

AD/A-004 090

**VIPS: A VISUAL IMAGERY AND PERCEPTION
SYSTEM; THE RESULT OF A PROTOCOL
ANALYSIS. VOLUME I**

Arthur Melvin Farley

Carnegie-Mellon University

Prepared for:

**Air Force Office of Scientific Research
Advanced Research Projects Agency**

May 1974

DISTRIBUTED BY:

NTIS

**National Technical Information Service
U. S. DEPARTMENT OF COMMERCE**

**Best
Available
Copy**

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR - TR-75-0062	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER AD/A-004090
4. TITLE (and Subtitle) VIPS: A VISUAL IMAGERY AND PERCEPTION SYSTEM; THE RESULT OF A PROTOCOL ANALYSIS, VOL I		5. TYPE OF REPORT & PERIOD COVERED Interim
7. AUTHOR(s) Arthur Melvin Farley		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Carnegie-Mellon University Computer Science Dept. Pittsburgh, PA 15213		8. CONTRACT OR GRANT NUMBER(s) F44620-73-c-0074
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, Virginia 22209		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61101D AO-2466
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Air Force Office of Scientific Research (NM) 1400 Wilson Boulevard Arlington, Virginia 22209		12. REPORT DATE May 1974
		13. NUMBER OF PAGES 158
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		16. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Reproduced by NATIONAL TECHNICAL INFORMATION SERVICE U S Department of Commerce Springfield VA 22151 PRICE SUBJECT TO CHANGE		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) VIPS is a program (implemented in LISP 1.5) which embodies a psychological theory of visual imagery and form perception. An experimental task was designed which required subjects to perform perceptual activity over an extended time frame. The subject was required to move a small viewing hole about a presented line drawing until being able to produce a verbal recall description and drawn reproduction of the whole picture. The subject was also instructed to "think aloud" during the hole movement procedure. Transcriptions of selected video-tape protocols served as the immediate basis for VIPS design. Relevant research results of cognitive psychology provided addition		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

(continued)

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. abstract (continued)

al design criteria. Behavioral correspondences between protocols and equivalent traces of VIPS activity indicate theory sufficiency. The theory proposes that visual form perception is a constructive activity involving the coordinated use of several memories and processes, which are defined. The resultant visual image (the perception) is realized by the integration of a succession of views of (fixations upon) the external environment stimulation. A proposal for the form and content of symbolic visual imagery is developed as a necessary part of VIPS. A comparison of recent research indications leads to the proposal that flexibility of information representation is a fundamental aspect of human cognition.

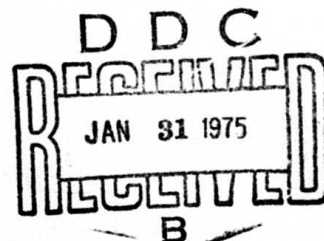
037068

VIPS: A Visual Imagery and Perception System;
the result of a protocol analysis

Volume I

Arthur Melvin Farley

Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
May, 1974



AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)

NOTICE OF TRANSMITTAL TO DDC

This technical report has been reviewed and is
approved for public release (AWR 100-12 (7b)).
Distribution is unlimited.

D. W. TAYLOR

Technical Information Officer

DEPARTMENT of COMPUTER SCIENCE

Approved for public release;
distribution unlimited.



NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, VA 22161

Carnegie-Mellon University

AD A004090

**VIPS: A Visual Imagery and Perception System;
the result of a protocol analysis**

Volume I

Arthur Melvin Farley

**Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
May, 1974**

**Submitted to Carnegie-Mellon University in partial
fulfillment of the requirements for the degree of
Doctor of Philosophy.**

**This work was supported by the Advanced Research Projects Agency of
the office of the Secretary of Defense (F44620-73-C-0074) and is
monitored by the Air Force Office of Scientific Research. This doc-
ument has been approved for public release and sale; its distribution
is unlimited.**

PREFACE

THIS WORK WAS PROMPTED BY A NUMBER OF EXPERIENCES WHICH INDICATED TO ME THAT VISUAL PERCEPTUAL EXPERIENCE IS NOT (EVER) MERELY A REFLECTION OF THE STATE OF AN EXTERNAL STIMULATING ENVIRONMENT. THIS LED TO A DESIRE TO UNDERSTAND HOW THE PERCEPTUAL PROCESSES OF VISION CREATE PERCEPTIONS FROM THE STRUCTURED STIMULATION REACHING THE RETINA. THE INVESTIGATION IS AS YET INCOMPLETE, VIPS BEING AN INITIAL MODEL REQUIRING FURTHER EXTENSION AND SPECIFICATION.

I GRATEFULLY ACKNOWLEDGE THE HELP OF MS. MERCEDES "BUNNY" KOSTKAS IN THE PREPARATION OF THIS MANUSCRIPT AND THE HELP OF DR. THOMAS MORAN IN THE IMPLEMENTATION OF THE IMAGERY AND PERCEPTION SYSTEM.

iii

Reproduced from
best available copy.

