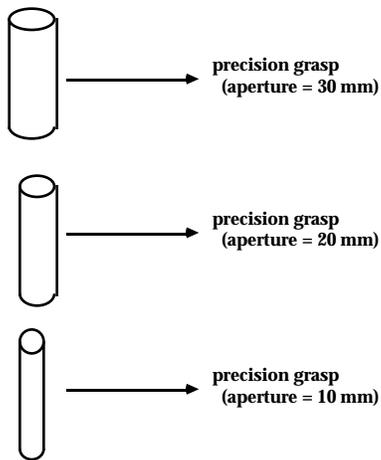


Figure 6.31: Cylinders of different widths map to a precision grasp of varying aperture size.

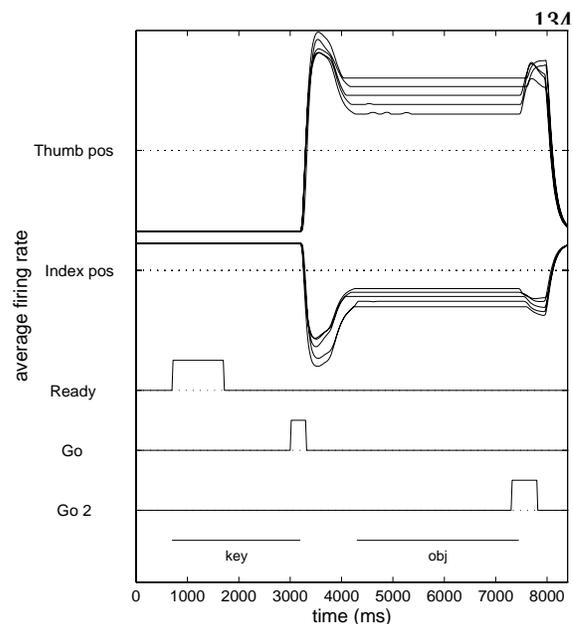


Figure 6.32: Thumb and index finger temporal behavior as a function of cylinder size.

### Prediction #3: Object Size Coding

When humans reach to grasp an object, the maximum aperture achieved by the hand during preshape is related to the anticipated size of the object (Paulignan, et al., 1991a). How is this information encoded in the various grasping motor programs? (Murata, et al., 1993) has reported that within AIP, some cells are modulated by the size of the object that is presented to the monkey. As discussed in chapter 5, we extrapolate this observation to F5, assuming that sub-populations of grasp-specific units encode the aperture of grasp (using a population encoding scheme).

In this experiment, we present the model with a set of cylinders which range in diameter from 10 to 40mm (Figure 6.31), and examine the resulting responses of a set of AIP and F5 units. Figure 6.32 compares the behavior of the thumb and index finger under these different conditions. As the cylinder diameter increases, both the maximum aperture and point of contact anticipated by the model increase accordingly.

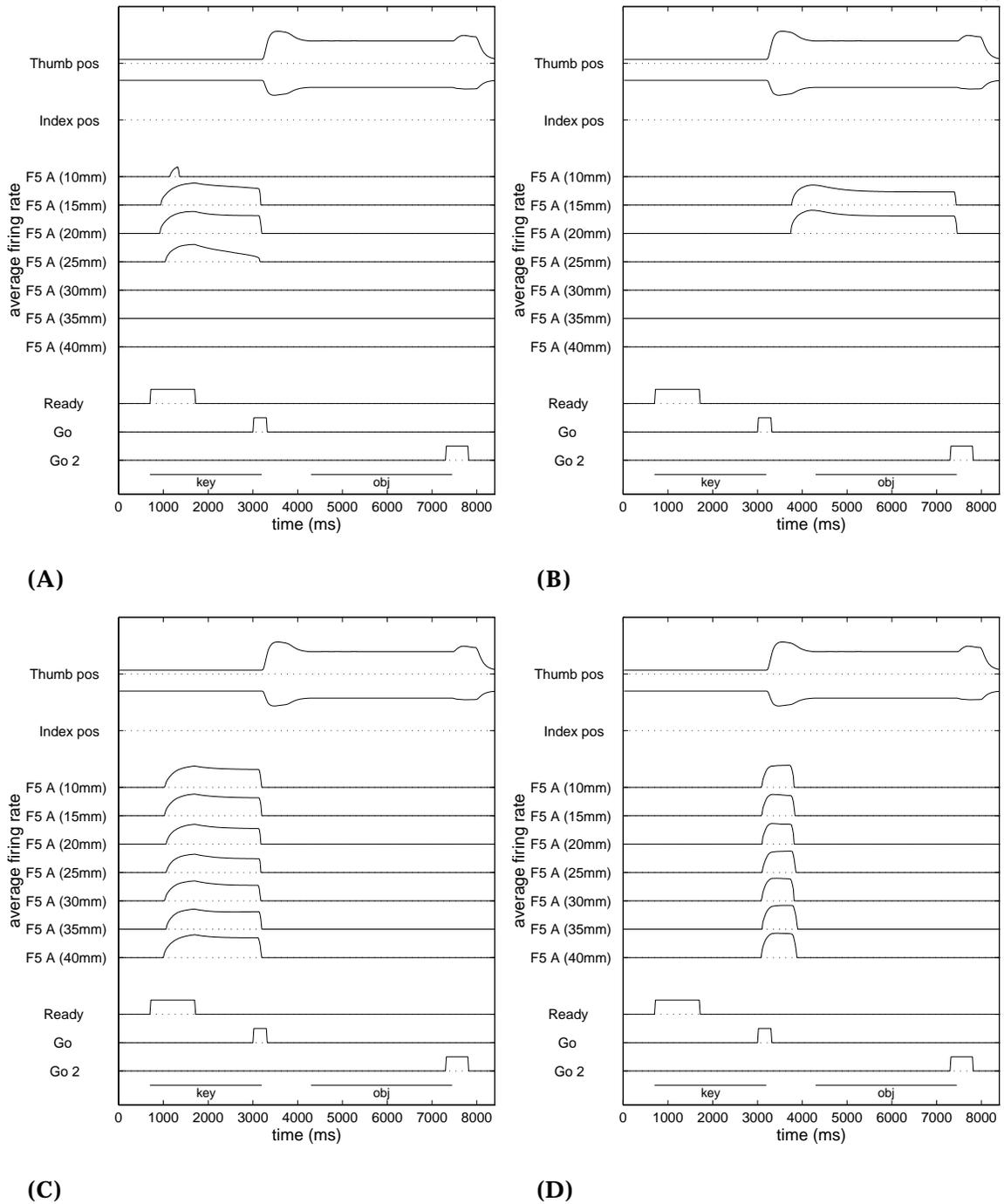


Figure 6.33: F5 cell responses during precision grasps of different apertures. Aperture-specific set (A) and hold (B) units; and aperture non-specific set (C) and movement-related (D) units.

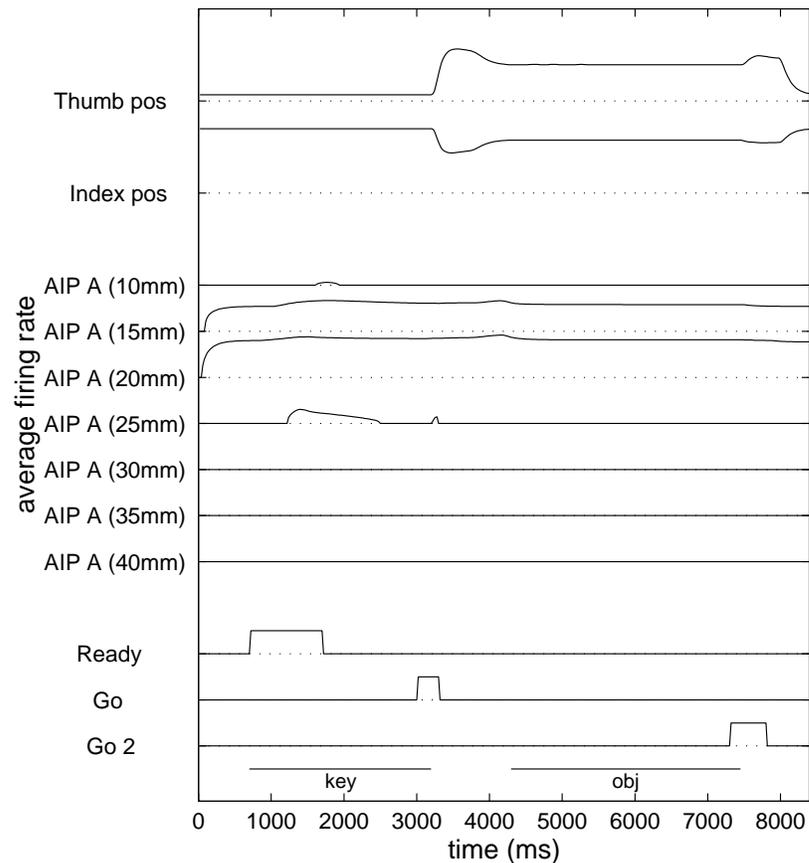


Figure 6.34: Visual-related AIP cell that demonstrates object size specificity (preferred aperture  $\approx$  20mm).

F5 cells exhibiting both aperture-specific and aperture insensitive properties are illustrated in Figure 6.33A-D. Note that both aperture sensitive cells (A and B) are responsive over a range of grasp apertures. In addition, each cell responds maximally for one or two object sizes, and as the objects become larger or smaller than these *preferred sizes*, the response level drops off rapidly. In contrast, cells C and D are not significantly modulated by the size of the object being grasped.

AIP cells also demonstrate a range of behaviors relative to the size of object presented to the system. Sub-populations of both visual- (Figure 6.34) and motor-related (Figure 6.35) units encode object (or grasp aperture) size. Other sub-populations of units encode the precision grasp in general and are only modulated to a small degree by object size (Figures 6.36 and 6.37). Note

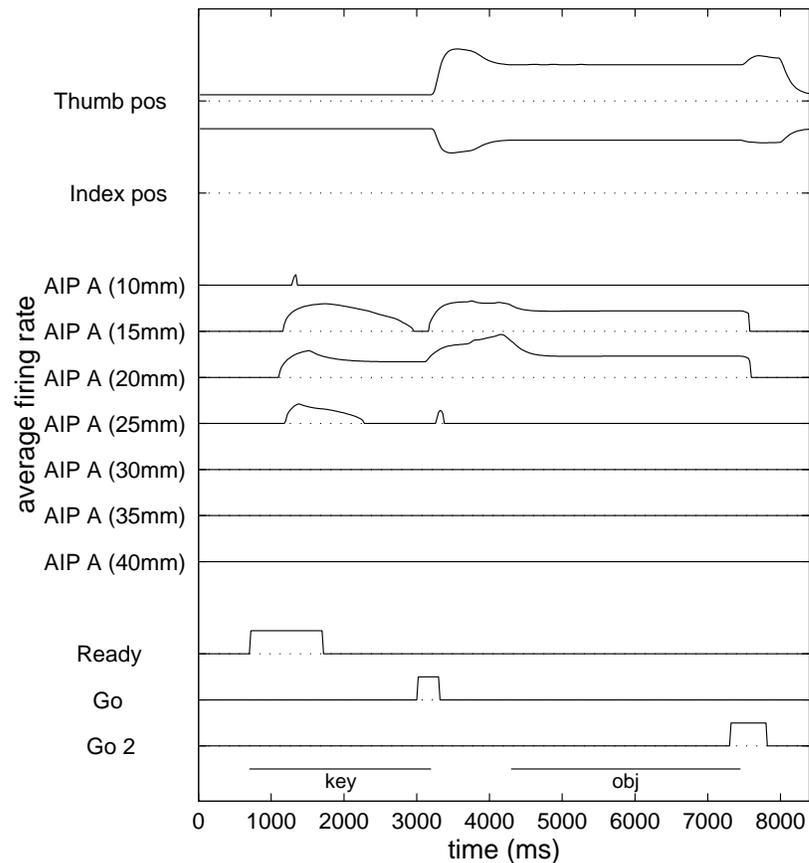


Figure 6.35: Motor-related AIP cell (preferred aperture  $\approx$  20mm).

that in the latter case, the visual/motor cell is shunted during the hold phase of the task for mid-range cylinder sizes, but is active for small and large sizes of cylinders. In addition, the cell exhibits a late onset for only the cylinder of width 35mm. These effects are due to a combination of the specific group of F5 units from which this cell receives recurrent projections, and to the general competition that takes place within the AIP layer (via the BG inhibitory loop).

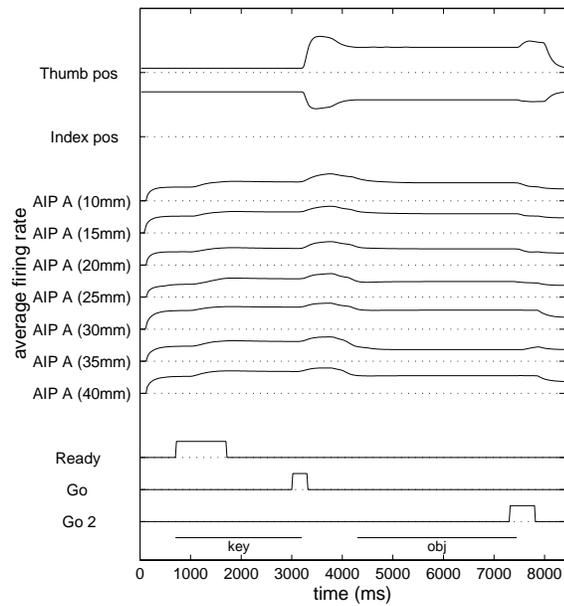


Figure 6.36: Visual-related AIP cell that demonstrates little modulation due to object size.

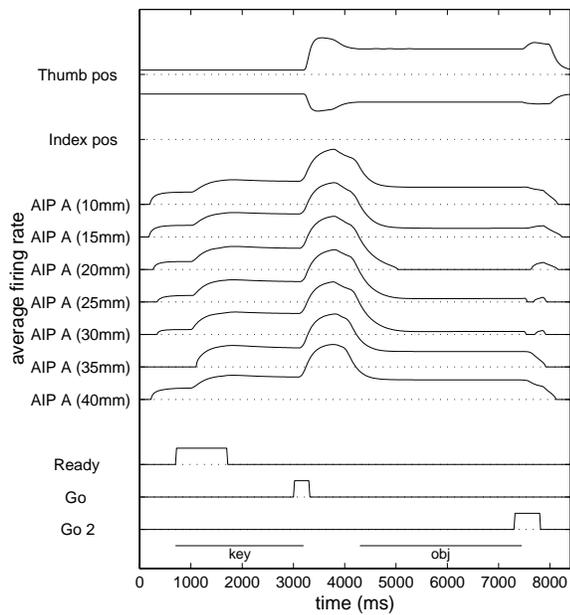


Figure 6.37: Visual/motor AIP cell that is inhibited during the hold phase of movement for mid-range cylinders.

#### **Prediction #4: Object Size Perturbation**

One perspective on AIP's role in the grasping process discussed earlier is the idea that AIP is maintaining an *active memory* of the grasp that is about to be or is currently being executed. If this is the case, then we expect that any changes in the executed grasp (from the planned grasp) must somehow be detected and then reported back to AIP. Take for example the case in which there is a miscalibration between the visually estimated size of the object and the appropriate grasp aperture (as in the size perturbation task of (Gentilucci, et al., 1995)). In the model, this sort of perturbation is sensed by hyperfeature detectors in SII at the time of contact with the object. Because the hyperfeature units are sensitive to very specific combinations of tactile and proprioceptive (population-coded) inputs, a perturbation in object size results in a shift in the SII activity pattern (for small perturbations, the activity patterns for the perturbed and unperturbed cases can be overlapping). This shift in SII contact triggering input into the F5 layer causes a different set of F5 (hold) units to be activated than would normally be the case. This new set of F5 neurons comprises units that are tuned for holding an object that is the same size as the object actually encountered.

This shifting of activity from one grasping program to another is demonstrated in Figure 6.38. Average firing rate traces for two units (A and B) are shown for four different conditions (grasping of a 20mm cylinder, a 30mm cylinder, perturbation of 20 to 30mm and perturbation from 30 to 20mm). Unit A is a movement-related cell, and hence turns off as contact with the object is made. Because the perturbation is not detected until time of contact, the cell behaves in the same manner for both trials involving seeing the 20mm cylinder. Unit B, on the other hand, is a hold-related cell. Although its preferred aperture is 20mm, it also activates for a short period of time when a 30mm grasp is established.

When the cell is presented with a perturbation, its activity pattern switches to match the size of object that is actually grasped. Note, however, that the onset of the cell (comparing the

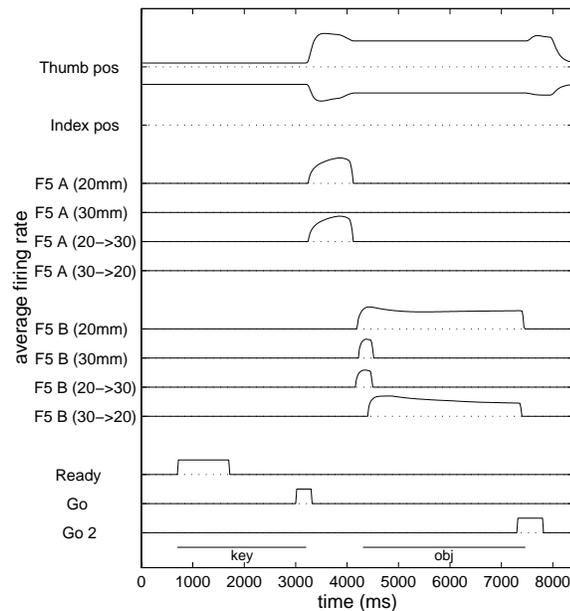


Figure 6.38: F5 movement-related cell (A) and a hold-related (B) cell during the perturbation experiment. 20mm/30mm traces correspond to presentation and grasping of a 20mm and a 30mm cylinder, respectively; traces labeled 20->30 and 30->20 indicate perturbation trials, in which a 20mm cylinder is switched for a 30mm cylinder, and a 30mm cylinder for a 20mm one, respectively.

20mm and 30->20 conditions) is delayed by about 200ms. This delay is due to a combination of 1) the priming of the SII unit in anticipation of contacting a 20mm object (which yields quicker activation of the SII unit in the non-perturbed case), and 2) the time required to "derail" (shunt) the ongoing execution of the 30mm grasping program in F5.

Examining the population responses for F5 cells, we see that not all F5 cells shift in response to a perturbation. In Figure 6.39, we plot the hold-related activity of F5 units that have a preferred aperture between 20 and 25mm. Comparing the responses for 20 and 30mm grasps, we see that the two grasps share some subset of cells (panel A), due to the population coding of grasp aperture. In general, however, these cells are more oriented towards coding a precision grasp of aperture 20mm. When a perturbation occurs (panel B), a subset of these cells react by becoming even more active, thus shifting the currently executing program from the 30mm precision grasp to the 20mm grasp.

This shift in F5 activity subsequently causes a change in the AIP representation of the ongoing grasp. Here we illustrate the main classes of AIP responses. The visual-related cell demonstrated in Figure 6.40 responds during the set phase to both objects nearly at the same level. However, during movement in the 20mm object (perceived size) conditions, this cell is shunted; the activity returns when a grasp is established with an object of size 20mm.

The motor related unit of Figure 6.41 demonstrates a clear movement-related response that is oriented towards objects of perceived size of 30mm. However, the cell maintains a hold response only for 20mm (actual size) objects. In contrast the visual/motor cell of Figure 6.42 is highly active in all four conditions, but maintains a significant level of activity during the hold phase only when an object of 30mm is being held.

Finally, some cells demonstrate a specific behavior for only one of the four tested conditions (Figures 6.43 and 6.44). In the first case, the unit activates during the set phase in 20mm conditions, but only maintains this state during the hold phase when the object size matches the perceived size. In the latter case, the cell is shunted during the hold phase of the task only when an object of 30mm is both perceived and grasped.