ABSTRACT

VIPS is a program (implemented in LISP 1.5) which embodies a psychological theory of visual imagery and form perception. An experimental task was designed which required subjects to perform perceptual activity over an extended time frame. The subject was required to move a small viewing hole about a presented line drawing until being able to produce a verbal recall description and drawn reproduction of the whole picture. The subject was also instructed to "think aloud" during the hole movement procedure. Transcriptions of selected video-tape protocols served as the immediate basis for VIPS design. Relevant research results of cognitive psychology provided additional design criteria. Behavioral correspondences between protocols and equivalent traces of VIPS activity indicate theory sufficiency. The theory proposes that visual form perception is a constructive activity involving the coordinated use of several memories and processes, which are defined. The resultant visual image (the perception) is realized by the integration of a succession of views of (fixations upon) the external environment stimulation. A proposal for the form and content of symbolic visual imagery is developed as a necessary part of VIPS. A comparison of recent research indications leads to the proposal that flexibility of information representation is a fundamental aspect of human cognition.

iv

VOLUME I

TABLE OF CONTENTS

CHAPTER 1.1 ... Introduction	... 1
1. Imagery ... 1	
HISTORY OF IMAGERY ... 1	
MEANING OF "IMAGE" ... 5	
VISUAL IMAGE SPECIFICATION ... 8	
VISUAL IMAGE CLASSIFICATION ... 10	
2. Visual Form Perception ... 11	
CLASSICAL THEORY ... 12	
GESTALT THEORY ... 13	
FORMAL AND MECHANICAL PATTERN RECOGNITION ... 14	
AN ALTERNATIVE - CONSTRUCTIVE THEORY ... 15	
3. Goals of the Thesis ... 18	
4. Thesis Content Guide ... 21	
CHAPTER 1.2 ... Experiments and Analysis	... 23
1. Experiments ... 23	
SUBJECTS ... 23	
EXPERIMENT I ... 25	
EXPERIMENT II ... 28	
STIMULI CHARACTERISTICS ... 29	
2. Analysis ... 31	
DATA TRANSCRIPTION ... 31	
TRANSCRIPTION USE ... 34	
CHAPTER 1.3 ... The Image Representation	... 38
1. Image Representation Criteria ... 38	
2. The Visual Image Representation ... 41	
SYMBOLIC REPRESENTATION ... 41	
IMAGE CHUNKS ... 42	
IMAGE CHUNK CHARACTERISTICS ... 44	
VERTEX CHUNK ... 47	

IMAGE BODY ACCESS - "TRAVERSAL" ...	48
OBJECT CHUNK ...	49
LINE CHUNK ...	51
ILINE ...	52
SIDE CHUNK ...	53
FACE CHUNK ...	54
INTERCHUNK RELATIONS ...	57
EQUIVALENCY AND REDUNDANCY ...	58
IMAGE REPRESENTATION FLEXIBILITY ...	59
3. The Image in VIPS ...	60

CHAPTER 14 ... The Form and Activity of VIPS ... 63

1. System Overview ...	63
THE MEMORIES ...	64
1. PIC (PICTure) ...	65
2. VR (Visual Register) ...	67
VI ...	68
EXP and ATT ...	69
DIR and RAN ...	70
3. STM (Short Term Memory) ...	71
Components of Long Term Memory ...	73
4. LTM (Long Term Memory) ...	73
5. RLTM (Recognition Long Term Memory) ...	74
6. ITM (Intermediate Term Memory) ...	76
THE PROCESSES ...	78
1. VI-SYS (Visual Input Process) ...	79
Process Similarities ...	80
Process Control Structure ...	82
2. AA-SYS (Assimilation - Accomodation Process) ...	83
3. REC-SYS (RECOgnition Process) ...	84
4. INC-SYS (INCorporation Process) ...	85
2. The System in Action ...	86
A PROTOCOL - PROGRAM TRACE SEQUENCE ...	86
THEORETICAL IMPLICATIONS ...	92
GOAL RELATED EPISODES ...	93

CHAPTER 15 ... VIPS Extensions ... 95

1. Memory Extensions ...	95
ICLTM (Temporal - Contextual Long Term Memory) ...	95
SLTM (Semantic Long Term Memory) ...	98
2. Three - Dimensional Picture Perception ...	99
RELEVANT PROPOSALS ...	100
A THEORY ...	101
PROTOCOL INDICATIONS ...	103

vi

SPATIAL CONSISTENCY . . .	104
FURTHER PROTOCOL INDICATIONS . . .	105
3. Generalized Visual Image and Perception System . . .	109
IMAGE GENERALIZATION . . .	109
ELEMENT GENERALIZATION . . .	110
PERCEPTION SYSTEM GENERALIZATION . . .	114
THE PERCEPTUAL GOAL . . .	114
PERIPHERAL VISION . . .	116
PERIPHERAL INFORMATION REPRESENTATION . . .	118
THE SINGLE (INITIAL) FIXATION . . .	119
EYEMOVEMENT SEQUENCES . . .	123

CHAPTER 1.6 . . . Conclusion . . . 129

1. Study comparison . . .	124
BAYLOR'S BLOCKS . . .	129
NEWELL'S STIMULUS ENCODING . . .	131
MORAN'S PATH IMAGE . . .	133
REPRESENTATION COMPARISON . . .	134
2. Flexibility of Representation . . .	135
LANGUAGES OF ART . . .	136
A THEORY OF COGNITION . . .	138
PSYCHOLOGICAL INDICATIONS . . .	144
RELATED PROPOSALS . . .	145
THE ROLE OF REPRESENTATION . . .	147
3. Future Research, Conclusion . . .	148

REFERENCES . . . 151

ALL APPENDICES ARE FOUND IN VOLUME II

VIPS: A Visual Imagery and Perception System;
The result of a protocol analysis

Chapter 1.1 Introduction

The title of this thesis conveys the dual nature of its content. Both the imagery system used to represent visual information and the perceptual processes used to understand the external visual environment are investigated. These two concerns are not independent but rather, they are fundamentally interrelated.