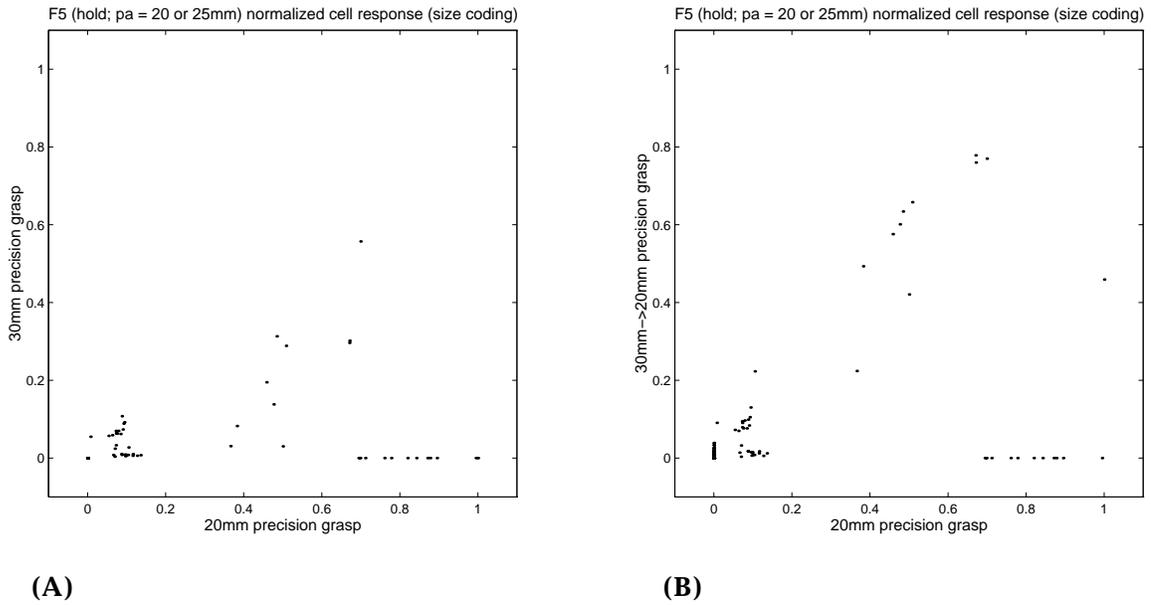


Figure 6.39: F5 cell responses (of cells with preferred aperture between 20 and 25mm only) during execution of the hold phase of a 20mm precision grasp, compared with both the execution of a 30mm precision grasp (A) and a perturbed grasp (B).

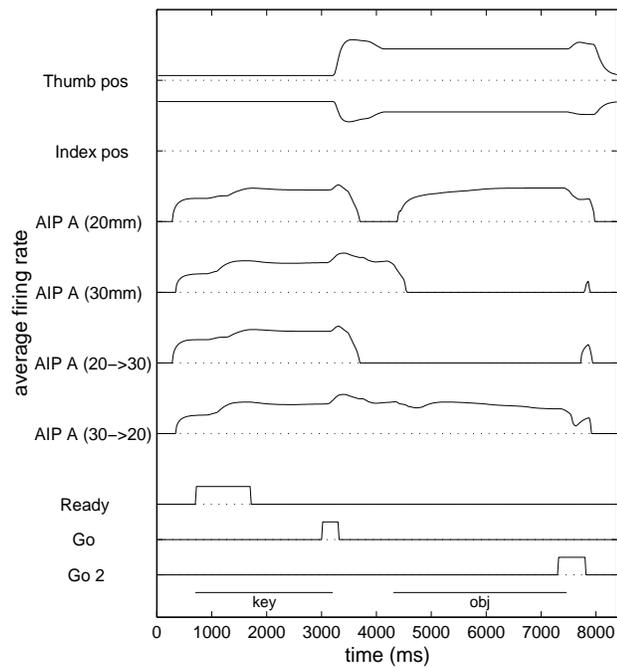


Figure 6.40: AIP visual cell that is active for all four conditions, but is shunted during the hold phase of a 30mm object.

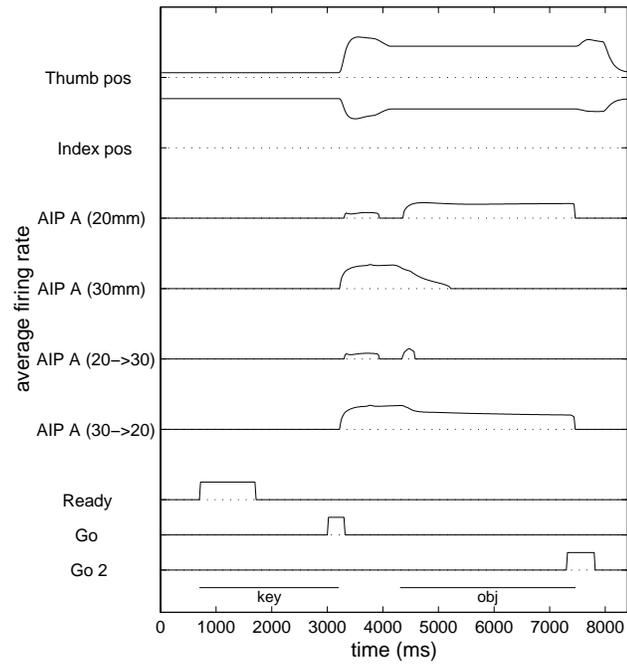


Figure 6.41: AIP motor-related cell that is most active during the movement phase of 30mm aperture grasps, but only continues to be active through the hold phase of 20mm grasps.

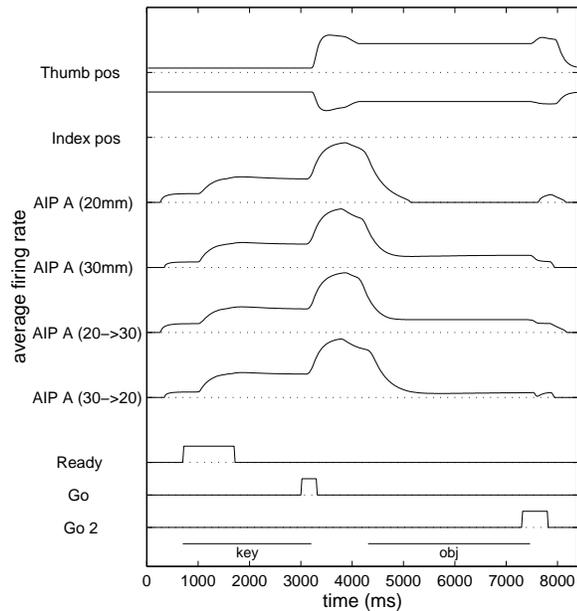


Figure 6.42: AIP visual/motor cell that is active in all conditions. Hold phase activity is most significant in cases in which a 30mm grasp is executed.

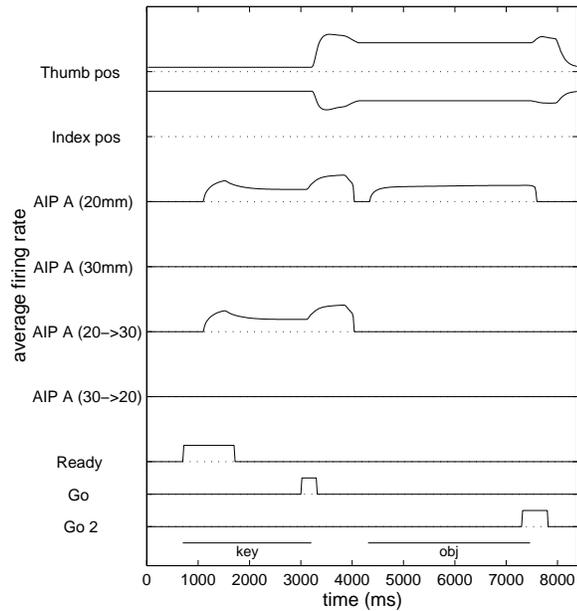


Figure 6.43: AIP motor-related cell that initially responds to the presentation of a 20mm stimulus, but only continues to be active through the hold phase in the unperturbed case.

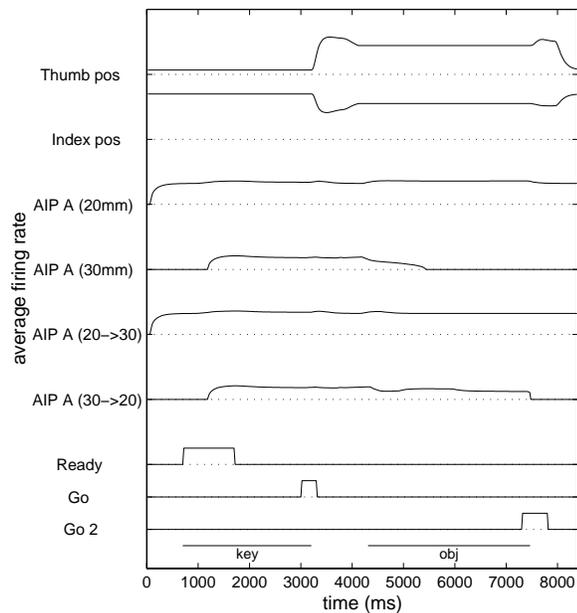


Figure 6.44: AIP (primarily) visual cell that is shunted during the hold phase only when a 30mm object is presented and manipulated.

### Prediction #5: Object Parameter Coding

In looking carefully at object size-related activity described by (Murata, et al., 1993), one observes that during the experiments, when objects of different sizes were tested, the objects were scaled simultaneously in all dimensions. If we believe that AIP is really computing something like the affordances of the attended object, then we would expect that only the properties of the object that are relevant for physically interacting with it should be captured in AIP. For example, when objects are grasped in particular ways, certain dimensions of the objects are not as relevant as others to selecting the grasp aperture.

In this experiment, we present the model with a series of four different box-shaped objects, each of which must be grasped along the indicated horizontal axis (Figure 6.45). The two boxes of the leftmost column are to be grasped using a precision pinch of aperture 10mm, although they differ significantly in size in the other two dimensions. Likewise, the boxes of the rightmost column must be grasped using a precision pinch of aperture 20mm. The key question to be asked is how these four objects are coded by the AIP cell visual and motor responses.

Figure 6.46 A compares the visual-related population responses for the two objects in the leftmost column. All active AIP units demonstrate a significant level of activity for both objects, with little variation between the two. This is the case even though the two objects match in size along only the width dimension. However, comparing the responses towards the two objects in the bottom row of Figure 6.45, we see a much wider distribution of responses (with some cells responding exclusively to one object or the other) (panel B). These differences in response properties are due to the fact that in the model, objects represented in the

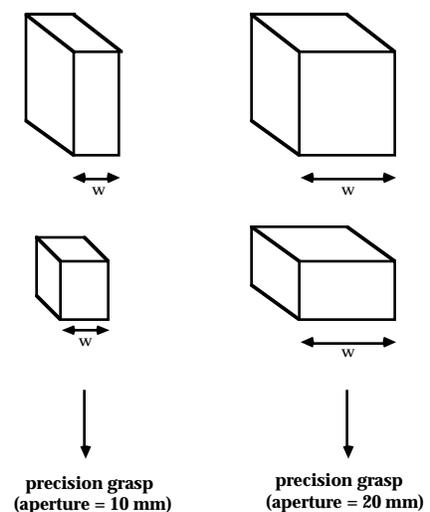


Figure 6.45: Four boxes of different dimensions.

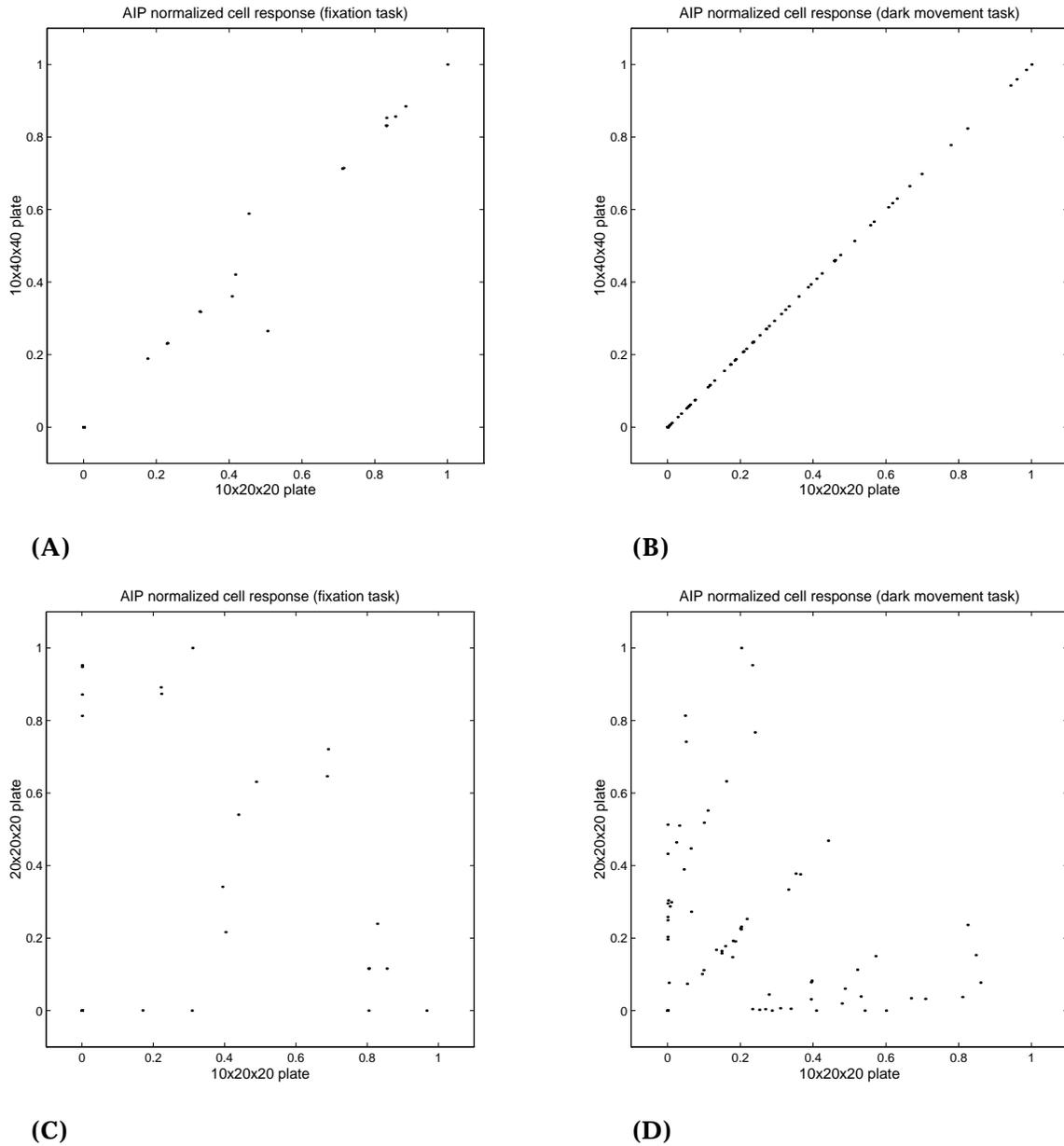


Figure 6.46: Comparison of AIP visual responses for objects of the same (A) and different (B) widths; and AIP motor-related responses (dark movement condition) for objects of same (C) and different (D) widths.

visual areas map to AIP cells that capture the critical parameter for physical interaction (the grasp aperture in this case).

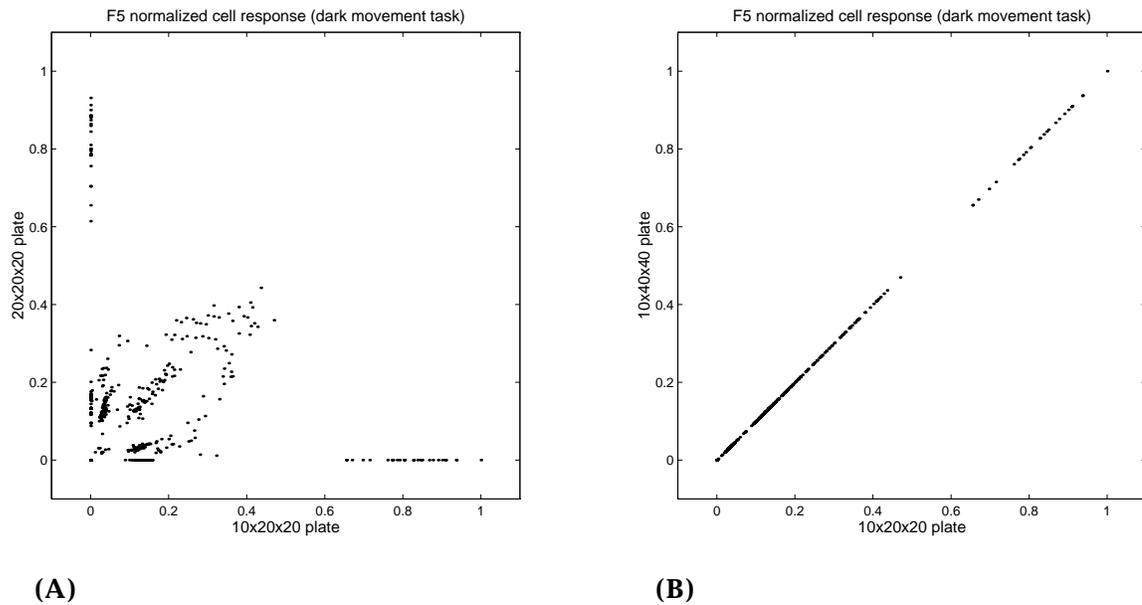


Figure 6.47: F5 population responses for two objects of the same width as measured along the grasp axis (A), and two objects of different widths (B).

Figures 6.46 C, D demonstrate the motor-related AIP population responses for the same pair of objects. As was seen in an earlier section (Prediction #1), the motor related responses tend to shed any object-specific properties not related to the grasp. For objects requiring identical grasps, the motor-related responses are identical (C); objects mapping to different grasps demonstrate grasp-specific modulation (D). In this case, however, many AIP motor-related cells are shared between the two objects. These shared units are of two types: 1) general precision grasp units (with no aperture specificity), and 2) aperture-specific cells that respond to both conditions due to the population code used to capture the aperture parameter. These latter responses compare closely with the F5 population responses of Figure 6.47.

## Object versus Grasp Coding in the Monkey AIP

The monkey experiment reported in (Taira, et al., 1990) looked at AIP responses to a variety of objects and conditions. In one case, a pull-knob was presented to the monkey; the knob was mounted such that it protruded from a flat plane (fourth row in Figure 6.48). In general, the monkey preferred to grasp the knob using a side opposition. In a second case, an identical knob was presented, but was inset into a vertically cut slot (third row). In this case, the slot forced the monkey to adopt a precision pinch in order to successfully grasp the knob.

This task compares to some degree to the conditional experiment described earlier that was performed with the model (Prediction #2). In the monkey experiment, an object (the knob) affords two different grasps (the side opposition and the precision pinch). The selection of one of the two possibilities, however, is not done as a function of an abstract instruction stimulus, but rather as a function of task constraints (in this case, the reachability of the knob). Because the visual description of the knob is very similar in the two cases (discounting the representation of the slot), the model predicts that visual-related responses in AIP should be the same. However, because two different motor responses are generated, the model predicts that movement-related AIP firing patterns should reflect the type of grasp, and hence the condition that is presented to the monkey.

The AIP movement-related cell shown in Figure 6.48, which was recorded during the monkey experiments, demonstrates a selectivity (in the movement conditions) towards the pull-knob. However, the cell is very significantly active when a precision pinch is elicited, and is only slightly active during execution of the side opposition. On the other end of the spectrum, the visual-related cell of Figure 6.49 demonstrates only a small difference in firing rate during object fixation of the two pull-knobs. But, when movement is generated in the light, a significant difference emerges in the two cases.