This introductory chapter first presents a brief history of the ideas about visual imagery and perception. Then there follows a statement of the goals of the thesis. Finally, an outline of content is presented to guide the reader through or to the chapters that follow.

1. Imagery

HISTORY OF IMAGERY

Imagery has had its ups and downs in psychology. Prior to the twentieth century, psychology was essentially a partition of philosophy concerned with the workings of the mind. The primary basis for any theory at that time was introspection. Imagery and related mentalistic concepts reigned supreme as theoretical constructs.

Champions of such introspective theories, including William James and E. B. Titchener, consistently noted the important role played by imagery in thought and perception. James described the meaning of concrete words to be "sensory images

awakened" (James, 1890, p. 265). Titchener (1909) developed explanations extending this view to include abstract term meanings as well.

But at the very time these two men were developing such ideas, imagery and related mentalistic concepts were beginning their rapid descent from prominence. They were soon to be almost completely abandoned as viable theoretical constructs in American psychology. This rapid fall was primarily the result of the emerging behavioristic outlook. One basic goal of this outlook was to minimize the inferential constructs existing between a given input stimulus and the resulting response. Being such an inferential construct, imagery came under attack.

Imagery had been already shaken by several experimental results, which further acted as a catalyst for the emergence of behaviorism. Experiments in the rote learning of nonsense syllables, first carried out by Ebbinghaus (1885), seemed to indicate the expendability of image constructs. Fernald's (1912) results indicated that memory images were not actually faithful reproductions of reality, even for people claiming "photographic" memory capability. The Würzburg experiments supplied further evidence for thinking in the absence of conscious imagery.

Such damaging experimental findings provided an impetus for the rise of behaviorism. Behaviorism immediately set to work to develop a more parsimonious theory of thought and memory involving only verbal concepts. One of its leading representatives John Watson wrote:

"I should throw out imagery altogether and attempt to show that practically all natural thought goes on in terms of sensori-motor processes in the larynx [word encodings]" (Watson, 1913, p. 174).

Another primary goal of behaviorism was the objectification of psychological observation. Thus, the method of introspection fell to near disuse and general suspicion. As Watson further wrote:

"Introspection forms no essential part of its [psychology's] methods, nor is the scientific value of its data dependent upon the readiness with which they lend themselves to interpretation in terms of consciousness" (Ibid, p 158)

Behaviorists such as Watson believed that the nervous system did not have even limited activation independence, that all stimulation (activation) derived necessarily from sense organ processes. Thus, imagery and introspection were contradictory to basic premises.

Experimental psychologists in the first half of this century tended to work exclusively with easily controllable verbal associations as experimental material. This often involved the use of nonsense syllables, creating a situation which is not conducive to the use or report of imagery (Pavlov, 1971). This research framework played a significant role in keeping imagery dormant in psychology. In contrast, the several introspective results of the period still utilized image constructs.

Around 1960, imagery began to re-emerge, due to an accumulation of both positive and negative factors. Experimental results indicating the existence and use of imagery provided the positive factors. The growth of psychopharmacology and sensory deprivation research provided numerous reports of hallucinatory images. Neurological investigations, particularly those involving direct stimulation of the brain (Penfield and Jasper, 1954), also provided reports of imagery.

Concurrently, the goals of psychological research experienced radical change. The goal for research now became the realization of detailed and operative models of the behaving organism. Model building had not been a goal of either introspective or behavioristic psychology, but did borrow criteria and techniques from both. Model inputs and outputs could be checked objectively against data (behavioristic), but the

psychologist again became a source of checks for the model as designer (introspective). This provided a framework for the incorporation of imagery into an objectively measured, model-based theory.

The emergence of computing machinery and the early successful developments of artificial intelligence served to forward the view of human thought and perception as being activities of information processing. This mentalistic outlook provided further grounds for renewed investigation of imagery.

Negative factors were also involved in imagery's re-emergence. The primary one was a growing dissatisfaction with the progress of behaviorism. As R. Bugeiski (1970) notes: "the nonsense syllable had had its day", and "the promises of behaviorists remained unfulfilled". This dissatisfaction led to a closer scrutiny of the arguments against imagery and the realization that in some ways they were vacuous.

The inaccuracy of visual memory had been held up as a disproof of visual imagery. But all modes of memory had repeatedly been shown to be inaccurate, not just visual. Bartlett (1932) noted this, characterizing memory as a partial construction of reality and not as a reproduction. Bartlett was one of the few psychologists of the time to give imagery a role in human behavior.

The realization also came that verbally oriented thought processes were as much inferred or implicit as those using imagery. Verbal responses could be just as well mediated by inferred mental images as by inferred mental words. In fact, results which indicate the existence of visual recognition memory in infants provide an example of a visual memory (imagery) with the possibility of a verbal encoding eliminated (Fagan, 1972). A third factor was that verbal-based theories were not necessarily more parsimonious. In visual recognition tasks and visual matching tasks one indeed

appeared to be proposing unnecessary processing in the act of translating all input to names or verbal descriptions for comparison.

Experiments noting reaction time differences between name and visual same-different comparisons of visually presented stimuli have produced results favorable to imagery (Tversky, 1969; Posner and Keele, 1968; Posner and Mitchell, 1967). Other experiments involving concept associations have appeared to indicate the utilization of non-word, imaginal mediators (Paivio, 1971). Paivio (1971) has presented a most valuable survey and discussion of experimental results relevant to imagery. Other recent books (Richardson, 1970; Segal, 1970; Chase, 1973) present further results and thought relevant to imagery and the processing of visual information. Imagery has returned as a concept to be investigated in cognitive psychology.

MEANING OF "IMAGE"

What is the meaning of the concept "image" (or "imagery")? The task of definition has not proven easy for psychology. According to a consensus of dictionary definitions an image is a mental representation or imitation of anything. This definition is not adequate for present psychological purposes. As further interest in imagery has developed, problems concerning its precise definition have arisen.

Two definitions representative of differing approaches to the specification of imagery follow. The modality approach to the definition is illustrated as follows, defining the image or imagery as:

"non-verbal memory representations of concrete objects and events, or non-verbal modes of thought (e.g. imagination) in which such representations are actively generated and manipulated by the individual." (Paivio, 1971, p. 12)

The second approach, one stressing modality independence, is as follows:

"The function of memory imagery is to put us in direct contact with HOW things looked, or sounded, or felt, or tasted, as distinct from WHAT they resembled, what they sounded like, looked like, felt or tasted like."

(Bower, 1972, p 51)

Neither definition is satisfactory. Paivio's definition leaves open the critical question of how to distinguish between "verbal" and "non-verbal" representations in the head. This definition of imagery directly confronts problems of epistemological status, ambiguity, and possible contradiction. Bartlett notes these problems to result from the fact that:

"most statements that have been made about images in traditional psychology concern their nature rather than their function, what they are rather than what they make it possible to do" (Bartlett, 1932, p215).

Bower's definition also relies heavily upon phenomenological intuition for its meaning. Bower's definition again concerns primarily the nature of the image and faces the same problems as Paivio's. Furthermore, it fails by implying that the image is an internally observable entity offering a means of "direct contact" with sensory related information. A visual image is not what the little man sees inside but what is inside the little man that sees.

Even accepting these two definitional specifications as being adequate does not yield a favorable situation. The differing degrees of importance placed upon representation modality leads to differing and divergent experimental concerns. Thus, current imagery research is being conducted without the benefit of an adequate and accepted definition as basis. Each investigation must attempt to clarify the concept, both as a preconsideration and as a result for discussion.

The definition offered here specifies the structure, content, and function of the

image. An image (in general) is an internal symbolic semantic representation of information which is capable of determining (guiding) behavior (enabling a response), and which has an internal modality characteristic. "Semantic" here means "meaningful". In the case of an image resulting from perceptual activity, "semantic" implies that the input meaning is preserved in the representation, though not necessarily in its exact form or order. That the information of an image is adequate for behavior determination indicates a basic functional equivalence between the image and sensory input. Bartlett (1932) notes this ability of imagery to free human behavior from strict dependence upon the continuous temporal sequence of environmental sensory input.

Images are but one form of representation within the internal "sensory order". The "sensory order" (Hayek, 1952) bears a limited (but often sufficient) correspondence to the "physical order" of the external environment. The elements of this internal representational order develop to reflect differences between sensory input which are necessary for the realization of adequate (successful) behavior. The "raison d'être" of any image is to provide a basis for adequate behavior.

The internal modality characteristic distinguishes images from other memory representations of meaning and provides a basic relationship to an associated perceptual process (sensory modality). An operational or functional definition of internal modality has been stated by Simon, as follows:

"Information is stored in such a way that it can be retrieved by internal processes isomorphic to the external processes that are used to retrieve information in the corresponding modality from external sources" (Simon, 1972, p. 97).

Therefore, the information within an image of a given modality is structured to be

Reproduced from
best available copy.

effectively and efficiently (straight-forwardly) retrievable by processes isomorphic to those of that perceptual modality. The information is "perceptually structured". The modality attribute of an image's information structure provides a means for the evaluation of any proposed visual image representation.

VISUAL IMAGE SPECIFICATION

This general image definition must be further qualified in visual terms to be sufficient for the purposes of the thesis. As a first step, the visual image is defined as the internal semantic representation of visual information capable of determining visually-related behavior. It remains to specify visual information and visually-related behavior in order to complete the definition.

The information within a visual image is structured so as to be straight-forwardly accessible by processes isomorphic to those which access external information (sensory stimulation) during visual perception. Visual stimulation is accessed by a series of spatially related samplings (scanning) during visual perception. Visual image access must be isomorphic in meaning to visual scanning. A visual image therefore has an implicit, spatially defined point of view. Internal modality produces a criterion for the content and structural characteristics of the visual image, and indicates an initial interrelationship between the visual image and the visual perception process.

Still to be specified are the classes of informational values (symbols) imbedded in the perceptually-based structure. Internal modality influences this image property also. Neurological experimentation has indicated the feature detection capabilities of the frog, cat, and monkey vision (eye) systems (Lettvin, et al, 1962; Hubel and Wiesel, 1962, 1965, 1968). Psychological and psychophysical results have indicated similar

ability and activity within the human vision system (Bartlett, 1932 ; N. Weisstein, 1973). Therefore, one class of information values imbedded in the visual image is that of the symbols representative of detectable visual features.

The other class of image symbols is best specified through a definition of visually-related behavior. The visual image may serve as a source of information to guide two basic types of activity. One is motor activity within a local, spatial environment. This includes not only coordinated physical body action, but also the control of the head and eye movements associated with the visual perception process. The other type of activity is cognitive behavior involving spatial and visual feature content. Examples are the architectural planning and verbal describing of a spatial (visual) environment or scene. The prime cognitive example is visual form perception. A visual image is the basis for perceptual inferences, hypothesis, and expectations.