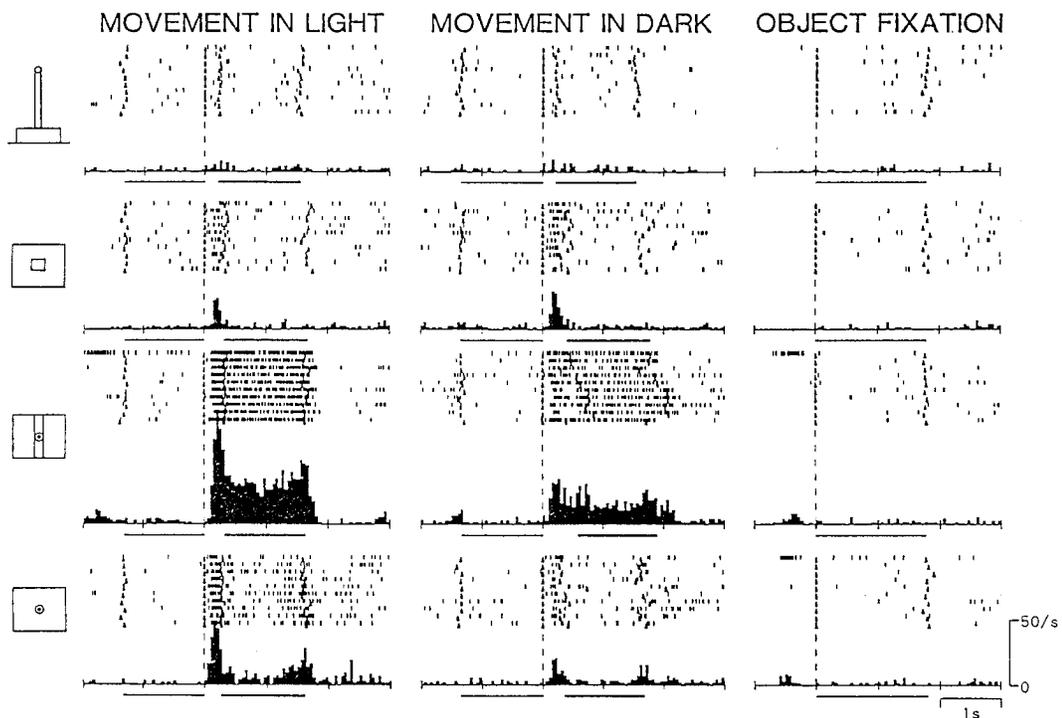


Figure 6.48: Response of a single monkey AIP cell to four objects (joystick, push-button, pull-knob in a slot, and a pull knob) and three conditions. The cell demonstrates motor-related behavior, and is most responsive when a precision pinch is used to grasp a knob. The experimental details are described in (Taira, et al., 1990); actual data from (Sakata, 1994).

Finally, the cell shown in Figure 6.50 is active in both fixation and dark movement conditions (and hence is considered a visual/motor cell). Although the cell demonstrates some difference in the firing rate for the fixation condition depending upon which knob is presented, the difference becomes much more significant when a movement is executed (either in the light or the dark).

In a more recent recording experiment, (Murata, et al., 1993) has looked at AIP responses to a set of different geometric primitives. In some cases, different primitives (e.g., a cylinder, and a rectangular block), which were approximately of the same dimensions, elicited the same type of grasp from the monkey. Under these conditions, the model predicts (Prediction #1) that visual-related responses can show alignment towards one or the other object. However, because

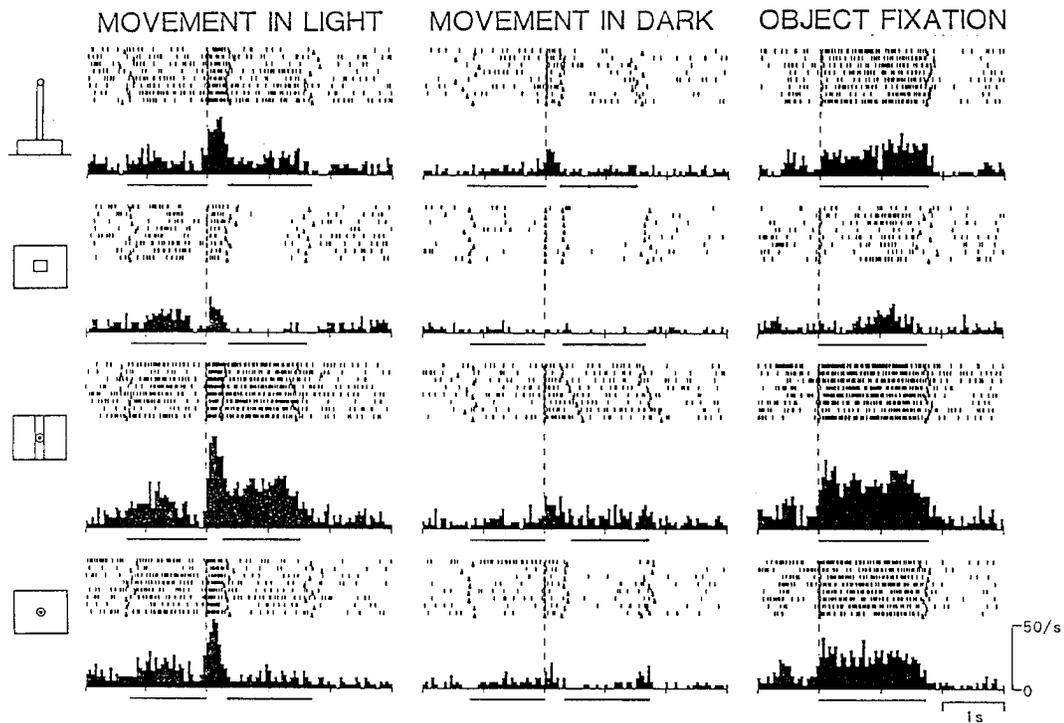


Figure 6.49: Response of a primarily visual-related AIP cell. Data from (Sakata, 1994).

the generated movement is identical in the two conditions, movement-related responses will lose the object specificity. Although the full analysis of this effect in the monkey data has not been performed, the few cells that have been studied in detail demonstrate the predicted behavior (Murata, 1995).

## Summary

In this chapter, we have demonstrated the neural-level behavior of the model in a variety of experiments. We first illustrated the general behavior of the model and then made comparisons to single unit recordings performed in monkey F5 and AIP. Finally, we performed several novel experiments, making explicit predictions about expected single unit and population behavior. These experiments addressed the following key questions:

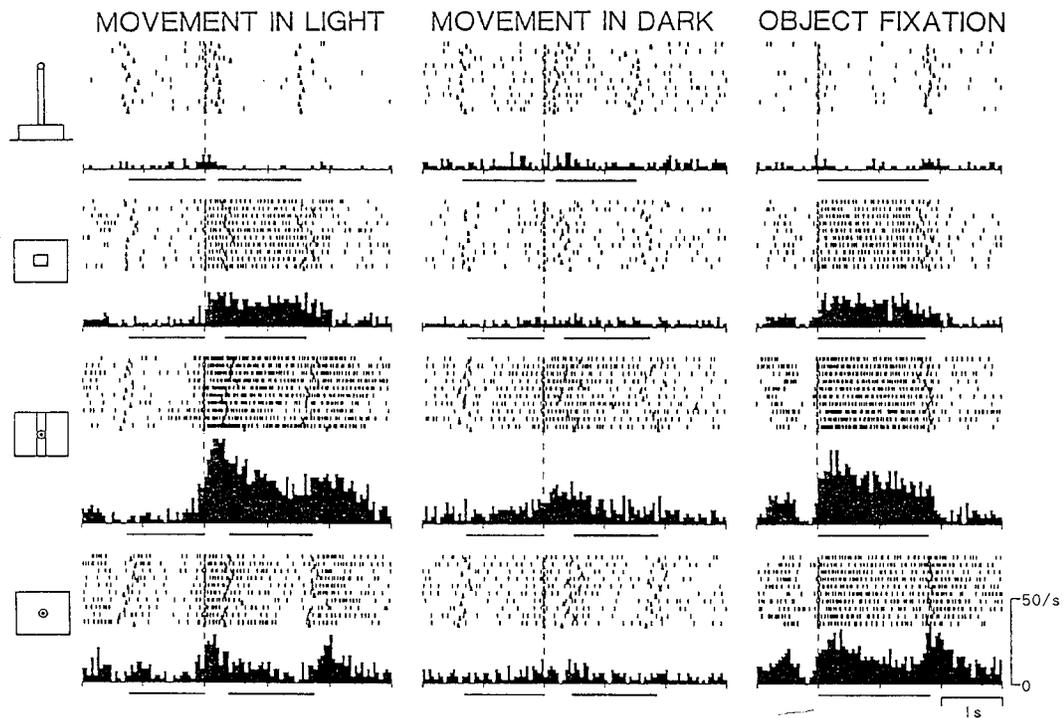


Figure 6.50: Response of a visual/motor AIP cell. From (Sakata, 1994).

1. What is the neural code in AIP that is used to represent different objects that map to the same grasp? The model predicts that an AIP cell's orientation (either object- or grasp-centered) is directly related to its visual/motor classification (Taira, et al., 1990). This arises from the fact that in the model visual-related responses are due to direct projections from the visual areas (PIP or IT), in which different objects can be coded using significantly different patterns of activity. The motor-related responses rely on recurrent projections from F5, which do not carry non-relevant object properties (from the point of view of the grasp).

2. What is the representation and conflict resolution process necessary for dealing with objects that can be grasped in more than one way? When the model is presented with an object of this type, and an instruction stimulus (IS) is used to determine which of the grasps should be used, then a) prior to the presentation of the IS, activity patterns in F5 and AIP reflect all

grasp/affordance possibilities; and b) after presentation of the IS, all but the selected grasp/affordance are shunted.

3. How are object/grasp parameters encoded in AIP and F5 (specifically object width and grasp aperture)? Aperture size is encoded in F5 and AIP (in the model) using sub-populations of cells (belonging to populations that are specific to the grasp/affordance pairs). The sub-populations encode the aperture parameter using a spatial code, in which each cell has a "preferred" aperture. Note that this effect is built directly into the model.

4. What is the behavior of the system when the object size is perturbed? When the apparent size of the object does not match the actual size, we see a shift in activity at the time of contact (first in F5, then in AIP) from the pattern that represents the grasp program for the expected object to the pattern of activity that represents the program for the object that was actually grasped.

5. Are all object parameters encoded with equal weight in AIP? A natural consequence of the wiring scheme is that when object parameters are explicitly relevant for programming the movement, the model predicts a higher concentration of cells that are involved in encoding that parameter. Parameters that are not relevant for physical interaction with the object will tend to not be captured in AIP.

## Chapter 7: Linking PET Imaging in Humans to the Grasping Model

---

In this chapter, we relate the model of grasp generation, derived in large part from monkey neurophysiological data, to the human functional anatomy involved in grasping. First, a technique referred to as *synthetic PET imaging* is applied to the model, from which we derive a set of predictions for what we expect to observe in a human experiment. We then describe a human PET experiment that looks at the processing of instruction stimuli in a conditional task, as well as at the relative representation of different grasp programs. Finally, through comparison of the synthetic PET predictions and human experimental results, we reflect on how the model may be further refined.

---

### Synthetic PET Imaging

$\text{H}_2^{15}\text{O}$  positron emission tomography (PET) is an imaging technique that allows the measurement of regional cerebral blood flow (rCBF). Because rCBF appears to be related to the local synaptic activity within a region (Brownell, Budinger, Lauterbur, & McGeer, 1982), it is possible to observe differences in the measured rCBF as a function of the task that a subject performs. As a result, the technique has significantly contributed to our recent understanding the functional anatomy of the human brain (e.g., (Barbur, Watson, Frackowiak, & Zeki, 1993; Decety, et al., 1994; Grafton, et al., 1995; Silverman, Grossman, Galetta, Liu, Rosenquist, & Alavi, 1995)).

Current resolution of the technique is on the order of 1.5mm in each spatial dimension, and temporally is about 80 seconds. This level of information differs considerably from what is available in neurophysiological experiments in monkey, where we are able to examine individual cells and resolve single spikes, but have tremendous difficulties in examining entire circuits. It is thus desirable to develop techniques that allow us to draw conclusions in one domain from experimental results in the other.

Synthetic PET imaging (Arbib, Bischoff, Fagg, & Grafton, 1995) has been proposed as one such mechanism for relating these different levels of experimentation through the construction of a computational model. The low-level details of the model described in this thesis have been derived primarily from a combination of neurophysiological (monkey) results and computational constraints. From the model, we extract a measure of regional synaptic activity, which can then be compared to rCBF results in human PET experiments.

### Computation of the Synthetic PET Measure

The synaptic activity of a region A under condition X is computed from the model as follows:

$$SA_A^X = \int \sum_{j \in A} \sum_i a_i(t) |w_{ij}| dt \quad (7.1)$$

where:

$a_j(t) |w_{ij}|$  is the activity of the synapse connecting unit j to unit i at time t.

$i$  iterates over all presynaptic units to unit j (which may or may not be members of region A; see Figure 7.1).

$j$  iterates over all neurons in region A.

The measure of instantaneous synaptic activity in region A is then integrated over the time required to perform a task (which might involve multiple trials).

The simulated synaptic activity of a region can then be compared over several conditions. Here we define the *relative synaptic activity* for region A under conditions X and Y as:

$$rSA_A^q = \frac{SA_A^q}{\text{Max}_{z \in \{X, Y\}} \{SA_A^z\}} \quad (7.2)$$

where:

$$q \in \{X, Y\}$$

And the *change in relative synaptic activity* from Y to X is:

$$\Delta rSA_A^{Y \rightarrow X} = rSA_A^X - rSA_A^Y \quad (7.3)$$

Note that the positive and negative projections into a region are treated equally in their contribution to the synthetic PET measure (due to the absolute value operator applied to the connection weight in equation 7.1). Thus, from the PET measure alone, it is impossible to distinguish the case in which a positive input is given from the situation where an equally negative input is given.

### Synthetic PET Imaging for the Grasping Model

Note that in the model, each column receives three distinct types of projections (priming, support, and trigger). In computation of the synthetic PET measures, the contributions of these different inputs are assumed to be weighted equally, and are simply summed in computing the column's contribution to the region's rCBF measure.

In what follows, we present the results of two different synthetic PET experiments, which

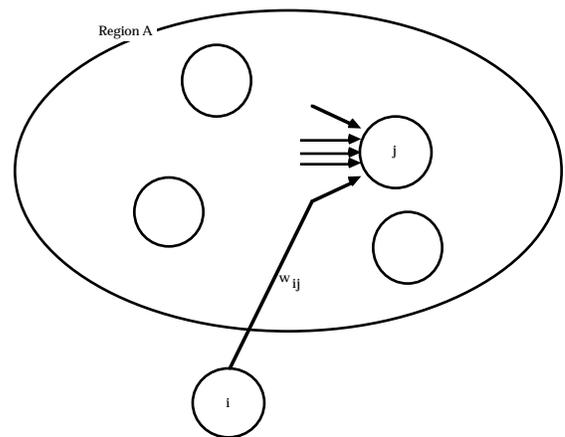


Figure 7.1: Neural elements involved in computing the synaptic activity for a simulated brain region.

serve as predictions for what we expect when the experiments are performed in the human.

### **Conditional Task**

When a delayed instruction stimulus is used to inform a subject how to grasp an object, what brain regions are involved above and beyond the case where the subject knows a priori which of the two grasps to perform? In chapter 6 (prediction #3), we analyzed how the model performs in such a situation, and examined the single cell and population-level behaviors of several key regions. Here, we present the synthetic PET measures taken during execution of a precision grasp under the conditional and non-conditional tasks.

Figure 7.3 A plots the relative synaptic activity measures for the two tasks. Only regions in the model that demonstrate a change in synaptic activity from one task to the other are shown. The most significant change predicted by the model is the level of activity exhibited by area F2 (dorsal premotor cortex). Its high level of activity in the conditional task is due to the fact that this region is only involved when the model must map an arbitrary stimulus to a motor program (in this case, a grasp). In the non-conditional task, the region does not receive IS inputs, and thus its synaptic activity is dominated by the general background activity in the region.

The additional IS inputs in the conditional task have a second-order effect on the network, as reflected in the small changes in activity in F5, BG, and AIP (Figure 7.2). The increased synaptic activity in F5 is due to the additional inputs from F2 (into the supporting inputs of some columns in F5). These inputs also cause an increase in the region's *activity level*, which is passed on through excitatory connections to both AIP and BG (the latter appears in the figure as the recurrent inhibitory connections back to F5 and AIP).

An important ability of the synthetic PET technique is that the positive and negative contributions to the synthetic PET measure can be differentiated. Figures 7.4A,B demonstrate the positive and negative contributions to the overall PET measure in the conditional/non-conditional task comparison. Note that although the positive contributions to the F5 and AIP essentially dominate the full PET measure, we also see small increases

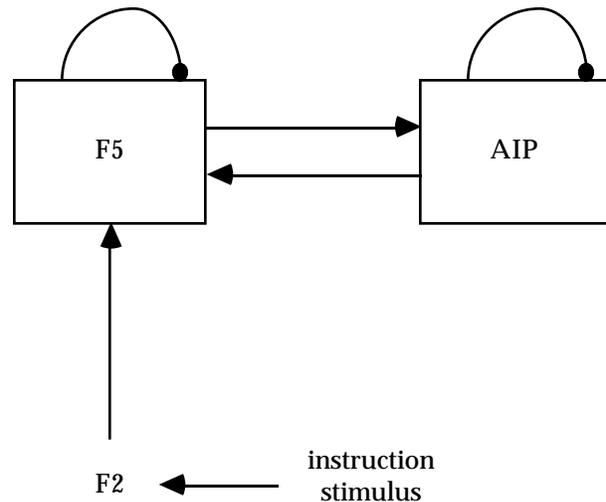


Figure 7.2: Functional interactions of F2, F5, and AIP.

in the negative inputs into these regions. These inhibitory signals are due to the recurrent inputs through the BG. This is additional evidence that both F5 and AIP are experiencing increases in their overall activity (not just of synaptic activity).

### Precision versus Side Opposition

One observation made by (Rizzolatti, et al., 1988) was that in F5, the number of neurons involved in the execution of the precision pinch was greater than those involved in either the side opposition or the power grasp (the fewest neurons were observed in the latter case). In the construction of the model, this information was used to select the distribution of neurons for each type of grasp. An important question is how this distribution is reflected in the PET measures. Here, we compare these measures for the precision pinch and side opposition (the protocol is the same as in prediction #3, although we compare the two non-conditional tasks).