To be adequate for the determination of these behaviors, a class of spatial information values (symbols) must also necessarily be imbedded into the structure of the visual image. Values sufficient for the specification of position, direction, and distance (range) make up this class of spatially meaningful symbols. Lee Brooks (1968,1973) has reported results from several task interference experiments which indicate both the existence of internal memory modalities and the spatial content of the internal visual representation. These results also indicate that imaging in a modality interferes with activity involving perception in that modality. The interference indicates the close relationship between imagery and perception.

In conclusion, the fully specified definition of the visual image is the internal semantic representation of spatial and visual feature information, structured so as to be straight-forwardly retrievable by internal processes isomorphic to those of visual

perception. A visual image is furthermore capable of acting as a basis for motor and cognitive behavior with regard to the visual environment represented.

The visual perception process has a fundamental definitional influence upon the structure and content of the visual image. The visual image has a correspondingly significant effect upon the perception process, acting as an informational guide for the perceptual activity. The image serves as source of feature expectations and eye (head) movement control specifications during visual perception. Finally, perceptual activity is a primary source of visual images. The resultant perception of the current environment is a visual image.

VISUAL IMAGE CLASSIFICATION

The degree of correlation with external events is the determining factor for one scale of image classification. In order of decreasing external correspondence, a visual image may be classified as a perception, an illusion, a hallucination, or a dream, memory, or mental image. As Segal states:

"When we have been able to ascertain that a given impression shows adequate correlation with external events to satisfy our current needs, it is conventional to call it a 'perception'." (Segal, 1970)

Another comparable classification is based upon the degree to which an image depends upon external (sensory) as opposed to internal sources of information for its content (Horowitz, 1972). At one end of the scale, dreams or mental images rely solely upon internal information sources. Richardson has defined mental imagery to be those conscious mental experiences which "exist for us in the absence of those stimulus conditions that are known to produce their genuine sensory or perceptual

counterparts" (Richardson, 1970). However, perceptions, at the other end of the scale, do not rely solely upon external information. Goals, expectancies, and experience (internal states) have consistently been shown to play significant roles in the realization of perceptual images and to effect their content (Berkeley, 1710; Segal, 1970; Chase and Clark, 1972).

These image classification scales must be continua, with the boundaries between each class being indistinct. The classic, and often replicated, study by Perky (1910) has shown that the conscious experience of imaging and perceiving are not noticeably distinct. In that study, perception is mistaken to be imagery. As for the inverse, studies by Goldiamond and Hawkins (1958) have shown that imagery can be mistaken to be perception. Visual perception and visual imagery appear to be inseparably entangled.

2. Visual Form Perception

As visual imagery and perception are so interrelated, consider now several proposed theories of the objective visual perception of form. Note that visual perception has been narrowed to the area of form perception. The thesis is not directly concerned with brightness, color, or movement perception, though these may be basic participants and/or resultants of form perception. The thesis is an investigation of how the representation of the environment as spatially interrelated forms is realized.

CLASSICAL THEORY

Classical or atomistic theory proposes that the visual perception of form is an "unconscious conclusion" realized by "unconscious inferences" which are based upon the values of the smallest discriminable or homogeneous patches of the stimulus (Helmholtz, 1962). The raw color sensation atoms are automatically associated with and recall the non-visual sensations of form. The inferences are based upon pre-existent knowledge of kinesthetic, factual, and eyemovement experiences which has become associated with the sensations (Gregory, 1970).

One source of criticism for the theory is the proposed unconscious nature of the perceptual inferences. This seemingly implies that the process is not analyzable and that any intermediate stages of processing and information representation are not detectable. To the contrary, recent research has indicated the existence of several perceptual stages (Sperling, 1960; Sternberg, 1967).

Another problem for the theory is the evidence for the feature extraction ability and activity of the vision system noted above. Thus, the basic visual sensation is not of the form of a retinal picture of homogeneous color atoms. Furthermore, there is evidence that distance features may be included in the "sensation". Barlow (1967) has determined that there is retinal activity directly responsive to sufficient location disparity between light input to the two eyes. Distance perception is therefore not solely a result of "unconscious inference" as is proposed by the classical theory. Classical theory's account of depth perception had been one of its major explanatory successes (Gogel, 1968).

GESTALT THEORY

Gestalt theory developed as a reaction to atomistic theory's failure to note and explain the effect of stimulation context upon the valuation of any stimulation atom. Gestalt psychology views visual perception to be the result of the application of major organizational laws to the complete encountered visual field. Such laws as proximity, similarity, continuity, and closure are defined to be the fundamental bases of form perception. These laws predict, with a high degree of satisfaction, which of many logically possible perceptual organizations will in fact be realized for a given presented visual field (Koffka, 1935).

One great problem facing Gestalt theory is the specification of these proposed organizational laws in terms which are not merely intuitive. Realization of this problem has led to attempts to quantify these aspects of the visual environment. One such approach has been to apply information theoretical techniques in the analysis of presented visual environments. Attneave (1954, 1959) provided the first completely developed attempt of such an approach, considering the visual field in terms of its contained redundancy.

Other similarly derived means for precisely defining the Gestalt organizational laws have concerned themselves with measuring the amount of information necessary to describe or specify a figure. The probability of perception, given a presented field, is then seen to be inversely related to such an information measure. This principle of perception, derived from Gestalt origins as noted, has been called the "minimum principle." Hochberg has been involved in several attempts to explain the three dimensional perceptions of straight line drawings in light of this minimum principle (Hochberg and McAllister, 1953; Hochberg and Brooks, 1960). The problem of

adequately specifying these intuitive organizational laws has not yet been solved. Any measure of stimulus characteristics upon which to base the Gestalt law performance necessarily introduces that measure or informational encoding as an intermediary construct, a foreign addition to the basic theory proposal.

Another problem for Gestalt theory is that the organizational "laws" are not actually determinants of behavior. An observer can voluntarily reverse figure and ground or choose between alternative figural perceptions. The means for such activities are not part of the Gestalt theory of perception.

One final, most important failure, is that the "whole" to which the laws apply is not specified or readily specifiable. The fact ignored by Gestalt perceptual theory is that a figure (object) is normally perceived through a succession of glimpses. Partial figural regions are foveally fixated at different times. Partial regions may have significant effects upon the complete figural perception (Smon, 1967; Hochberg, 1968; Gregory, 1970). If Gestalt laws are only applicable to the total visual field, the effects of these partial regions are not explained. If these smaller, contained regions are the "whole" units for law application, Gestalt theory fails to offer any means by which to integrate them into complete scene perceptions.

FORMAL AND MECHANICAL PATTERN RECOGNITION

The program implementation and mathematical formalization of two methods of pattern recognition indicate some inherent capabilities and limitations of these as possible explanations of visual form perception.

Template matching was one of the earliest methods of pattern recognition to be implemented (Rabinow, 1957). A stimulus is matched against a set of recognizable

patterns, with the "best match" being the one recognized. The method was found to have limited applicability due to the adverse effects upon performance by small variances of form and increasing recognition set size. Neisser (1967) discusses template matching as a possible explanation of form perception. He notes significant problems in accounting for observed recognition invariances over retinal position, size, and orientation (in some cases).

Several early pattern recognition programs were based upon feature extraction. These systems incorporated abilities to create and adjust feature operators (Uhr and Vossler, 1961), to adjust feature weightings (L. G. Roberts, 1960), and to form valuable feature combinations (Selfridge, 1959). These abilities have been recently rediscussed by Uhr (1973). A mathematical formalization of this method has indicated several inherent limitations (Perceptrons; Minsky and Papert, 1969). A system using diameter-bounded feature extractors is incapable of determining pattern connectedness. Recognition in context is beyond the scope of a system based solely upon feature extraction.

AN ALTERNATIVE - THE CONSTRUCTIVE THEORY

The Classical and Gestalt theories of form perception and the template and feature extraction methods of pattern recognition are all seen to have serious shortcomings as theoretical explanations. An alternative theory has been developing, based upon the positive and negative results generated by the theories and methods discussed. This theory is called the constructive theory of visual form perception. Its basic premise is that the activity of form perception is the construction of an internal representation which embodies the structure (content) of the external visual

environment (the perception). This constructive act involves the integration of a sequence of informational inputs derived from a series of eye fixations upon the external environment. Representation construction does not imply the continuous generation of new, unique symbol structures. This construction involves the selective activation and association of pre-existent images or image components (symbols and relations). That unique (not previously seen) environments can be perceived indicates the ability to create new structures from available components to construct a visual image. That new visual discriminations can be learned indicates the ability to create new components for image construction.

The constructive theory of visual perception provides a straight-forward explanation for the effects of internal states (i.e. goals) upon perception. Goals can effect the selection of activated image components. Experience has determined those components which are available for image incorporation. Appropriate knowledge may be activated to compensate for erroneous stimulation indications during perception. The effects of water (the bent-stick phenomenon) and eyeglasses (colored or prismatic) can be assimilated into the image construction. Similarly, the ability to voluntarily alter the perception of an environment presents no problem for the theory.

The resonance theory of perception (Gibson, 1956) proposes that perception results from a direct analysis of stimulation structure. In contrast to the constructive theory, resonance theory encounters difficulty in the explanation of the effects of internal states and the ability to voluntarily alter perceptions. Rather than being a phenomenon of resonance, visual perception is better described as a phenomenon of interference or interaction between internal states and sensory stimulation. Cassedy (1942) has even described internal visual representations as a form of "interference

wave patterns". That proposal is not made here, but constructive activity through assimilation and accommodation can be described as a form of interference.

Hebb (1949) began the modern psychological formulation of a constructive theory of form perception. Hebb's model includes the integration of inputs and the direction of eye attention as basic elements. He proposes two classes of higher order neural units: cell assemblies and phase sequences. Cell assemblies are groupings of neurons which are fired frequently and simultaneously by high-probability stimulus patterns. Phase sequences are the further joining of several such cell assemblies which have often been activated in temporal sequence. This joining acts to facilitate such sequential firing (somewhat like an expectation) and to direct attention upon the environment to accomplish the same. One problem is that Hebb's proposal does not relate directly to a level of sensory input, only dealing with learned central processes. The basic idea of information integration during perception is introduced.

Johan Hochberg (1905, 1970, 1973) has become a primary figure in the investigation and specification of the constructive theory. He proposes that a basic aspect of visual perception is an underlying structural organization which makes possible the selective attention upon and successive integration of the visual environment. This organization is embodied by "schematic maps" which are constructed to guide the perceptual activity. These maps are generated according to organizational laws of both Classical (experience) and Gestalt ("minimum principle") origins.

This thesis accepts and attempts to further specify this theoretical proposal. It then equates the "schematic map" with a class of visual images (visual form perception). This is an extension of Hochberg's "schematic map" concept and of its



role in cognition. The perceptual image is constructed both as guide for the perceptual activity and as product of that behavior.