Figure 7.3B illustrates a general increase in synaptic activity in many of the model's regions for the precision pinch case. Although there is an increase in the number of active units only in F5 and AIP (from one condition to the other), we observe a subsequent increase in synaptic

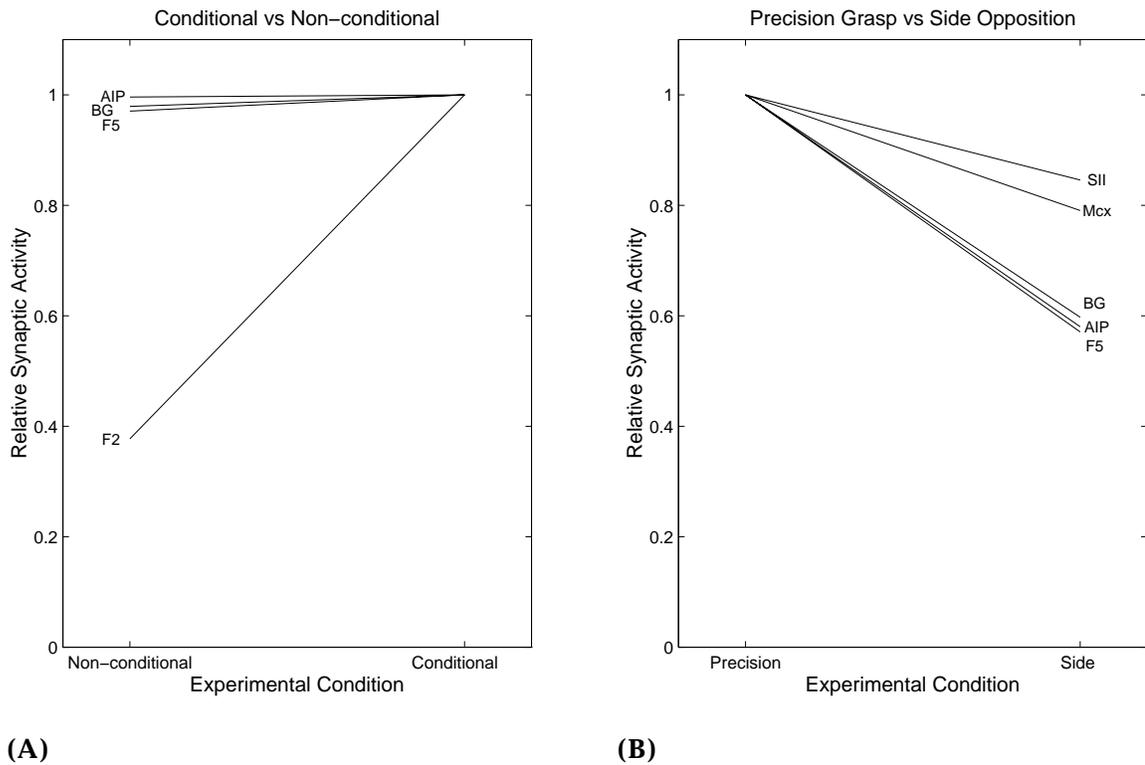
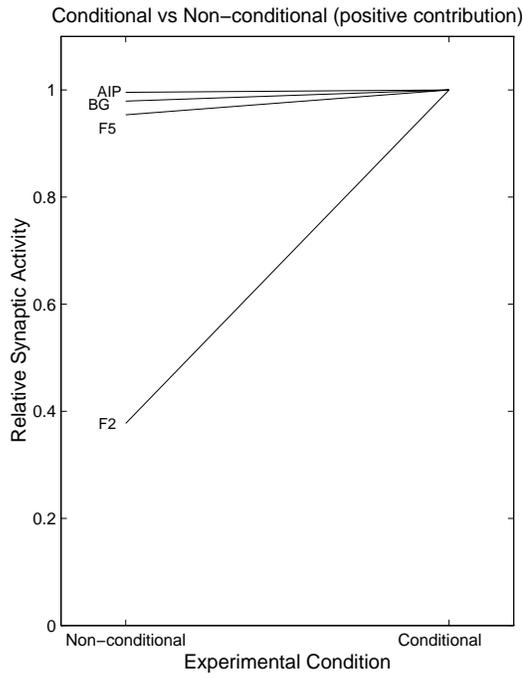
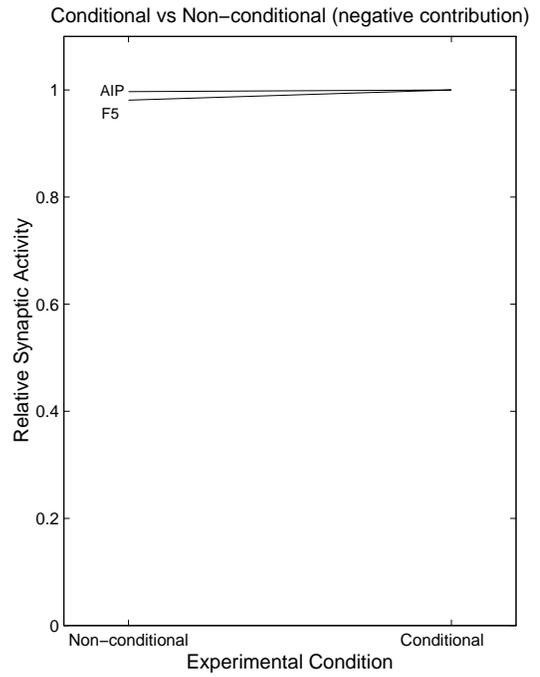


Figure 7.3: Predictions of relative synaptic activity changes for (A) a non-conditional task and a conditional task, and (B) a precision pinch and a side opposition.

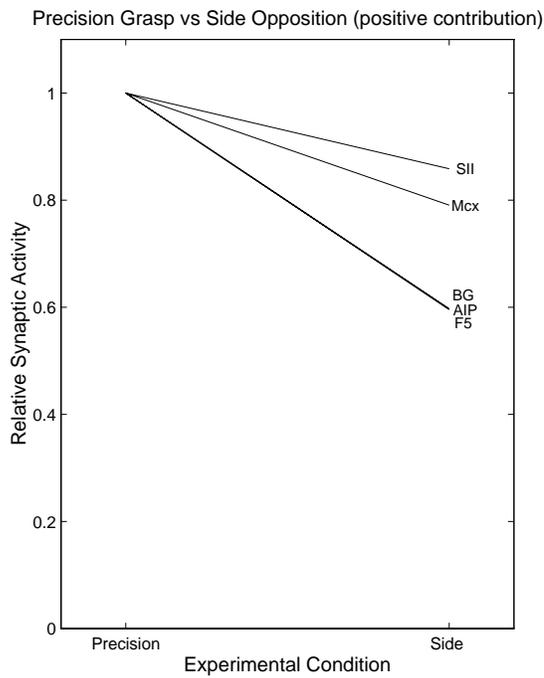
activity in BG, Mcx, and SII. The increase in synaptic activity in BG actually reflects a true increase in the region's activity level, as indicated by the increase in inhibitory inputs back into AIP and F5 (Figure 7.4D).



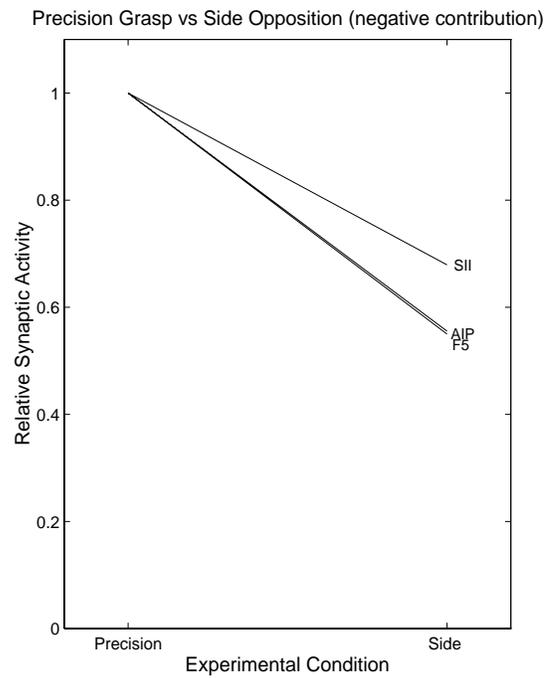
(A)



(B)



(C)



(D)

Figure 7.4: Positive and negative synapse contributions to the synaptic activity measure for the conditional/non-conditional task (A and B), and for the precision pinch/side opposition task (C and D).

## Human Grasping Experiment

The synthetic PET experiments that have been presented have asked some important questions about how instruction stimuli are mapped to arbitrary motor programs, and about the relative representation of different grasps. In this section, we present a human PET experiment in which both of these questions are addressed.

### Materials and Methods

#### Subjects

Six normal, right-handed subjects participated in the study.

#### Tasks

Subjects were asked to repeatedly perform grasping movements over the 90 second scanning period. The targets of grasping were mounted on the experimental apparatus shown in Figure 7.5. Each of three stations mounted on the apparatus consisted of both a rectangular block that could be grasped using a power grasp, and a pair of plates (mounted in a groove on the side of the block; see inset of Figure), which could be grasped using a precision pinch (thumb and index finger). A force sensitive resistor (FSR) material, mounted on the front and back of the block, detected when a solid power grasp had been established. The two plates were attached to a pair of mechanical micro-switches, which detected when a successful precision pinch had been executed. A bi-colored LED at each station was used to instruct the subject as to the next target of movement. A successful grasp of this next target was indicated to the subject by a change in the color of the LED. The subject then held the grasp position until the next target was given. Targets were presented every  $3 \pm 0.1$  seconds.

In the first grasping condition, subjects repeatedly performed a power grasp to the indicated block. The target block was identified by the turning on of the associated LED (green in color). When the subject grasped the block, the color of the LED changed from green to red. For the second condition a precision pinch was used; the target was identified in the same manner as the first condition. In the third grasping condition (conditional task), the initial color of the LED instructed the subject to use either a precision pinch (green) or a power grasp (red). When contact was established, the LED changed to the opposite color.

In the fourth (control) condition, the subjects were instructed to simply fixate on the currently lit LED and not make movements of the arm or hand. The lit LED changed from one position to another at the same rate and variability as in the grasping task.

Each of the four conditions were performed by each subject three times. The order of condition presentation was selected at random. In addition, the order of trial presentation was randomly selected, as was the LED color in the conditional task.

Prior to scanning, subjects practiced each of the grasping conditions for several minutes.

### Performance Measures

Reaction time data was collected by the control computer as the subjects performed the task. The computer logged time required to release the current object in response to a new target, the time required to establish a grasp on the new

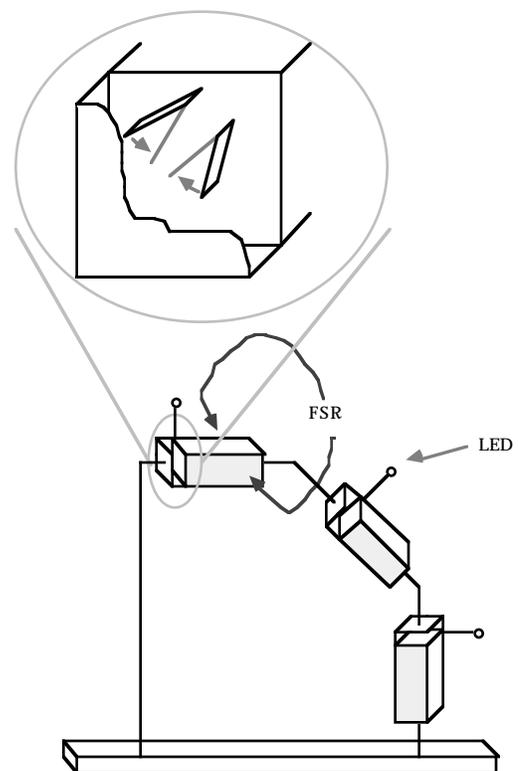


Figure 7.5: Apparatus used in PET experiment. Each of three stations can be grasped in two ways: precision pinch of the two plates in the grove (inset), or power grasp of the block.

object, and whether the established grasp was correct. In addition, the subject's hand position was tracked using a Polhemus tracking device (Polhemus, Colchester, VT), which measures six DOF of wrist position. A VPL DataGlove (VPL Research, Inc) measured the joint angle of the two proximal joints of each finger. Both types of data were sampled at 30 Hz.

### **Imaging**

rCBF images were collected using a Siemens 953/A tomograph. Subjects were injected intravenously with  $\text{H}_2^{15}\text{O}$  at the start of scanning and performance of the behavioral task.

### **Image Analysis**

Prior to statistical analysis, the rCBF images were aligned to a reference MRI atlas that was defined relative to the Talairach coordinate system (Talairach & Tournoux, 1988). This alignment was accomplished by first performing a within-subject alignment of the subject's 12 scans using an automated registration algorithm (Woods, Cherry, & Mazziotta, 1992). Next, a mean PET image for each subject was computed. The mean PET image was then coregistered with the standardized MRI atlas (Woods, Mazziotta, & Cherry, 1993). Finally, the individual rCBF images were resliced (combining the within-subject and mean PET image-to-atlas transformations) into the coordinate frame of the standardized MRI atlas.

The rCBF images were then smoothed. A common volume mask was computed, which marks the voxels for which data is available in all scans (from all subjects). The smoothed images were then normalized to each other.

### **Statistical Tests**

Task-related effects in rCBF were extracted using a 2-way analysis of variance (ANOVA) applied to the image set. The sources of variance in the model were *task* and *subject* (repetitions were assumed to not be a source of variance).

An F-map of task effects was finally computed for each voxel, the results of which are thresholded and overlaid with the MRI atlas.

## Results

### Behavioral Results

For each trial during the scanning, subject reaction time to target presentation (turning on of the LEDs) were measured. Both the time required to release the currently held object and the time required to grasp the next were recorded. Figure 7.6A demonstrates that there was relatively little effect on reaction time as a function of task, although the conditional task required on average about 119ms longer than the non-conditional tasks (well within 1 standard deviation). Hence, the amount of time involved in transport of the hand and establishment of the grasp is essentially constant across the tasks. In addition, little effect was observed when precision grasp trials were compared to power grasp trials in the conditional task (B).

Finally, little change was seen in performance as a function of task repetition (C, D, and E). This indicates that little or no learning occurred during the course of the scans, and we can discount any repetition effects in our analysis.

### Imaging Studies

The set of regions more significantly active ( $p < 0.005$ ) in the conditional task as compared to the movement conditions (combination of precision and power tasks) are shown in Figure 7.8 <sup>1</sup>; the most significant ( $p < 0.001$ ) regions are summarized in Table 7.1.

The most highly active region in this comparison is the contralateral dorsal premotor cortex. In monkey, this region is thought to be involved in the arbitrary association of stimuli with the preparation of motor programs (Evarts, et al., 1984; Kurata & Hoffman, 1994; Kurata & Wise, 1988; Mitz, et al., 1991; Wise & Mauritz, 1985), although this region has been traditionally thought of as a limb-related area, and not one involving distal movements. The ipsilateral dorsal premotor cortex is also active, although the degree of activity is

---

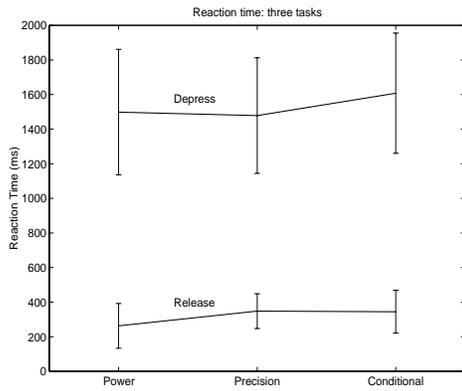
<sup>1</sup> The color plates corresponding to Figures 7.11 and 7.10 may be found at <http://www.usc.edu/dept/robotics/personal/af0a/thesis/>

significantly less, and the centroid of activity is more mesial and frontal from the contralateral side.

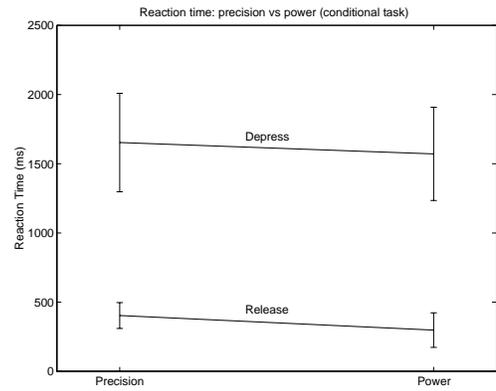
A site in the contralateral inferior parietal cortex is also significantly active. This region is located near the anterior bank of the intra-parietal sulcus, and may be related to the monkey AIP. In addition, several visual processing sites are reasonably active: two distinct locations in the ipsilateral area 18, and one area in the superior occipital lobe (near the area 18/19 boundary). Finally, a site within the cerebellar cortex shows moderate activation.

The areas more significantly active ( $p < 0.01$ ) in the precision grasp condition as compared with the power grasp condition are shown in Figure 7.9; the most significant ( $p < 0.005$ ) sites are summarized in Table 7.2. The most significant sites of activation are in the contralateral superior occipital lobe (identical site as in the above comparison), and the contralateral supplementary motor area (SMA-proper; near the boundary with area 4).

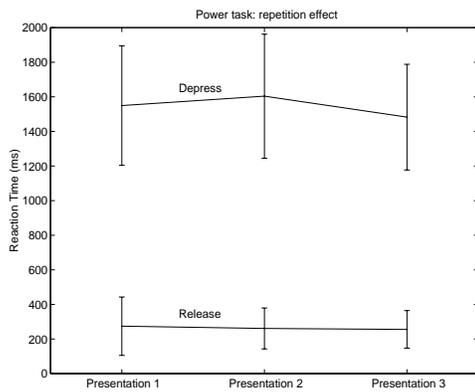
Less active, but still above threshold, include the ipsilateral dorsal premotor cortex (different site than in the conditional comparison), the contralateral cerebellar vermis, contralateral inferior parietal cortex, contralateral area 18, and contralateral area 7. As in the conditional comparison, the inferior parietal cortex site is located near the anterior bank of the intra-parietal sulcus. However, this site is more lateral and slightly more frontal.



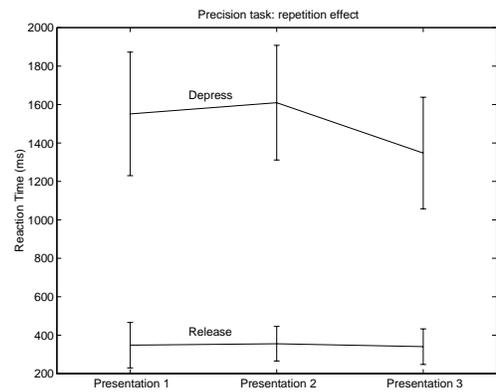
(A)



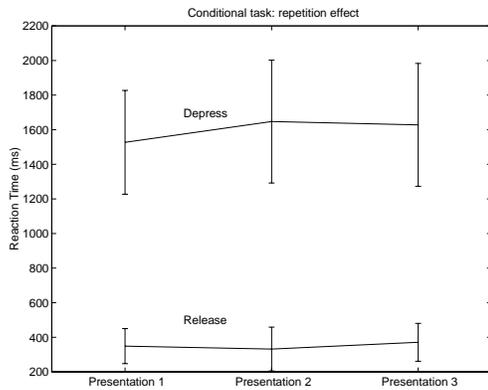
(B)



(C)



(D)



(E)

Figure 7.6: Reaction times for both the release of the switch and the depression of the next switch: (A) comparison between the three movement tasks; (B) comparison of precision and power trials within the conditional task; and repetition effects for power grasp (C), precision pinch (D), and the conditional task (E). The error bars indicate 1 STD.

Anatomic Location	Talairach Coordinates (mm)			Volume (mm <sup>3</sup> )
	X	Y	Z	
L Dorsal Premotor Cortex (postcentral gyrus)	-25.35	-10.52	51.34	1340
L Inferior Parietal Cortex (supramarginal gyrus; anterior bank of the intra-parietal sulcus)	-32.98	-56.86	33.00	520
R Area 18 (occipital lobe, lingual gyrus)	7.66	-98.67	-6.44	476
L Area 18/19 (superior occipital lobe)	-24.91	-72.72	25.34	192
R Area 18 (occipital lobe, lingual gyrus)	8.37	-69.81	-11.68	74
R Cerebellar Cortex*	7.5	-66.0	-22.5	433
R Dorsal Premotor Cortex (frontal medial gyrus)*	15.21	-2.27	49.81	165

Table 7.1: Significant regions of activity ( $p < 0.001$ ): conditional - non-conditional (combined precision and power). \* = only significant ( $p < 0.001$ ) in 5 subjects (certain regions were missed on the scanner with one subject).

Anatomic Location	Talairach Coordinates (mm)			Volume (mm <sup>3</sup> )
	X	Y	Z	
L Area 18/19 (superior occipital lobe)	-26.74	-74.39	25.74	176
L (somewhat bilateral) SMA Proper (frontal medial gyrus)	-6.74	-37.22	66.56	493
R Dorsal Premotor Cortex (frontal medial gyrus)	18.28	-12.01	65.10	179
L Cerebellar Vermis*	-4.69	-60.09	-20.12	57
L Inferior Parietal Cortex (supramarginal gyrus; anterior bank of the intra-parietal sulcus)*	-50.21	-47.27	37.34	196
L Area 18 (occipital lobe, lingual gyrus)*	-11.24	-80.84	0.11	216
L Area 7 (superior parietal lobe)*	-28.52	-56.87	54.76	209

Table 7.2: Significant regions of activity ( $p < 0.005$ ) : precision pinch - power grasp. \* = only significant ( $p < 0.005$ ) in 5 subjects (certain regions were missed on the scanner with one subject).

## **Comparison of PET and synthetic PET Results**

### **Precision versus Side Opposition/Power Grasp**

In the comparison of precision pinch and side opposition, the model predicts an increase of synaptic activity in both F5 and AIP in the precision pinch case. Although we see an increase in activity in the inferior parietal area in the human experiment, we fail to see any such change in the ventral premotor cortex (specifically F5). Two explanations are possible for this negative result. First, it is possible that for stereotypical grasps, that F5 does not play a significant role. Rather, it is the cerebellum that is that is responsible for the execution of the grasp.

The second possibility assumes that F5 is involved in the grasp, but the effect is masked by force-related activity in the region. In the human experiment, performance of the power grasp required a reasonable level of force to be applied to the block before the LED would indicate to the subject that a grasp had been detected. In monkey, force-related activity has been observed in F5 (Heppreymond, Husler, Maier, & Qi, 1994). The implication is that even though there are fewer neurons involved in encoding the power grasp, they are achieving a higher level of activity because of the force requirements of the task. Thus, the rCBF measures could still be relatively similar in both conditions.

### **Conditional vs Non-Conditional Task**

The model predicts that in the conditional task, we would expect to see a much higher level of activation in F2 (dorsal premotor cortex), some activation of F5, and a slight activation of AIP. The human experiment confirmed the F2 result, but failed to do so in F5 and AIP. In fact, the results were essentially opposite: in human, we see an activation of the inferior parietal cortex (AIP), but no significant activation of ventral premotor cortex.