3. Goals of the Thesis

The research which is reported here is a further investigation and specification of the constructive theory of visual form perception. More specifically, the goals have been: (1) to investigate the nature of the processes and memories which are involved in the fixation and integration of successive views of the environment; (2) to investigate the nature of the internal representation (symbolic visual image) which is capable of embodying the necessary partial and complete perceptions; (3) to specify the results of the investigations in the form of an operational, computer-implemented visual imagery and perception system (VIPS). VIPS is such a program which has been implemented in LISP 1.6 on the CMU PDP-10.

The motivation for the two investigations is abundant. The need is best expressed by the following two statements:

"There is little evidence to guide our thinking on how these integrations and constructions take place. In fact, —, little attention has been paid to how such processes occur at all." (Haber and Hershenson, 1973, p174)

"What we need is a set of operations for defining and studying the kind of visual storage that will build up the structures of perceived forms out of momentary glimpses" (Hochberg, 1968, p322).

As for the goal of specifying the results in the form of a computer-implemented model, I believe that one major factor which has contributed to the fluid state of affairs at the theoretical level is the failure to attempt complete specifications and

implementations of the imagery and perceptual processes implied by a stated theory. Accepting the view that perception and imagery are basically symbolic, information processing activities, the computer and several existing programming language systems (ie. LISP) now provide a valuable means for the further investigation of theoretical proposals. Not only can some light be shed on the sufficiency questions for any proposal, but through implementation design and subsequent actual system behavior, other theory implications possibly not realized earlier appear for consideration. What control information is necessary and can it be feasibly included? Is the information representation concise enough to fit into a limited STM (active memory), yet rich enough to afford the inferences which subjects display? These are only two of the many questions which cannot really be answered for a perceptual (cognitive) theory until a completely operational specification and implementation has been realized, or at least attempted, for a reasonable subset of activity. Experimental research in psychology has provided the necessary critical mass of design relevant information to attempt an initial implementation of the theory such as VIPS.

The implementation goal here has been more extensive (inclusive) than that of achieving simple input-output equivalence or correspondence. VIPS employs a number of memory components and perceptual processes which explicitly contain all of the symbolic information involved in the perceptual system's operation. The memory components and their associated characteristics have been inferred from experimental results of cognitive psychology (icon, STM, LTM). Thus, the model attempts to realize correspondence at a less superficial, internal level. The programmed system's output is a trace of memory contents at selected points of processing. A common criticism of computer-implemented models (simulations) of human cognitive behavior has been that

they necessarily employ many hidden, computer-oriented, non-psychological operations to achieve input-output equivalence. This and other recently implemented systems (Newell, 1972; Moran, 1973) attempt to make theoretical implications explicit in the implementation to a greater extent than before and to explore (confront) the consequences.

4. Thesis Content Guide

To accomplish the basic investigatory goal, the necessary first goal of the thesis was to develop an experimental situation which would yield data of value to image and perceptual process specification. Chapter 12 provides a description of the two experiments carried out, of the form of the data acquired, and of the analysis procedure which led to the resultant system specification.

Chapter 13 describes the image representation which has been proposed. A discussion of the basic design criteria utilized is followed by a full explanation of the resulting image specification.

Chapter 14 is an exposition of the proposed perceptual processor at an overview level. Constituent memories and processes and their interrelationships are fully described. Much of this level of specification will be seen to result from a desire to remain consistent with external, relevant psychological findings and theories, especially concerning the memory proposals. The specific operational characteristics and contents of these memories and processes are direct results of the experimentally acquired protocol data. This chapter is the overall theoretical statement of the thesis. The second half of the chapter serves to illustrate the proposed perceptual system and image representation in action.

Chapter 15 discusses possible extensions to the implemented system to account for the subject behavior upon line drawings yielding three-dimension (projectional) perceptions. The perceptual theory embodied by the implemented system is then discussed and shown to be satisfactorily generalizable. In such a general form, it is shown to be useful in accounting for both directly acquired and externally noted eye-

movement data and their associated visual perceptions. Accompanying proposals are also made to realize the extension and generalization of the fully instantiated image representation for line drawings to engulf all visual environments and peripheral vision considerations.

Chapter 16, as the conclusion, contrasts the proposed imagery system and its representational structures to other comparable proposals. A brief description and discussion of a general theory of cognition, research suggestions, and an evaluation of thesis goal realization conclude this report.

In the second volume, Chapter 111 gives complete expositions of the three supplementary perceptual processes and of the two implemented assimilation-accommodation process instantiations which have been inferred from the obtained data. The chapter references appendices displaying the actual implementation listings to avail fuller understanding.

In Chapter 112, behavior of the fully specified and program-implemented perceptual system is contrasted to the observed subject's activity of four selected protocols. A qualitative evaluation, based on several relevant quantitative measures is given. The chapter references appendices containing transcriptions of subject behavior and traces of program activity at several levels of specificity.

Chapter 12 Experiments and Analysis

This chapter is a description of the experimental and subsequent analytical procedures employed in the development of the perceptual processing system proposed by this thesis. The first part of the chapter considers the experiment in terms of subjects, stimuli, required tasks, and acquired data. The second part discusses how this directly acquired data has been processed to serve as a basis for the proposed system's design.

1. Experiments

Two experiments, each consisting of several parts, were conducted to provide the data necessary to investigate those aspects of human visual perceptual behavior of interest within the stated goals of this thesis. The primary design criteria for the experiments used are naturally derivatives of the thesis goals. As such, Experiment 1 first familiarizes the subject with a class of line drawings which are used throughout the following tasks as stimuli. A task situation is then created which forces the subject to integrate a succession of partial picture views to realize a complete perception of the stimulus. The subject must move a small (3 degree visual angle) hole about a line drawing in order to realize its perception. This task yields data relevant to determining the constructive means and capabilities of the subject's perceptual system.