Consider the functional connectivity of these regions in the model (Figure 7.2). In practice, the strength of projection from F2 to F5 is essentially a free parameter. In other words, there is a wide range of values over which the model will correctly perform the conditional and non-conditional tasks. The implication is that by tuning this parameter, we can control this projection's contribution to the synaptic activity measure in F5. However, it is always the case that the difference in

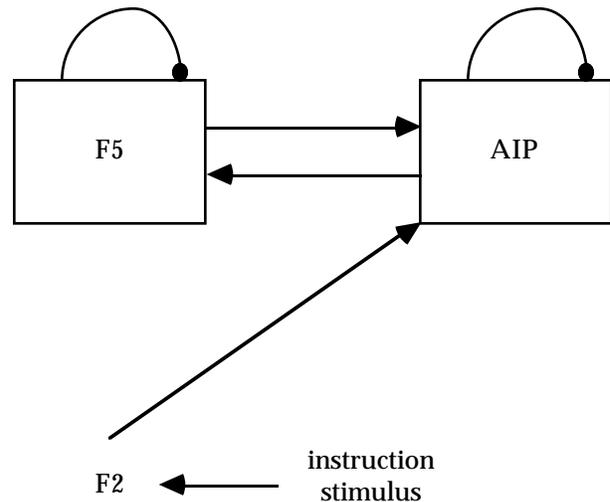


Figure 7.7: Updated functional model. The information from F2 flows into the circuit through a projection into AIP.

AIP synaptic activity from the non-conditional to the conditional task will be less than the difference observed in F5. Why is this the case? By increasing the projection strength from F2 to F5, we observe both an increase in F5 activity and synaptic activity. The increase in F5 activity, however, is attenuated by the recurrent inhibitory projection through the BG. Thus the signal that is then passed on to AIP from F5 does not reflect the full magnitude of the signal received from F2.

The conclusion is that although we can adjust the free parameter to match one or the other observations in the human experiment (of either F5 or AIP changes), the model cannot reflect both at the same time. One possibility for repairing this problem in the model is to reroute the F2 information so that it enters the grasp decision circuitry through AIP, rather than F5 (Figure 7.7). As a result, we will see an increase in activity in AIP due to F2 activation, and only an attenuated signal will be passed on to F5 (and thus, only a small increase in synaptic activity will be observed).

## Summary

Neurological experiments performed in monkey and human fundamentally yield different types of information about the computational structure and function of the brain. In human, PET and fMRI techniques allow us to achieve a global view of the systems involved in performing a task, but at the expense of a very coarse spatial and temporal resolution. In monkey, on the other hand, recording of single unit (or small populations of units) presents significant challenges as we attempt to understand what a region is computing, let alone an entire circuit of regions. It is thus important to develop techniques that allow experimental results at both levels to be brought together as we attempt to understand the different systems.

In this chapter, we have explored the use of a technique referred to as synthetic PET imaging for just this purpose. The low-level details of the grasping model presented in earlier chapters has been derived primarily from neurophysiological results obtained in monkey. Using the synthetic PET approach, we extract from the model measures of regional synaptic activity as the model performs a variety of tasks. These measures of synaptic activity are then compared to rCBF (regional cerebral blood flow) observed during human PET experiments as the subjects perform tasks similar to those performed by the model. In some cases, the human results provide confirmation of the model behavior. In other cases, where there is a mismatch between model prediction and human results, it is possible (as we have shown) to use these negative results to further refine and constrain the model.

Figure 7.8: Significant activity ( $p < 0.005$ ) for conditional - non-conditional tasks

Insert cond - movement here

Insert precision - power here

Figure 7.9: Significant activity ( $p < 0.01$ ) for precision pinch - power grasp.

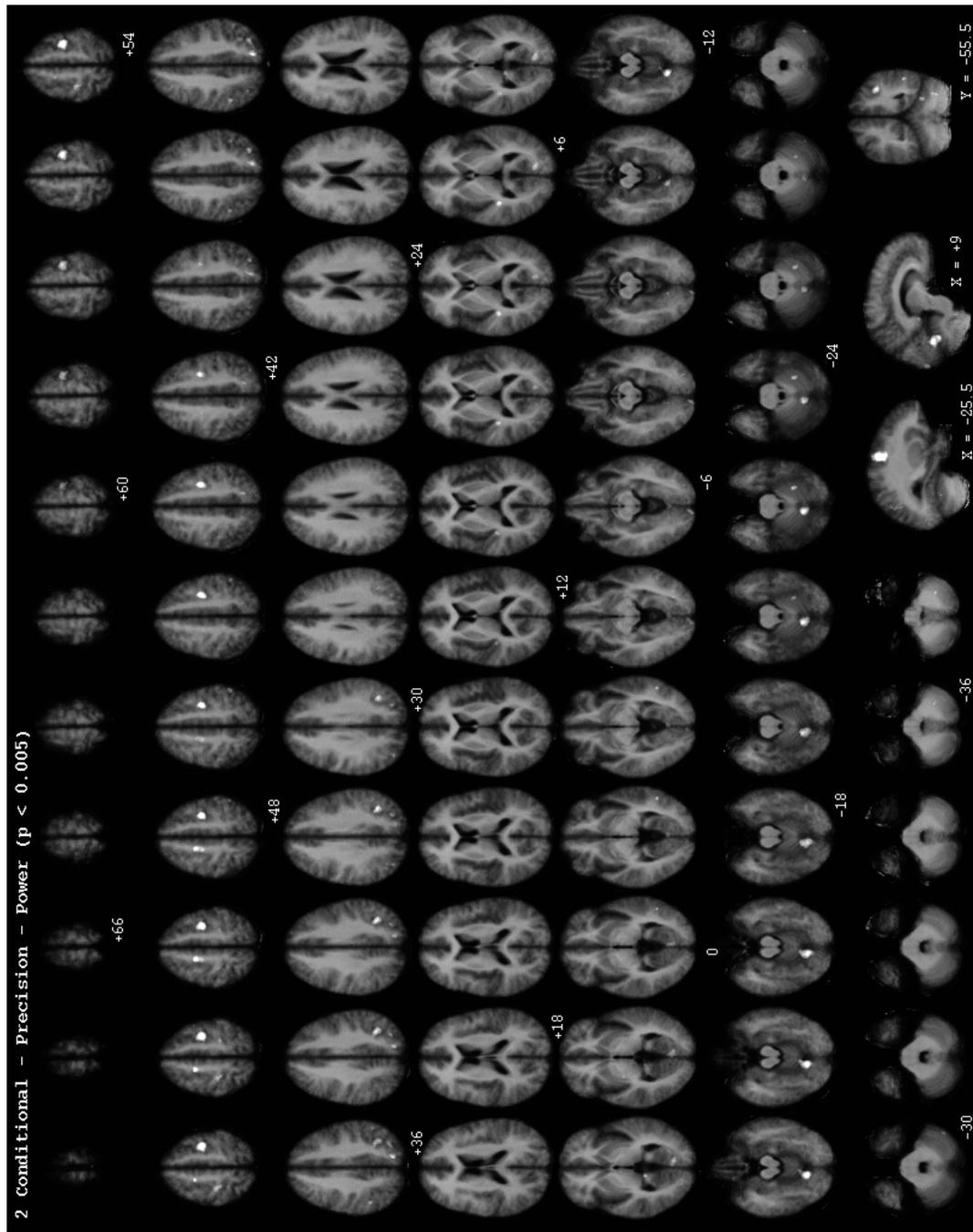


Figure 7.10: Significant activity ( $p < 0.005$ ) for conditional - non-conditional tasks.

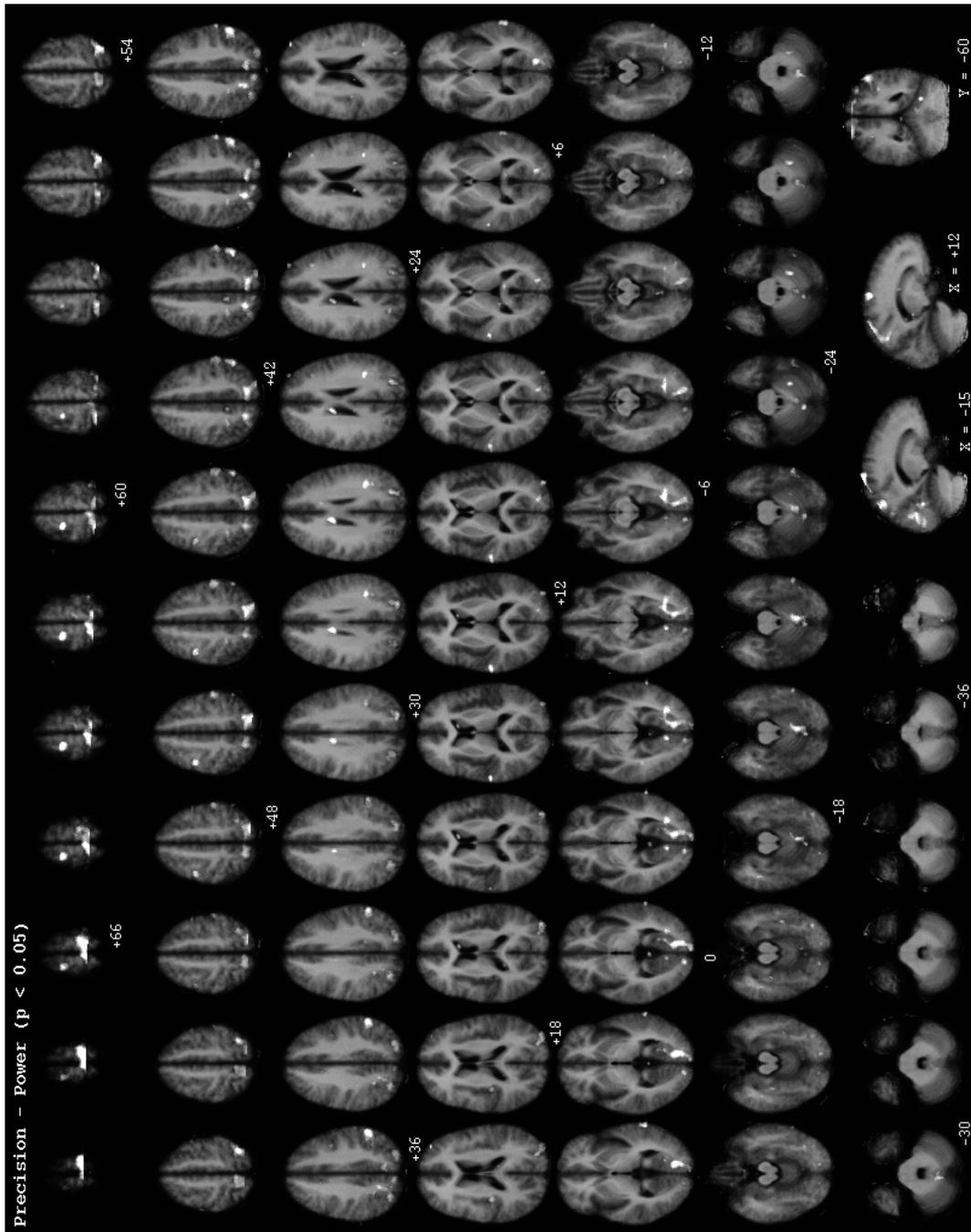


Figure 7.11: Significant activity ( $p < 0.01$ ) for precision pinch - power grasp.

## Chapter 8: Grasp Configuration Learning

---

In this chapter, we present a model that learns the transformation from visual and task parameters to an appropriate grasp. The transformation is acquired through a process of trial and error: the simulated robot makes successive attempts at grasping objects and observes the outcome of the attempts. The degree of success of each attempt is scored by a teacher, which is then used by the *reinforcement learning algorithm* to adjust the visual/motor mapping. After repetitively experiencing several different situations, the network learns to perform the appropriate transformations. In addition, individual neurons in the network either align themselves with one of the two grasps (precision pinch or power grasp), or come to represent the aperture of the grasp that is executed.

---

### Introduction

The grasping model presented in previous chapters of this thesis has relied on a heuristically wired network for the transformation from visual and task information into a grasp description. The wiring algorithm requires that the gross behavior of each F5 and AIP module be specified a priori by the network "programmer." In this chapter, we demonstrate that such a transformation may be automatically acquired by a network through the use of a learning algorithm. Rather than specifying the intermediate representation of the mapping (behavior of individual neurons), the *teacher* provides only information about the input/output pairs of the desired transformation. The learning algorithm makes use of the teaching

information in adjusting the behavior of individual neurons such that the desired transformation is ultimately achieved.

In the experiment presented below, it is assumed that the teacher actually does not know the exact form of the transformation. In other words, the teacher cannot say what the exact motor commands are that should be sent to the fingers of our simulated robot given any particular situation. However, the teacher is capable of evaluating the performance of a grasp that is executed by the robot. The high-level performance measure of *how well* the robot did, referred to as the *reinforcement* feedback signal, is given in the form of a scalar score. In the case of the grasping task, this score measures two elements: the success and the efficiency of the executed grasp (described below in more detail). Based upon this feedback information, the reinforcement learning algorithm (Barto, Sutton, & Anderson, 1983; Sutton, 1988) adjusts the connection strengths of the artificial neural network so as to maximize the success and efficiency of the grasps that are executed.

## Neural Dynamics

The neural architecture (as shown in Figure 8.1), which is significantly simplified from that presented in earlier chapters, has been adapted from our work on primate visual-motor conditional learning (Fagg & Arbib, 1992) and work on reinforcement learning for reactive control of a mobile vehicle (Fagg, Lotspeich, Hoff, & Bekey, 1994). The visual and task information that is input into the model is represented as an activity pattern across a set of neurons ( $V$ ). A total of 11 neurons are used in this layer: three represent object type (cylinder, cube, and plate), three represent length (short, medium, long; note that this is a discrete representation), three for diameter (narrow, medium, wide), and two for task requirements (manipulability and stability). In the monkey, this visual information is provided by a sub-region of the posterior parietal cortex (specifically PIP); task information can be derived from a number of different regions including prefrontal cortex (including area 46), and preSMA (F6).

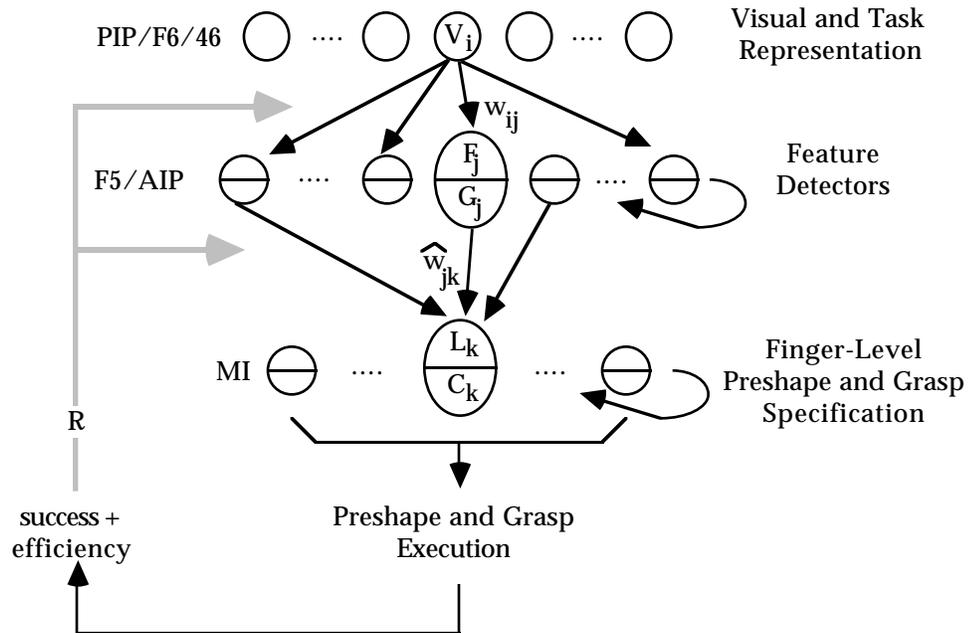


Figure 8.1: Schematic view of the architecture for the grasp configuration learning model. Visual parameters (from PIP) and task requirements (from F6 and area 46) are combined at the feature detector layer (representing the AIP/F5 complex). The activated feature detectors in turn select a hand configuration by specifying how individual fingers will behave during the preshape and grasp (primary motor cortex; MI). This program is executed and then evaluated by a teacher. The evaluation ( $R$ ) is used to update the inter-layer connection strengths.

This information ( $V$ ) is projected through a set of synapses ( $W$ ) to a feature detector layer ( $F/G$ ), in which each neuron represents some higher-order feature of the original description. For example, a feature detector representing *cylinder and manipulability* would receive connections from the corresponding input neurons. The activity level of the feature detector unit ( $F$ ) is computed according to the following equation:

$$F_j = \sum_i (V_i * w_{ij}) + Noise_j \quad (8.1)$$

where:

$F_j$  is the activity level of feature detector unit  $j$ .

$V_i$  is the activity level of input unit  $i$ .

$w_{ij}$  is the strength of connection from input unit  $i$  to feature detector  $j$ .

$Noise_j$  is a random signal that is injected into feature detector j.

These feature detector units then interact through a local competition mechanism to contrast-enhance the incoming activity pattern. This is implemented as a one-pass local-maximum operation: a neuron produces a non-zero output if and only if it is the most active neuron within a small neighborhood. The output of the feature detector units ( $G$ ) is computed as follows:

$$G_j = \begin{cases} F_j & \text{if } F_j = \underset{j-N \leq l \leq j+N}{\text{Max}} \{F_l\} \\ 0 & \text{otherwise} \end{cases} \quad (8.2)$$

where:

$G_j$  is the output of unit j.

$N$  defines the size of the region of competition. Note that we have implicitly assumed that the neurons are arranged in a circular array.

The active feature detectors then vote for the configuration to be imposed upon the hand by passing activation to the output units ( $L$ ):

$$L_k = \sum_j (G_j * \hat{w}_{jk}) + Noise_k \quad (8.3)$$

where:

$L_k$  is the activity level of output unit k.

$\hat{w}_{jk}$  is the strength of connection from feature detector unit j to output unit k.

$Noise_k$  is a random signal that is injected into output unit k.

The configuration specifies which fingers will be actively participating in the grasp and how the fingers should be positioned during the actual execution of the preshape and grasp. As this model is designed to drive the Belgrade/USC Hand (Bekey, Tomovic, & Zeljkovic, 1990), the index/middle and the ring/little finger pairs are considered single entities to be controlled. The output ( $C$ ) consists of seven separate sub-vectors, each specifying a different detail of the

grasp configuration. Three of the sub-vectors (each consisting of two units) specify the participation of the thumb, I/M fingers, and R/L fingers in the grasp, respectively. One sub-vector (also consisting of two units) determines whether the thumb will be abducted or not. The three remaining sub-vectors (each consisting of three units) determine the degree of flexion (small, medium, and large) during preshape of the thumb and of the I/M finger and R/L finger pairs, respectively.

For each sub-vector, a winner-take-all circuit computes the single most active unit of the set:

$$C_k = \begin{cases} 1 & \text{if } L_k = \underset{m \in S(k)}{\text{Max}}\{L_m\} \\ 0 & \text{otherwise} \end{cases} \quad (8.4)$$

where:

$S(k)$  is the set of units that are in the same sub-vector as unit  $k$ .

$C_k$  indicates whether configuration bit  $k$  is a winner.

It is this resulting pattern of activity ( $C$ ) that is used by the non-neural execution system to generate the motor commands that result in the preshape and enclose movements of the hand.

## Learning Dynamics

After execution of the specified grasp, a teacher evaluates the performance of the system. There are two elements to this evaluation: success and efficiency. *Success* tells us whether or not the grasping movement was able to pick up the object. If the robot is unable to accomplish this, then a reinforcement signal of  $R=-0.1$  is given by the teacher. If the grasp is successful, then a positive reinforcement signal ( $R=1$ ) is given, but discounted if the grasp is *inefficient*. A grasp is considered inefficient if the fingers preshape to a larger extension than is necessary for the presented object (this is related to Hoff's notion of *hand comfort* (Hoff & Arbib, 1993)). The

discount factor is set such that if the system produces a preshape of maximum possible aperture for a narrow cylinder, a reinforcement of 0.4 is given (assuming success).

The reinforcement signal is used by the learning algorithm to update the connection strengths in the projections from the visual/task representation to the feature detectors ( $W$ ), and from the feature detectors to the action units ( $\hat{W}$ ), with the goal of ultimately maximizing the level of reward that is received. The adjustments to the connection strengths are done using a Hebbian/Anti-Hebbian learning algorithm, as follows:

The system is presented with a certain input, for which a grasp plan is computed by the network and then executed. Suppose that the teacher gives a positive reinforcement signal for the system's performance on the trial. In this case, we would like to ensure that the next time the system is presented with the same input, the same plan is output. This is accomplished by:

- 1) Making sure that the same set of feature detectors is activated by increasing the connection strength from active input units ( $V$ ) to active feature detectors ( $G$ ) (thereby increasing their response level the next time).
- 2) Increasing the active feature detectors' support of the selected grasp plan by increasing the connection strengths from the active feature detectors ( $G$ ) to the active output units ( $C$ ).

On the other hand, suppose negative reinforcement is received. This could be due to the fact that either the wrong set of feature detectors was selected, or just that the specified grasp plan was incorrect. Since we do not know which is the case, both are assumed:

- 1) The connection strengths from the active input units ( $V$ ) to the active feature detectors ( $G$ ) are reduced, thereby giving other feature detectors an opportunity to become active next time (implementing a search for the correct set of feature detectors).
- 2) The support of the active feature detectors for the specified configuration is also reduced. This allows other configurations to be tried.

These rules are captured in the following connection strength update equations:

$$\begin{aligned}\Delta w_{ij} &= \alpha R V_i G_j w_{ij} \\ \Delta \hat{w}_{jk} &= \alpha R G_j C_k \hat{w}_{jk}\end{aligned}\tag{8.5}$$

where:

$\Delta w_{ij}$  and  $\Delta \hat{w}_{jk}$  are the changes to the connection strengths.

$\alpha$  is a learning rate parameter.

$R$  is the teacher's reinforcement signal.

In a final step, the weight updates are integrated into the weight matrices  $W$  and  $\hat{W}$  according to:

$$w_{ij} \leftarrow \frac{w_{ij} + \Delta w_{ij}}{\sum_i (w_{ij} + \Delta w_{ij})}\tag{8.6}$$

$$\hat{w}_{jk} \leftarrow \frac{\hat{w}_{jk} + \Delta \hat{w}_{jk}}{\sum_k (\hat{w}_{jk} + \Delta \hat{w}_{jk})}\tag{8.7}$$

The combination weight update and normalization not only maintains individual weights within a limited range, but also implements a form of competition between the different connections. Relative to a single feature detector unit, when a set of input connections are incremented (in response to positive reinforcement), the remaining connections are subsequently decremented (equation 8.6). This encourages the development of feature detector units that are responsive to a small number of situations. On the other hand, when a set of weights are decremented (negative reinforcement), the remaining connection strengths are increased, making the feature detector responsive to a larger number of situations (forcing the system to continue searching for the appropriate combination of feature detector units).

The set of outputs leading from an individual feature detector unit also compete with one-another (equation 8.7). This can be interpreted as an individual feature detector unit having a limited number of votes that can be distributed across all possible configurations.

## Reinforcement Prediction

As presented, the reinforcement learning algorithm has a tendency to converge on the first successful solution that it discovers for each situation. This is the case because the reception of positive reinforcement always results in an increase of some set of connection strengths, even after the network reliably produces the successful solution. The result is that the activity level of the feature detector and output units is not significantly affected by the injected noise (equations 8.1 and 8.3), and hence the system does not continue the search process, even in the vicinity of the successful solution. Here, we outline a modification to the learning algorithm that attempts to focus learning to situations in which there is an increase or decrease in performance, rather than being sensitive to the absolute performance of the network (more details of this formulation can be found in (Fagg, Lotspeich, Hoff, & Bekey, 1996)).

First we assume that for any situation that is presented to the network (situation being object and task), we have an estimate of the network's average performance, referred to as  $\hat{R}$ . We can then define *internal reinforcement* as the deviation of the network's performance (on a single trial) from its average performance:

$$R' = R - \hat{R} \tag{8.8}$$

and replace the weight update equations (8.5) with:

$$\begin{aligned} \Delta w_{ij} &= \alpha R' V_i G_j w_{ij} \\ \Delta \hat{w}_{jk} &= \alpha R' G_j C_k \hat{w}_{jk} \end{aligned} \tag{8.9}$$

The effect of this algorithmic change is that when the network performs better (or worse) than expected, a positive (negative) internal reinforcement signal is generated. In addition, when the network performs as expected, no learning occurs.

The average performance of the network is estimated as a linear function of the input and feature detector states:

$$\hat{R}(V, G) = M^T \bullet \begin{bmatrix} V \\ G \end{bmatrix} \quad (8.10)$$

where:

$M$  is a column vector of parameters.

After observation of the outcome of a grasp, the estimate of average performance is updated according to the incremental LMS algorithm (Widrow, 1962):

$$M \leftarrow M + \hat{\alpha} \Delta M \quad (8.11)$$

$$\begin{aligned} \Delta M &= (R - \hat{R}(V, G)) \nabla_M \hat{R}(V, G) \\ &= (R - \hat{R}(V, G)) \begin{bmatrix} V \\ G \end{bmatrix} \end{aligned} \quad (8.12)$$

where:

$\hat{\alpha}$  is the learning rate.

The effect of this reinforcement prediction algorithm is that when the system discovers a new (and better) solution to a certain situation, the network will receive positive internal reinforcement for several trials. This positive reinforcement will cause the network to favor the new solution over the old. However, the reinforcement predictor will respond (with a lag) by increasing the estimate of average performance. This means that after the first few trials, the new solution will result in a near-zero internal reinforcement signal, and therefore learning is stopped unless an even better solution is discovered. This allows the search process to continue (due to the injected noise) around the vicinity of the currently favored solution.

Note that this process of first estimating and then climbing the reinforcement gradient is a related to Sutton's temporal difference algorithm (Sutton, 1988), and its application to learning of optimal controllers (Barto & Bradtke, 1991).

## **Experimental Results**

During the training process, the system was presented with a sequence of situations selected from six possibilities. The presented object was a cylinder of one of three widths. In addition, the task requirement could be either manipulability or stability. The experiments were performed in simulation. Evaluation of the grasps was accomplished by checking that the fingers were opened wide enough to clear the object and that the task requirements were suitably satisfied. If manipulability was a requirement, the system was expected to generate a pad opposition grasp configuration (with the thumb opposing the finger tips). If stability was required, then the system had to have produced a palm opposition in order to be considered successful. The protocol randomly presented the system with one of the six possible situations. However, if the system failed to grasp the object, then the same situation was presented until a successful grasp was obtained.

For the results presented below, the network required 2500 trials before its performance peaked. After training, all grasps were successful but not all were completely efficient. Of the 80 feature detector units, only 18 achieved very significant response levels. Figure 8.2 shows the post-learning response curves of five of these units as a function of the situation that is presented. Units A-D are all selective for a pad opposition, although C and D are also selective to some degree for the width of the object. Unit E is selective for palm opposition. These grasp specific neurons, similar to those reported by Rizzolatti (Rizzolatti, et al., 1988) in area F5, resulted from the learning process and the system's interaction with its environment, even though the architecture did not inherently contain these concepts.

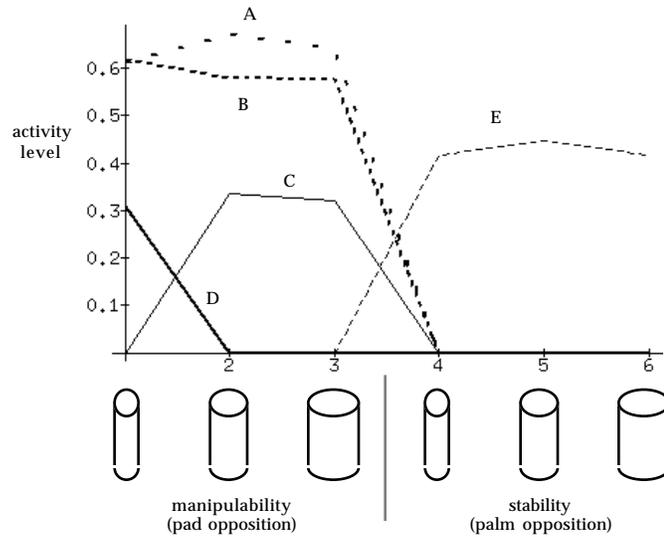


Figure 8.2: Responses of several feature detector units (A-E) given the six different situations. All units are selective for either pad or palm opposition.

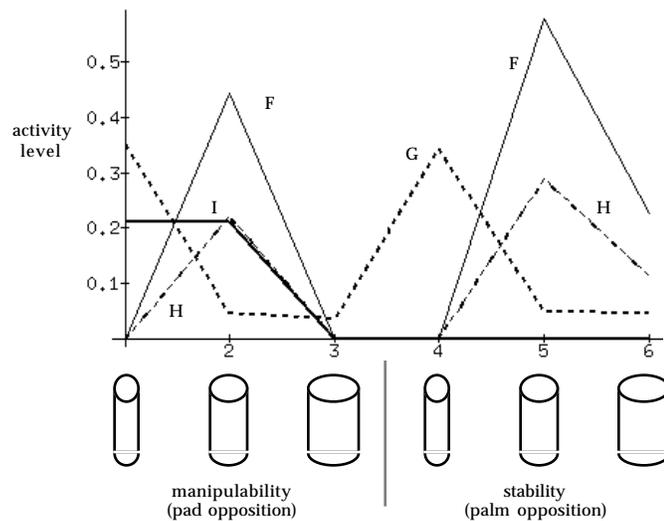


Figure 8.3: Responses of four additional feature detector units (F-I) from the same experiment. Units F, G, and H are selective for the width of the cylinder. Unit I is selective for small & medium cylinders for pad opposition.

Finally, it is important to note that the firing pattern of some cells did not follow a simple symbolic rule such as *fire during execution of the pad opposition*. Rather, these cells responded

to a combination of several different input dimensions. Despite this apparent difficulty of placing symbolic labels on these cells, their interaction with other cells in the feature detector layer was able to yield a correct and consistent grasp plan.

In addition to grasp selective cells, other types of units were also seen in this experiment. Figure 8.3 demonstrates several such cells. Cells F and H are selective for objects of medium width regardless of the grasp type that is made. Cell G is selective for small objects, but does show some activity for the other situations. In all, nine units demonstrated purely opposition-related activity, where only four showed object size selectivity. The remaining five either showed selectivity to exactly one situation or to an arbitrary combination of multiple situations. Although size-specific responses have not been observed in F5 (due to experimental design), Sakata has observed such responses in AIP (Sakata, 1994).

## Summary

In this chapter, we have demonstrated that a transformation from visual and task parameters into a motor plan can be acquired automatically. The network programmer (or teacher) need not specify the computational processes behind the transformation (i.e., the behavior of individual neurons), but rather may specify the expected results of the computation. However, unlike algorithms that learn to mimic human grasp plans (Iberall, 1987; Iberall, 1988; Uno, Fukumura, Suzuki, & Kawato, 1993), the planner actually learns by observing the results of executing its own plans. Note that there is still a teacher involved in the process - instead of telling the robot exactly *what* to do, the teacher only tells the robot *how well* it did in performing its task. In this way, it is possible to construct plans that are best oriented towards the actuation and sensing capabilities of the robot. This is a property that is not necessarily achievable when only mimicking human performance, since underlying the human behavior may be a program that relies on sensory feedback that is not available to the

robot. When this happens, the robot is unable to distinguish situations in which different motor decisions must be made.

It is important to note, however, that the network architecture presented in this chapter is significantly simplified from that used in the main model (Chapters 4, 5, and 6). First of all, the mapping from visual and task parameters into motor outputs is implemented as a one-pass neural algorithm; any recurrent interactions are simulated through the use of winner-take-all mechanisms. Second, the output of the network is a static representation of the grasp. The actual execution of the preshape and grasp are assumed to be implemented by a separate, non-neural module. Although these simplifications are justified from a computational efficiency argument, it is also not immediately clear how the temporal aspects of the neural firing patterns (in the original model) can be shaped by an associative algorithm. This is the subject of ongoing research.

## Chapter 9: Conclusions and Future Work

The planning and execution of grasps requires the integration of multiple sources of information. The visual system provides the object's identity, as well as an estimate of its shape, size, and localization in space. Although this information can be used to specify the fine details of a grasp plan, a single object actually affords many different grasps. This choice of grasp is dependent upon the behavioral context, which can include instructions from an experimenter, working memory inputs, or a specification of the task that is to be performed with the object once it is grasped.

This thesis has presented a neural-level model of the cortical processes that are involved in the generation and execution of grasp plans. The model addresses the issues of neural-level encoding of object, grasp, and phasic information, and shows how populations of cells can capture the critical information that is required to perform the necessary computations. In addition to reproducing single-cell behavior that has been observed experimentally in F5 and AIP in monkey, the model makes several more general hypotheses about the computations that are being performed by these neural regions:

1. The neural activity patterns observed in AIP reflect the extraction of action-relevant features (or affordances) from the visual representation of the attended object.
2. During execution of the grasp, AIP maintains an *active memory* of the affordance that corresponds to the grasp that is being executed. This active memory is maintained in part by recurrent projections from F5.

3. During grasp plan formation, F5 selects from the possible grasps, as represented in working memory or in AIP (corresponding to the affordances). This selection process takes into account task-related constraints (received from either area 46 or F6).
4. F5 is responsible for the high-level execution and monitoring of the grasp.
5. Parameters of an object or grasp are reflected in sub-populations of F5 and AIP cells (this hypothesis is explored in Prediction #3 of Chapter 6).

These hypotheses lead to a number of predictions about the single cell- and population-level behaviors under several novel tasks.

1. In the case where two distinct objects are presented and subsequently grasped in an identical manner, hypotheses 1 and 2 imply that motor-related cells in AIP will tend not to reflect differences in the two situations. However, visual-related cells (which derive many of their inputs from PIP and IT) can reflect these differences, and hence appear to demonstrate an object-specific code (Prediction #1 of Chapter 6).
2. Hypothesis 1 implies that different parameters of an object will be inequitably reflected in AIP. Specifically, we expect to see only those parameters that are relevant for programming of the grasp (Prediction #5).
3. In the case where an object is to be grasped in one of two ways, but information about which of the two grasps that must be selected (e.g. via an instruction stimulus) is withheld for some delay period, hypothesis 3 implies that we would expect both possible grasps to initially be prepared, and hence we predict activity patterns in F5 that correspond to both possibilities. When the IS is given, the population of cells not corresponding to the selected grasp are inhibited (Prediction #2). Hypothesis 2 implies that AIP will behave in a similar manner.
4. Because F5 monitors the ongoing grasp (hypothesis 4), when there is a miscalibration between the visually-estimated and actual size of an object, we predict a shift in the executing program at the time that the more precise tactile information becomes

available (Prediction #4). Hypothesis 2 implies that this shift in program will also be reflected in AIP.

As discussed at the end of Chapter 6, some observations made of AIP cell responses are consistent with Predictions #1 and #2.

### **PET Imaging**

In addition to the predictions made by the model at the neuronal level, we have shown how it is possible to extract other measures of model behavior. Specifically, through the application of the synthetic PET imaging technique (Arbib, et al., 1995), in which the regional synaptic activity is measured over the course of performing a task, we are able to make predictions about the expected results from human PET experiments.

Two sets of predictions made by the model were tested in a human PET experiment. In the first experiment, we looked at the relative synaptic activity during execution of a conditional task (in which an IS informed the subject as to which grasp to use) as compared to execution of the two grasps without using an IS (see Chapter 7 for more details). The most significant result from this comparison was that an area of the contralateral postcentral gyrus (corresponding to dorsal premotor cortex, or F2) was very active. This result was predicted by the model. In addition, the model predicted a reasonable activity increase in F5, and a slight increase in AIP (increase from non-conditional to the conditional task). The human PET result did not confirm these predictions: no change in activity was seen in F5, and only a small amount of activity was observed in AIP.

The second of the comparisons looked at synaptic activity differences between the precision and power grasps (in the model, we compared precision pinch and side opposition). The model predicted an increase in both F5 and AIP (from side opposition to precision pinch). The human PET result showed an increase in AIP activity, but no change in F5.

We concluded the discussion on PET and synthetic PET imaging by showing how the PET results can be used to further constrain the model. Because PET imaging inherently provides a global measure of system behavior, the constraints that it provides are also at a global level. These global constraints allow us to make conclusions about the existence or non-existence of projections from one layer to another, or even about the relative strengths of inhibitory and excitatory projections between regions.

### **Grasp Configuration Learning**

The primary model in this thesis relied on a network that was, for the most part, wired by hand. However, in Chapter 8, we demonstrate how the problem of transforming visual and task parameters into an appropriate grasp description can be learned without prewiring a network. The transformation was learned in a trial-and-error fashion: for a given situation, the network produced a grasp plan and then observed the outcome of the plan. Using a reinforcement-learning algorithm, the success and efficiency of the executed grasp was used to update the transformation. Most interestingly, after learning was complete, some neurons within the network became sensitive to the type of grasp that was performed (precision pinch or power grasp), as we have seen in the monkey F5 (Rizzolatti, et al., 1988). In addition, other cells reflected the aperture width of the grasp. In addition, the behavior of some cells could not be described using a simple symbolic rule, but demonstrated rather complex behaviors (as we see in the biology).

## **Related and Future Work**

### **Control of Wrist Orientation**

Some cells observed in AIP are responsive to the orientation of the target object (Taira, et al., 1990). Although this information might be relevant to the set of grasps that are afforded by the object, it is also likely that this information is used in programming of the wrist position.

Note that the necessary wrist position is not simply determined by the orientation of the target object in space, but varies as a function of the affordance that is ultimately selected. Including object orientation in AIP is consistent with the notion that F5 is involved in controlling the position of the wrist (Rizzolatti, 1995).

### **Planning of Arm Movements**

A related problem is that of positioning of the arm. Although the location of the arm is tied to the position of the object in space (as represented in VIP), depending upon the grasp that is selected, the wrist might be positioned differently (consider grasping a cylinder with either a precision pinch or a power grasp). Thus, there must be some mechanism by which the selection of an affordance/grasp pair is integrated into the goal position of the arm. Likewise, this mechanism must also permit arm position constraints to be reflected in the decision of the grasp. Possible pathways for these interactions include AIP's connections with VIP and F4.

The transformation of object position from a body-centered coordinate system into a program for arm movement has been studied extensively (Kalaska & Crammond, 1992; Stein, 1992; Wolpert, Ghahramani, & Jordan, 1995). In particular, (Burnod, Grandguillaume, Otto, Ferraina, Johnson, & Caminiti, 1992) have proposed a model based on monkey neurophysiological studies of visual and motor encoding of movements. The model learns to combine a visual representation of target position and a kinesthetic representation of the current arm state in order to compute arm movement commands. The movements are represented using a population coding scheme.

### **Coordination of Reach and Grasp**

During execution of the reach and grasp, it has been observed in human that there is a close temporal coordination between the reach and grasp processes (Paulignan, et al., 1991a; Paulignan, MacKenzie, Marteniuk, & Jeannerod, 1991b). (Hoff & Arbib, 1993; Hoff, 1992) have proposed a schema-level model that describes how the reach, preshape, and enclose modules

exchange information in order to achieve this coordination. During execution, each module estimates the amount of time that is required in order to complete its goal. This timing information is compared by a coordinating module, which in turn modulates the rate of execution for each module. In cases where the reach process requires much more time than the preshape/enclose processes (as in the case where the target position is perturbed), their state is reset to the beginning of the preshape phase. Although this coordination mechanism most likely involves cerebellar processes, the cerebellum has access to the hand primarily through F5 and MI. Within the model presented in this thesis, modulation of the 'rate' of execution can be achieved through a global facilitation or depression of F5 or MI activity. The resetting of hand control state could be achieved through the shunting of flexion phase cells (in F5) and activation of the extension phase cells.

### **Learning of Motor Programs**

In Chapter 8, we presented a second model in which the mapping of visual and task parameters to a selected grasp was learned through a trial-and-error procedure. One of the significant simplifications of the model was that the mapping was static -- i.e., the grasp plan that was output was independent of the module that executed the plan. Although it is not immediately clear how the set of grasp programs in F5 can be learned, once the set of programs exist, it is possible for the mapping between visual parameters (as represented in PIP and IT) to be associated with the affordance representation within AIP. This learning scheme relies on two features in the model. First, that F5 will automatically adjust its program to the actual situation based on tactile and kinesthetic inputs received via SII. Second, changes in the F5 program are reported back to AIP through the recurrent cortico-cortical projections. Thus, once the grasp is established, AIP contains a representation of the grasp that was *actually executed*, which can then be associated with the visual representation of the object.

## Encoding of Grasp Programs in F5

As we observed in Chapter 6, individual F5 units in the model were often active in more than one phase of the grasping program to which they belong. In addition, a number of cells were observed as participating in more than one grasp. Despite the temporal and programmatic overlap, as implemented, the model does not make functional use of this feature. By *functional use* we mean that the output motor command generated by a single F5 unit (to MI) is specific to one grasp type/phase pair<sup>1</sup>. A more satisfying grasp coding scheme, and one that is better related to Rizzolatti's notion of a *language of motor acts* (Rizzolatti, et al., 1988), is one which allows individual F5 units to be responsible for small pieces of a grasp, which can then be assembled with other pieces to form a complete grasp. For example, an F5 unit that participates in both a lateral and precision pinch might generate motor outputs that are common to both grasps, such as the control of thumb flexion. Although it is possible to approach this issue through modification of the network wiring heuristics (see Appendix A), it is probably best addressed in the context of learning the motor programs.

## Perception of Grasp

Rizzolatti, Fadiga, Gallese, & Fogassi (1996) have recently reported that many neurons in the convexity of area F5 are not only active as the monkey executes a specific grasp, but also when the monkey observes the same motor act being executed by the experimenter or another monkey. It has thus been hypothesized that these neurons are involved in the perception of motor acts. These observations present a number of interesting challenges for our modeling work. First of all, how do these neural properties arise? One possibility is that the monkey first learns to associate the visual patterns created by movements of his own hand (possibly supplied by temporal cortex) with the descending motor command produced by the F5 neurons

---

<sup>1</sup> However this overlap does play a role in the cooperative interaction of groups of F5 units.

that are responsible for the generation of the movement (which are located in the bank of F5). These associations are then generalized to hand movements external to the monkey. Second, how is this information useful to the monkey? One important reason is that through understanding the motor acts that are being performed around him, it is possible for the monkey to learn when and how to perform the same motor acts. This ability to learn by example of others is also important for our work in the design of more intelligent robot controllers (Fagg, 1993; Kang & Ikeuchi, 1995).

## Appendix A: Model Implementation Details

Here, we present detailed descriptions of the network construction. The first step in constructing a given network involves specifying the type and high-level behavior of the neurons in the different layers, as well as specifying the set of grasps and objects that are to be encoded in the network. We first present the information that must be specified, and then show the number and distribution of neurons that are used in the experiments.

The second step in constructing a network makes use of the behavioral descriptions to heuristically generate the network wiring diagram.

Further details of this wiring process (including code), as well as access to the model itself can be found at <http://www.usc.edu/dept/robotics/personal/af0a/thesis/>

### Network Structure Description

The behavioral descriptions of individual AIP and F5 units have already been presented in Chapter 5 (Tables 5.1 and 5.2, respectively). In addition we have discussed the representation of motor commands in MI (also Chapter 5).

The grasp description defines the configuration of the hand during and after execution of the grasp (Table A.1); this information is used to establish connections from F5 to MI to implement the movements during grasp execution. The *jacobian* and *expected tactile input* vector is used in selecting the MI cells that will be involved in responding to contact with the object. The *center position* is intended for use in the generation of arm movements (although this is not currently used in the model).

<b>Variable</b>	<b>Description</b>
Name	
Grasp class	One of <i>precision</i> , <i>power</i> , or <i>side</i>
Aperture	Final width (at time of contact)
Joint positions	For each joint, the position at which contact with the object is made.
Max aperture	For each joint, the position at which the maximum aperture is achieved during the preshape
Jacobian	For each joint, a measure of the derivative of the joint with respect to the aperture (at the point of contact)
Expected tactile input	A bit mask representing those tactile pads that will be active when a grasp is established.
Center	The T-matrix describing the transformation from the wrist to the center of the grasp. The orientation represents the opposition vector (between VF1 and VF2).

Table A.1: Description of grasp.

<b>Variable</b>	<b>Description</b>
Name	
Shape	One of <i>cylinder</i> , <i>sphere</i> , or <i>block</i>
Parameters	List of parameters and their corresponding values. Depending upon object shape, the parameters may include: <i>width</i> , <i>length</i> , and <i>depth</i> .

Table A.2: Description of objects.

The object description table describes the shape and size of each object that is to be presented to the model (Table A.2). This information is used in deriving the representation of objects in PIP and IT.

Finally, the object/grasp table lists for each object the set of grasps that are appropriate for the object. This table is used in establishing the set of affordances for each object.

## Number and Distribution of Neurons

The number of neurons in each layer is presented in Table A.3. Due to computational limitations, the different experiments that are performed with the model rely on different network configurations. Although the classes of defined cells in AIP and F5 (grasp type, etc.) might change from one experiment to another (as do the affordance mappings from PIP to AIP), the ratios of basic cell types (e.g., in F5: set, extension, flexion, hold, release) remain essentially the same. In the bigger picture, we can imagine the network used for each experiment as being a component of a much larger network. The general robustness that has been observed in the model (with respect to numbers and types of cells) leads us to believe that the model will scale to this more general case.

Table A4 details the number and distribution of F5 and AIP neurons for two experiments described in Chapter 6.

Layer	Number of Units	Comments
PIP	183	
Object classes	3	cylinder, sphere, block
General shape neurons per class	1	
Neurons per parameter per class	30	
AIP	50 - 232	Number depends upon the number of grasps represented.
F5	213 - 750	Depends on number of grasps (type and aperture size) represented.
MI	480	divided evenly between position, force, and anti-force
SI	178	
Tactile cells	28	
Hand joints	15	
Proprioceptive neurons per joint	10	
SII	6 - 24	Depends on number of grasps (type and aperture size) represented.
BG	16	
Phasic	10	Two units for each motor program phase.
F5_Recurrent	1	Inhibitory feedback.
BG_Recurrent	1	Inhibitory
Grasp_Bias	4	
F2	4	One for each grasp that can be associated with an arbitrary stimulus.
F6	5	Four units for task-based selection of grasp type; one for general priming of the grasping program.
Area 46	213 - 750	Depends on number of F5 units.

Table A.3: Number of cells in each model layer.

<b>Description</b>	<b>1 object/2grasps: Prediction #2</b>	<b>Size Variation: Prediction #4</b>
F5 Cells	430	750
AIP Cells	110	232
F5 Cell Grasp Orientation		
Precision Pinch	242	750
Lateral Pinch (side opposition)	188	-
F5 Cell Parameter Coding		
General (no specificity)	170	82
Aperture	260	668
F5 Cell Phasic Responses		
set	56	100
extension	197	278
flexion	202	286
hold	65	139
release	50	115
AIP Parameter Coding		
General (no specificity)	51	24
Aperture	59	208
AIP Grasp Orientation		
Precision Pinch	63	232
Lateral Pinch	47	-
AIP Phasic Responses		
set	46	98
extension (early)	46	98
flexion (late)	18	36
AIP V/M Orientation		
Pure Visual	11	19
Visual Dominant (V>M)	49	135
Motor Dominant (V<M)	30	49
Pure Motor	20	29

Table A.4: Distribution of F5 and AIP cells for two experiments (Predictions #2 and #4 of Chapter 6).

## Network Wiring Rules

The following define the heuristic network wiring rules. Although the network is programmed using information from discrete tables (finite number of objects, grasps, and sizes), because the network makes use of a population encoding scheme to represent objects and grasps, a reasonable degree of generalization is achieved. Thus, rather than a discrete set of objects and grasps, the system is capable of representing entire classes.

### PIP -> AIP

For each affordance/AIP unit pair (set-related units only):

If the grasp type of the AIP unit matches the affordance grasp type:

If the AIP unit does not encode aperture:

With a given probability, associate the representation of the object in PIP to this AIP unit (connection is made to the signal input of the AIP column).

The strength of the connections is proportional to the AIP cell's *visual* response, and a random value (selected from a uniform distribution). So cells with a zero visual response will not receive PIP inputs.

Otherwise if the AIP unit aperture matches the object width\* (or other parameter encoding the object dimension along the opposition vector):

With a given probability, associate the PIP representation of the object parameter to the AIP unit (connection is made to the signal input).

The strength of connection is proportional to the AIP cell's *visual* response, the degree of match\* between the object width and AIP aperture, and a random value (selected from a uniform distribution).

\*Match is related to the difference between object width and grasp aperture. Degree of match is  $\text{gauss}(\text{width-aperture}, \text{std})$ , where *std* is a parameter that determines the sensitivity to the difference.

### IT -> AIP

For each affordance/AIP unit pair (set-related units only):

If the object is "known" by the model and the grasp type of the AIP unit matches the affordance grasp type:

If the AIP unit aperture matches the object width (or other parameter encoding the object dimension along the opposition vector):

With a given probability, associate the IT representation of the object parameter to the AIP unit (connection is made to the signal input).

The strength of connection is proportional to the AIP cell's *visual* response, the degree of match between the object width and AIP aperture, and a random value (selected from a uniform distribution).

### **AIP -> AIP**

Four possible conditions: S->S, S->G, G->S, and G->G (where S=specific - i.e. parameters, G=general - no parameters).

For each pair of AIP units:

If the AIP units match in grasp type and in aperture size (latter applies only to S->S):

With a given probability, a connection is established from one column to the support input of the other column. The probability is a function of the condition.

The strength of connection is proportional to a constant that is a function of the condition, a random value, and the degree of match between the two apertures (S->S case).

### **AIP -> F5**

Two conditions: S->S and G->G

For each AIP/F5 pair that falls into one of the two conditions:

If the grasp type matches:

If G->G case:

With a given probability, a supporting connection is established from the AIP cell to the F5 cell.

The strength of connection is proportional to a random value.

If S->S case and the apertures match:

With a given probability, a supporting connection is established from the AIP cell to the F5 cell.

The strength of connection is proportional to a random value, and to the degree of match between the two.

**F5 -> AIP**

Two conditions: S->S and G->G

For each F5/AIP pair that falls into one of the two conditions:

If the grasp type matches and the phases match\*:

If G->G case:

With a given probability, a signal connection is established from the F5 cell to the AIP cell.

The strength of connection is proportional to a random value.

If S->S case and the apertures match:

With a given probability, a signal connection is established from the F5 cell to the AIP cell.

The strength of connection is proportional to a random value, and to the degree of match between the two.

\* If the AIP cell is a set onset cell, then all phases match. If the onset is early, then no connections are established from F5 set cells; if the onset is late then no connections are established from F5 set or extension cells.

**F5 -> F5**

Four possible conditions: S->S, S->G, G->S, and G->G (where S=specific, G=general).

For each pair of F5 units:

If the F5 units match in grasp type and in aperture size (latter applies only to S->S):

With a given probability, a connection is established from one column to the support input of the other column. The probability is a function of the condition.

The strength of connection is proportional to a constant that is a function of the condition, a random value, and the degree of match between the two apertures (S->S case).

**F5 -> Mcx**

For each F5 unit that specifies an aperture:

Compute the goal joint positions for this cell (taking into account responsiveness to tactile inputs)\*. These goals are a function of the type of grasp that the cell codes and its phase.

Randomly select a subset of joints to drive.

Recruit a set of Mcx cells that achieve the goal positions for the selected joints. This computation is performed using a pseudo-inverse with corrections to remove negative weights.

\*The goal positions recruited by and F5 unit depends on the phase during which the cell responds (see Chapter 5):

Set phase: no connections

Extension phase: Mcx position cells are recruited that achieve the maximum aperture.

Flexion phase: Mcx position cells are recruited that achieve the expected object width. Mcx inverse-force cells are recruited in anticipation of contacting the object.

Hold phase: Mcx position units maintain the approximate finger positions, force and inverse-force cells are recruited such that a specific force is maintained.

Release phase: Mcx position and force cells are recruited that cause the fingers to open beyond the size of the object.

## **F5 -> SII**

For each grasp program (grasp type and aperture) that is specified by the grasp description table, three SII cells are allocated (which detect maximum aperture, object contact, and object release).

F5 aperture-specific cells provide positive priming connections to the corresponding SII units:

extension-related cells in F5 prime maximum aperture detectors.

flexion-related cells in F5 prime contact detectors.

release-related cells prime SII release detectors

## **SII -> F5**

Maximum aperture detectors provide signal inputs to F5 flexion units.

Contact detectors provide signal inputs to F5 hold units.

Release detectors inhibit F5 release units and cause the grasping program to shut down.

**SI -> SII**

Maximum aperture detectors receive proprioceptive information for all joints that correspond to the maximum aperture specified by the corresponding grasp table element.

Contact detectors receive both proprioceptive and tactile (force) inputs. The tactile inputs that are received correspond to those pad at which a tactile input is expected.

Release detectors receive inputs from force inputs.

**SI -> Mcx**

Discussed at the end of Chapter 5.

**F5 <-> Area 46**

Each unit in F5 is paired with a unit in area 46. During execution of a grasp in the light, the A46 unit remembers the highest value achieved by the F5 unit. This acts as a memory of the grasp that was executed. When a subsequent grasp is executed in the dark, the A46 unit provides positive support to the F5 unit that is proportional to the memorized activity level.

**AIP <-> BG**

Every AIP unit sends a positive connection to the BG\_AIP\_Recurrent unit. An inhibitory signal is returned to all AIP units.

**F5 <-> BG**

For the interaction between the F5 cells and the BG\_F5\_Recurrent unit, the interaction is the same as in AIP.

Also through the BG, it is possible to bias the selection of one of the grasp types. One of four BG units is allocated to each of the grasp types. Activation of one of these BG cells provides positive support to those F5 units that code the grasp (but do not encode a specific aperture).

**F6 -> F5/AIP**

The F6 unit responsible for detecting the context of the experiment primes all units in F5 and AIP.

**F6/F2 -> BG**

Task-related constraints (as represented in F6) and abstract stimuli (as detected by F2) can bias the grasp selection.

**Protocols****Vanilla Protocol**

The basic protocol (grasping in the light) proceeds as follows:

1. At the beginning of the trial, the object is presented by setting up the relevant representations in PIP and IT.
2. After a delay period, the *ready signal* is presented.
3. After another delay period, the *go signal* is given. The system responds by producing a preshape and a grasp.
4. After a hold period, the *secondary go signal* is given. The system responds by releasing the object.

**Grasping in the Dark**

1. Object is presented and grasped in the light, during which area 46 collects a representation of the F5 program.
2. Object representations in IT and PIP are cleared (turning out the lights). The memory in area 46 is imposed on F5. Protocol then proceeds the same as in steps 2-4 in the vanilla protocol.

## **Object Fixation**

At the beginning of the trial, the object is presented by setting up the relevant representations in PIP and IT. No other stimuli are presented during the trial.

### **1 Object/2 Grasps (conditional case)**

1. At the beginning of the trial, the object is presented by setting up the relevant representations in PIP and IT.
2. After a delay period, the *ready signal* is presented.
3. After a delay, the *instruction stimulus* is presented by setting up an activity pattern in F2.
4. After another delay period, the *go signal* is given. The system responds by producing a preshape and a grasp.
5. After a hold period, the *secondary go signal* is given. The system responds by releasing the object.

### **1 Object/2 Grasps (non-conditional case)**

1. At the beginning of the trial, the object is presented by setting up the relevant representations in PIP and IT. The pattern corresponding to the task-constraints is set in F6.
2. After a delay period, the *ready signal* is presented.
3. After another delay period, the *go signal* is given. The system responds by producing a preshape and a grasp.
4. After a hold period, the *secondary go signal* is given. The system responds by releasing the object.

## References

- Alexandre, F., Guyot, F., Haton, J.-P., & Burnod, Y. (1991). The Cortical Column: A New Processing Unit for Multilayered Networks. Neural Networks, **4**, 15-25.
- Arbib, M. A. (1989). The Metaphorical Brain 2: Neural Networks and Beyond. New York: Wiley-Interscience.
- Arbib, M. A. (1990). Programs, Schemas, and Neural Networks for Control of Hand Movements: Beyond the RS Framework. In M. Jeannerod (Eds.), Attention and Performance XIII. Motor Representation and Control Hillsdale, NJ: Lawrence Erlbaum Associates.
- Arbib, M. A., Bischoff, A., Fagg, A. H., & Grafton, S. T. (1995). Synthetic PET: Analyzing Large-Scale Properties of Neural Networks. Human Brain Mapping, **2**, 225-33.
- Arbib, M. A., Iberall, T., & Lyons, D. (1987). Schemas That Integrate Vision and Touch for Hand Control. Cambridge, Massachusetts: The MIT Press.
- Barbur, J. L., Watson, J. D., Frackowiak, R. S., & Zeki, S. (1993). Conscious Visual Perception without V1. Brain, **116**(6), 1293-302.
- Barto, A. G., & Bradtke, S. H. (1991). Real-Time Learning and Control using Asynchronous Dynamic Programming (TR No. 91-57). Department of Computer Science, University of Massachusetts, Amherst.
- Barto, A. G., Sutton, R. S., & Anderson, C. W. (1983). Neuron-like Adaptive Elements That Can Solve Difficult Learning Control Problems. IEEE Transactions on Systems, Man, and Cybernetics, **SMC-5**, 834-46.
- Bekey, G. A., Tomovic, R., & Zeljkovic, I. (1990). Control Architecture for the Belgrade/USC Hand. In T. Venkataraman & T. Iberall (Eds.), Dextrous Robot Hands Springer-Verlag.
- Brinkman, C. (1984). Supplementary Motor Area of the Monkey's Cerebral Cortex: Short- and Long-term Deficits After Unilateral Ablation and the Effects of Subsequent Callosal Section. Journal of Neuroscience, **4**, 918-29.
- Brownell, G. L., Budinger, T. F., Lauterbur, P. C., & McGeer, P. L. (1982). Positron Tomography and Nuclear Magnetic Resonance Imaging. Science, **215**, 619-26.
- Burnod, Y., Grandguillaume, P., Otto, I., Ferraina, S., Johnson, P. B., & Caminiti, R. (1992). Visuomotor Transformations Underlying Arm Movements Toward Visual Targets: a Neural Network Model of Cerebral Cortical Operations. Journal of Neuroscience, **12**, 1435-1453.
- Colby, C. L., Duhamel, J. R., & Goldberg, M. E. (1993). Ventral Intraparietal Area of the Macaque - Anatomic Location and Visual Response Properties. Journal of Neurophysiology, **69**(2), 902-14.

- Corkin, S. (1978). The Role of Different Cerebral Structures in Somesthetic Perception. In C. E. Carterette & M. P. Friedman (Eds.), Handbook of Perception (pp. 105-55). New York: Academic Press.
- Cutkosky, M. R. (1989). On Grasp Choice, Grasp Models and the Design of Hands for Manufacturing Tasks. IEEE Transactions on Robotics and Automation, 5(3), 269-79.
- Dakin, G. (1986) The Role of Motor Schemas in the Selection and Execution of Skilled Prehensile Movements. Senior Honors Thesis, University of Massachusetts, Amherst.
- Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., Mazziotta, J. C., & Fazio, F. (1994). Mapping Motor Representations with Positron Emission Tomography. Nature, 371(6498), 600-2.
- Didday, R. L. (1976). A Model of Visuomotor Mechanisms in the Frog Optic Tectum. Mathematical Biosciences, 30, 169-180.
- Dominey (1993) Models of Spatial Accuracy and Conditional Behavior in Oculomotor Control. Ph.D., University of Southern California.
- Dominey, P. F. (1995). Complex Sensory-Motor Sequence Learning Based on Recurrent State Representation and Reinforcement Learning. Biological Cybernetics, 73(3), 265-74.
- Evarts, E. V., Shinoda, Y., & Wise, S. P. (1984). Neurophysiological Approaches to Higher Brain Function. New York: Wiley-Interscience Press.
- Fagg, A. H. (1993). Developmental Robotics: A New Approach to the Specification of Robot Programs. In G. A. Bekey & K. Y. Goldberg (Eds.), Neural Networks in Robotics (pp. 459-86). Norwell, Massachusetts: Kluwer Academic Publishers.
- Fagg, A. H., & Arbib, M. A. (1992). A Model of Primate Visual-Motor Conditional Learning. Journal of Adaptive Behavior, 1(1), 3-37.
- Fagg, A. H., Lotspeich, D., Hoff, J., & Bekey, G. A. (1994). Rapid Reinforcement Learning for Reactive Control Policy Design in Autonomous Robots. In Proceedings of World Congress on Neural Networks, San Diego, California:
- Fagg, A. H., Lotspeich, D. L., Hoff, J., & Bekey, G. A. (1996). Rapid Reinforcement Learning for Reactive Control Policy Design for Autonomous Robots. In T. Shibata & T. Fukuda (Eds.), Artificial Life in Robotics IEEE Press.
- Fogassi, L., Gallese, V., Fadiga, L., Luppino, G., Matelli, M., & Rizzolatti, G. (1995). Coding of Peripersonal Space in Inferior Premotor Cortex (Area F4). Journal of Neurophysiology, **submitted**.
- Gentilucci, M., Daprati, E., Toni, I., Chieffi, S., & Saetti, M. C. (1995). Unconscious Updating of Grasp Motor Program. Experimental Brain Research, **to appear**.
- Gentilucci, M., Fogassi, L., Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1988). Functional Organization of Inferior Area 6 in the Macaque Monkey I. Somatotopy and the Control of Proximal Movements. Experimental Brain Research, 71, 475-490.

- Gentilucci, M., Toni, I., Chieffi, S., & Pavesi, G. (1994). The Role of Proprioception in the Control of Prehension Movements - A Kinematic Study in a Peripherally Deafferented Patient and in Normal Subjects. Experimental Brain Research, **99**(3), 483-500.
- Georgopoulos, A. P., Ashe, J., Smyrnis, N., & Taira, M. (1992). The Motor Cortex and the Coding of Force. Science, **256**(5064), 1692-5.
- Georgopoulos, A. P., Kalaska, J. F., Caminiti, R., & Massey, J. T. (1982). On the Relations Between the Direction of Two-Dimensional Arm Movements and Cell Discharge in Primate Motor Cortex. Journal of Neurophysiology, **46**, 1527-37.
- Gibson, J. J. (1950). The Perception of the Visual World. New York: Houghton Mifflin Company.
- Giszter, S. F., Mussaivaldi, F. A., & Bizzi, E. (1993). Convergent Force-Fields Organized in the Frog's Spinal-Cord. Journal of Neuroscience, **13**(2), 467-91.
- Goodale, M. A. (1993). Visual Routes to Knowledge and Action. Biomedical Research, **14**(4), 113-23.
- Goodale, M. A., Jakobson, L. S., Milner, A. D., Perrett, D. I., Benson, P. J., & Hietanen, J. K. (1994a). The Nature and Limits of Orientation and Pattern Processing Supporting Visuomotor Control in a Visual Form Agnostic. Journal of Cognitive Neuroscience, **6**(1), 46-56.
- Goodale, M. A., Meenan, J. P., Bulthoff, H. H., Nicolle, D. A., Murphy, K. J., & Racicot, C. I. (1994b). Separate Neural Pathways for the Visual Analysis of Object Shape in Perception and Prehension. Current Biology, **4**(7), 604-10.
- Grady, C. L., Haxby, J. V., Horwitz, B., Schapiro, M. B., Rapoport, S. I., Ungerleider, L. G., Mishkin, M., Carson, R. E., & Hescovitch, P. (1992). Dissociation of Object and Spatial Vision in Human Extrastriate Cortex - Age-Related-Changes in Activation of Regional Cerebral Blood-Flow Measured with O-25 Water and Positron Emission Tomography. Journal of Cognitive Neuroscience, **4**(1), 23-34.
- Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of Implicit and Explicit Grasp Representations in Humans by PET. **in press**.
- Grafton, S. T., Fagg, A. H., Woods, R. P., & Arbib, M. A. (1995). Functional Anatomy of Pointing and Grasping in Humans. Cerebral Cortex, **to appear**.
- Heppreymond, M. C., Husler, E. J., Maier, M. A., & Qi, H. X. (1994). Force-Related Neuronal-Activity in 2 Regions of the Primate Ventral Premotor Cortex. Canadian Journal of Physiology and Pharmacology, **72**(5), 571-9.
- Hoff, B., & Arbib, M. A. (1993). Models of Trajectory Formation and Temporal Interaction of Reach and Grasp. Journal of Motor Behavior, **25**(3), 175-92.
- Hoff, B. R. (1992) A Computational Description of the Organization of Human Reaching and Prehension. Dissertation, University of Southern California.

- Horwitz, B., Grady, C. L., Haxby, J. V., Schapiro, M. B., Rapoport, S. I., Ungerleider, L. G., & Mishkin, M. (1992). Functional Associations Among Human Posterior Extrastriate Brain-Regions During Object and Spatial Vision. Journal of Cognitive Neuroscience, **4**(4), 311-22.
- Iberall, A. R. (1987) A Neural Model of Human Prehension. Ph.D. Dissertation, University of Massachusetts, Amherst.
- Iberall, T. (1988). A Neural Network for Planning Hand Shapes in Human Prehension. In Proceedings of the 1988 Automatic Controls Conference, (pp. 2288-93).
- Iberall, T., Bingham, G., & Arbib, M. A. (1986). Opposition Space as a Structuring Concept for the Analysis of Skilled Hand Movements. In H. Heuer & C. Fromm (Eds.), Generation and Modulation of Action Patterns (pp. 158-73). Berlin: Springer-Verlag.
- Jeannerod, M. (1994). The Hand and the Object - the Role of Posterior Parietal Cortex in Forming Motor Representations. Canadian Journal of Physiology and Pharmacology, **72**(5), 535-41.
- Jeannerod, M., Arbib, M. A., Rizzolatti, G., & Sakata, H. (1995). Grasping Objects: the Cortical Mechanisms of Visuomotor Transformation. Trends in Neuroscience, **18**(7), 314-20.
- Jeannerod, M., Decety, J., & Michel, F. (1994). Impairment of Grasping Movements Following a Bilateral Parietal Lesion. Neuropsychologia, **32**(4), 369-80.
- Kalaska, J. F., & Crammond, D. J. (1992). Cerebral Cortical Mechanisms of Reaching Movements. Science, **255**(5051), 1517-23.
- Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (Ed.). (1991). Principles of Neuroscience. Elsevier.
- Kang, S. B., & Ikeuchi, K. (1995). Toward Automatic Robot Instruction from Perception - Temporal Segmentation of Tasks from Human Hand Motion. IEEE Transactions on Robotics and Automation, **11**(5), 670-81.
- Katayama, M. (1993) A Computational Understanding of Motor Learning Control Using Neural Internal Models for a Multi-Joint Arm. University of Tokyo.
- King, S. M., & Cowey, A. (1992). Defensive Responses to Looming Visual-Stimuli in Monkeys with Unilateral Striate Cortex Ablation. Neuropsychologia, **30**(11), 1017-24.
- Kurata, K., & Hoffman, D. S. (1994). Differential-Effects of Muscimol Microinjection into Dorsal and Ventral aspects of the Premotor Cortex of Monkeys. Journal of Neurophysiology, **71**(3), 1151-64.
- Kurata, K., & Wise, S. P. (1988). Premotor Cortex of Rhesus Monkeys: Set-Related Activity During Two Conditional Motor Tasks. Experimental Brain Research, **69**(2), 327-43.
- Logothetis, N. K., & Pauls, J. (1995). Psychophysical and Physiological Evidence for Viewer-Centered Object Representations in the Primate. Cerebral Cortex, **5**(3), 270-88.

- Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape Representation in the Inferior Temporal Cortex of Monkeys. Current Biology, 5(5), 552-63.
- Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1993). Corticocortical Connections of Area F3 (SMA-Proper) and Area F6 (Pre-SMA) in the Macaque Monkey. Journal of Comparative Neurology, 338, 114-40.
- Luppino, G., Matelli, M., & Camarda, R. M., Gallese, V., Rizzolatti, G. (1991). Multiple Representations of Body Movements in Mesial Area 6 and the Adjacent Cingulate Cortex: An Intracortical Microstimulation Study in the Macaque Monkey. Journal of Comparative Neurology, 311, 463-82.
- Luppino, G., Matelli, M., & Rizzolatti, G. (1990). Cortico-Cortical Connections of Two Electrophysiologically Identified Arm Representations in the Mesial Agranular Frontal Cortex. Experimental Brain Research, 82(1), 214-18.
- MacKenzie, C. L., & Iberall, T. (1994). The Grasping Hand. North Holland.
- Matelli, M. (1994). personal communication
- Mitz, A. R., Godshalk, M., & Wise, S. P. (1991). Learning-Dependent Neuronal Activity in the Premotor Cortex. Journal of Neuroscience, 11(6), 1855-72.
- Miyashita, Y., Date, A., & Okuno, H. (1993). Configurational Encoding of Complex Visual Forms by Single Neurons of Monkey Temporal Cortex. Neuropsychologia, 31(10), 1119-31.
- Murata, A. (1995). personal communication
- Murata, A., Gallese, V., Kaseda, M., Kunimoto, S., & Sakata, H. (1993). Hand-Manipulation-Related Neurons of the Parietal Cortex of the Monkey: Further Analysis of Selectivity in Shape, Size, and Orientation of Objects for Manipulation. In Neuroscience, Washington D.C.:
- Napier, J. R. (1956). The Prehensile Movements of the Human Hand. Journal of Bone and Joint Surgery, 38B, 902-13.
- Paulignan, Y., Jeannerod, M., MacKenzie, C., & Marteniuk, R. (1991a). Selective Perturbation of Visual Input During Prehension Movements. 2. The Effects of Changing Object Size. Experimental Brain Research, 87, 407-20.
- Paulignan, Y., MacKenzie, C., Marteniuk, R., & Jeannerod, M. (1991b). Selective Perturbation of Visual Input During Prehension Movements. 1. The Effects of Changing Object Position. Experimental Brain Research, 87, 407-20.
- Picard, N., & Smith, A. M. (1992a). Primary Motor Cortical Activity Related to the Weight and Texture of Grasped Objects in the Monkey. Journal of Neurophysiology, 68(5), 1867-81.
- Picard, N., & Smith, A. M. (1992b). Primary Motor Cortical Responses to Perturbations of Prehension in the Monkey. Journal of Neurophysiology, 68(5), 1882-94.

- Pons, P. T., Garraghty, P. E., & Mishkin, M. (1992). Serial and Parallel Processing of Tactual Information in Somatosensory Cortex in Rhesus Monkeys. Journal of Neurophysiology, **68**(2), 518-27.
- Quintana, J., & Fuster, J. M. (1993). Spatial and Temporal Factors in the Role of PRefrontal and Parietal Cortex in Visuomotor Integration. Cerebral Cortex, **3**(2), 122-32.
- Rizzolatti, G. (1987). Functional Organization of Inferior Area 6. In Motor Areas of the Cerebral Cortex (CIBA Foundation Symposium 132), (pp. 171-186). Chichester: John Wiley and Sonns.
- Rizzolatti, G., Gentilucci, M., Camarda, R. M., Gallese, V., Luppino, G., Matelli, M., Fogassi, L. (1990). Neurons Related to Reaching-Grasping Arm Movements in the Rostral Part of Area 6 (area 6aB). Experimental Brain Research, **82**, 337-50.
- Rizzolatti, G. (1995). personal communication.
- Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional Organization of Inferior Area 6 in the Macaque Monkey II. Area F5 and the Control of Distal Movements. Experimental Brain Research, **71**, 491-507.
- Roland, P. E., Larsen, B., Lassen, N. A., & Skinhøf, E. (1980). Supplementary Motor Area and Other Cortical Areas in Organization of Voluntary Movements in Man. Journal of Neurophysiology, **43**, 118-36.
- Rolls, E. T. (1995). Learning-Mechanisms in the Temporal Lobe Visual-Cortex. Behavioral Brain Research, **66**(1-2), 177-85.
- Rolls, E. T., & Tovee, M. J. (1995). Sparseness of the Neuronal Representation of Stimuli in the Primate Temporal Visual-Cortex. Journal of Neurophysiology, **73**(2), 713-26.
- Sakata, H. (1991). personal communication.
- Sakata, H. (1994). personal communication.
- Sakata, H., Taira, M., Murata, A., & Mine, S. (1995). Neural Mechanisms of Visual Guidance of Hand Action in the Parietal Cortex of the Monkey. Cerebral Cortex, **5**(5), 429-38.
- Shadmehr, R. (1993). Control of Equilibrium Position and Stiffness through Postural Modules. Journal of Motor Behavior, **25**(3), 228-41.
- Shadmehr, R., Mussiavaldi, F. A., & Bizzi, E. (1993). Postural Force-Fields of the Human Arm and Their Role in Generating Multijoint Movements. Journal of Neuroscience.
- Silverman, I. E., Grossman, M., Galetta, S. L., Liu, G. T., Rosenquist, A. C., & Alavi, A. (1995). Understanding Human Visual-Cortex - The Role of Functional Imaging.
- Stein, J. F. (1992). The Representation of Egocentric Space in the Posterior Parietal Cortex. Behavioral and Brain Sciences, **15**, 691-700.

- Sutton, R. S. (1988). Learning to Predict by the Methods of Temporal Differences. In Machine Learning 3 (pp. 9-44). Boston: Kluwer Academic Publishers.
- Taira, M., Mine, S., Georgopoulos, A. P., Murata, A., & Sakata, H. (1990). Parietal Cortex Neurons of the Monkey Related to the Visual Guidance of Hand Movement. Experimental Brain Research, **83**, 29-36.
- Talairach, J., & Tournoux, P. (1988). Co-planar Stereotaxic Atlas of the Brain. New York: Thieme Medical Publishers.
- Tanaka, K. (1993). Neuronal Mechanisms of Object Recognition. Science, **262**(5134), 685-8.
- Tanji, J. (1994). The Supplementary Motor Area in the Cerebral Cortex. Neuroscience Research, **19**, 251-68.
- Toni, I. (1994). personal communication.
- Uno, Y., Fukumura, N., Suzuki, R., & Kawato, M. (1993). Integration of Visual and Somatosensory Information for Preshaping Hand in Grasping Movements. In Advances in Neural Information Processing Systems, 5 (pp. 311-8).
- Weiskrantz, L., Warrington, E. K., Sanders, M. D., & Marshall, J. (1974). Visual Capacity in the Hemianopic Field Following a Restricted Occipital Ablation. Brain, **97**, 709-28.
- Widrow, B. (1962). Generalization and Information Storage in Networks of Adaline "Neurons". In M. Yovits, G. Jacobi, & G. Goldstein (Eds.), Self-Organizing Systems (pp. 435-461). Washington: Spartan Books.
- Wise, S. P., & Mauritz, K. H. (1985). Set-related Neuronal Activity in the Premotor Cortex of Rhesus Monkeys: Effects of Changes in Motor Set. In Proceedings of the Royal Society of London, 223 (pp. 331-354).
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). Are Arm Trajectories Planned in Kinematic or Dynamic Coordinates - An Adaptation Study. Experimental Brain Research, **103**(3), 460-70.
- Woods, R. P., Cherry, S. R., & Mazziotta, J. C. (1992). Rapid Automated Algorithm for Aligning and Reslicing PET Images. Journal of Computer Assisted Tomography, **115**, 565-87.
- Woods, R. P., Mazziotta, J. C., & Cherry, S. R. (1993). MRI-PET Registration with Automated Algorithm. Journal of Computer Assisted Tomography, **17**, 536-46.
- Young, M. P., & Yamane, S. (1992). Sparse Population Coding of Faces in the Inferotemporal Cortex. Science, **256**(5061), 1327-31.

# Index

- A.T. 31, 73
- activating signal 47
- activation 24
- active memory 43, 139
- affordance 51, 71, 115, 145
- AIP 5, 40, 48, 54, 100, 115, 134, 139, 145, 156, 164, 168, 172, 185
  - behavioral parameters 71
  - motor-related activity 7, 106
  - object-specific activity 9
  - position-selective activity 11
  - size-selective activity 11
  - temporal behavior 11, 101
  - visual-related activity 7, 107
  - visual/motor-related activity 102, 117
- anterior intra-parietal area 5, See also
  - AIP
- area 4 - 44
- area 46 - 55, 105
- basal ganglia 50, 74, 81
- behavioral descriptions 193
- BG 50, 74, 81, 156, 168
- blindsight 32
- column 63
- computational constraints 51
- conditional task 121, 161
- coordinated control program 82
- corticospinal pathway 44
- D.F. 31, 35
- dark grasping task 104
- dorsal premotor cortex 19, 55, 56, 121, 156, 163, 167, See also F2
- dorsal stream 30, See also how system
- equilibrium position 88
- F1 - 21, 44, See also primary motor cortex
- F2 - 19, 55, 56, 78, 121, 156, 167, See also
  - dorsal premotor cortex
- F3 - 55, See also SMA-proper
- F4 - 5, 16, 39, 189, See also inferior
  - premotor cortex
- F5 - 5, 16, 19, 39, 48, 91, 100, 139, 156, 167, 172, 185, See also inferior premotor
  - cortex
  - behavioral description 75
  - temporal behavior 101
- F6 - 55, 121, See also preSMA
- feature detector 174
- fixation task, 104

- go signal 6, 203, 204
- grasp description 193
- grasp perception 191
- Hebbian/Anti-Hebbian learning
  - algorithm 177
- how pathway 30, 36, See also dorsal
  - stream
- hyperfeature 87, 139
- inferior parietal cortex 164
- inferior premotor cortex 5, 16, 39, See also
  - F4, F5
- inferotemporal cortex 30, See also IT
- instantiation 24
- instruction stimulus 121, 156, 204
- internal model 92
- internal reinforcement 179
- IT 30, 39, 69, 73, See also inferotemporal
  - cortex
- key phase 6
- language of motor acts 191
- lateral pinch 26
- leaky-integrator 62, 65
- learning algorithm 172
- $\mu$ -schema 62
- manipulability 173
- maximum selector 74
- MI 44, 91, See also F1, primary motor
  - cortex
  - motor assembly 45
  - motor-related cells 41
- muscimol 7, 19
- negative reinforcement 177
- network wiring rules 198
- normalization 178
- normalized cell response 109
- object description 194
- object parameter coding 145
- object phase 6
- object size perturbation 139
- object/grasp table 194
- opposition space 25
- output unit 63
- pad opposition 26
- palm opposition 26
- parameter coding 53
- perturbation 139
- PET 153, 160
- phasic projections 83
- PIP 38, 54, 68, 71, 173, See also posterior
  - intraparietal area

- population analysis 109
- population code 38, 39, 68, 74, 85, 86, 134, 140, 198
- positive reinforcement 177
- positron emission tomography 153, See also PET
- posterior intraparietal area 38, See also PIP
- power grasp 26, 157, 161
- pre-SMA 55
- precision pinch 115, 121, 157, 161
- premotor cortex
  - dorsal 56, See also F2
  - inferior 39, See also F4, F5
- primary motor cortex 21, 44, See also F1
- primary somatosensory cortex 86, See also SI
- priming 24
- priming signal 47
- priming unit 63
- protocol 5, 29, 203
- R.V. 31
- rCBF 153
- ready signal 5, 203, 204
- real object 34
- reflex loops 45
- region of competition 175
- regional cerebral blood flow 153
- reinforcement learning 173
- reinforcement prediction 179
- reinforcement signal 176
- relative synaptic activity 155
- revolute spring 88
- schema
  - activation 24
  - competition 25
  - cooperation 25
  - instantiation 24
- schema theory 24, 61
- secondary go signal 203, 204
- SI 44, 86, 91, See also primary somatosensory cortex
- side opposition 26, 121, 157
- signal unit 63
- SII 49, 86, 91, 92, 99, 139, 158
- size coding 134
- SMA 54, See also F3
- SMA-proper 55, See also F6
- somatosensory cortex 44
- sparse population code 39

stability 173

sub-schema 62

supplementary motor area 54

support unit 64

synaptic activity 154

synthetic PET imaging 154

Talairach coordinate system 162

teacher 172

temporal coordination 189

trigger stimulus 6, 29, 91

ventral intraparietal area 38, See also  
VIP

ventral stream 30, See also what  
pathway

VF1 26

VF2 26

VIP 38, 55, 189, See also ventral  
intraparietal area

virtual finger 26

virtual object 34

visual-related cells 41

weight normalization 178

what pathway 30, 36, See also ventral  
stream

where pathway 30, See also dorsal  
stream

wrist orientation 188