The primary perceptual task of the first experiment is necessarily artificial. The second experiment provides data from a more normal visual perception situation. It

involves the tracking of the subject's eye movements during the perception of a line drawing stimulus. These second results are then interpreted in light of the integrative and imaginal operations inferred from the prior results to further specify a comprehensive theory.

SUBJECTS

It is not a goal of this thesis to investigate the developmental aspects of acquiring the ability to effectively perceive straight line drawings. It is to investigate the nature of the processes and memories involved in the developed, effective perception of such pictures. Therefore, perceptually capable subjects were desired. This is not to deny that picture perception is a learned activity. This has been fairly well substantiated by child observation and experiments introducing pictures to people of cultures which do not use pictures as a spatial representation medium (Miller, 1973). Likewise, the possible influence of any preceding developmental period upon the mature pictorial perception process is not denied.

The desired subjects were persons familiar with the class of line drawings to be presented as stimuli, thus inferentially insuring their possession of a developed system of coordinated processes and memories applicable to the perception of such pictures. The chosen subjects possessed the desired experience, plus the characteristic that they were unaware, to a great extent, of the current research techniques and results in the field of the psychology of perception.

Four subjects were used. The theory to be developed here is of the type which has been called a mini-theory [Newell]. It accounts for few subjects on relatively few behavioral performances. The subjects employed were, at the time, all students of Carnegie-Mellon University, as follows:

Subject B (Bill) - male graduate student in computer science;

Subject P (Paul) - male graduate student in industrial administration;

Subject S (Steve) - male graduate student in computer science;

Subject G (Gwen) - female senior undergraduate in graphic design.

Throughout the experimental performances there appeared to be no great difference of capability between subjects, though they possessed differing backgrounds and current situations. One exception was that G showed some greater sensitivity to size (length) characteristics in her perceptions, possibly an influence of her graphic design study (this difference has been only noted qualitatively, from anecdotal experimental indications). Other subject differences, not in perceptual capabilities but in strategies employed, will be discussed later as results.

EXPERIMENT I

All four subjects completed the three parts which comprise this experiment. Before each part of the experiment the subject was presented with a copy of the instructions for that part. The subject read this aloud, and then was permitted to study the instructions until sure of understanding them. These instructions were all the information given concerning each experiment part.

The instructions for Part I are given in Figure II.1. Two of the reasons for inclusion of this experiment part are indicated in the last paragraph of these instructions. It was felt to be desirable to use a minimum of verbal descriptions of the pictures to be used, and as can be seen all that was said is "line drawing". Figure II.2 presents pictures of the twelve line drawing stimuli used in Part I. From this part, the

EXPERIMENT 1

PART 1

Instruction

You will now be presented with a series of line drawings. You are to look at each drawing until you believe you are capable of fully describing it. You are then to close your eyes and say the word, "ready", which will prompt me to remove the picture. After reopening your eyes you will be required to do three things, as follows:

1. orally describe the line drawing; there is no "correct" description, though completeness is desired
2. answer any questions I may have about the line drawing, or your description
3. draw the line drawing, as you recall it, on the pad in front of you

Do not start to describe the line drawing until after it has been removed. Do not start drawing the line drawing until after you have completed the oral description and answered any questions I ask.

This is a training period to familiarize you with the class of line drawings used throughout later experiments and also to acquaint you with the tasks of describing and drawing from memory.

25b

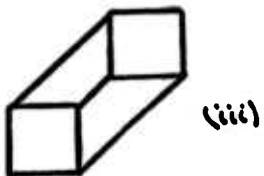
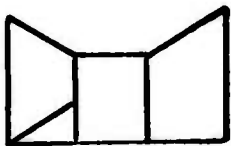
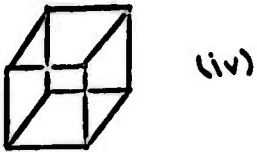
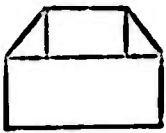
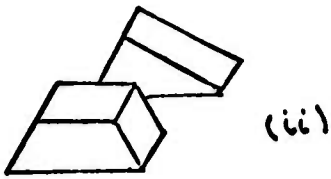
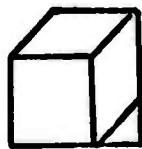
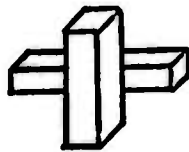
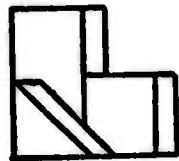
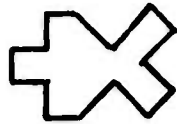
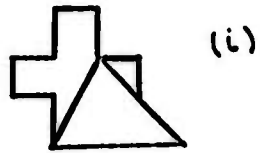


Figure 11.2



subjects learn that only straight lines are employed, and that several different types of these straight line drawings may be expected, these being:

- (1) simple two dimensional, geometric objects, with or without overlay (Figure II.2 (i));
- (2) straight-forward three dimensional projections (i.e. Figure II.2 (ii));
- (3) impossible three-dimensional projections (i.e. Figure II.2 (iii));
- (4) transparent (reversible) three-dimensional projections (i.e. Figure II.2 (iv)).

As noted, the subjects also practiced describing their perceptions and subsequently drawing them from their constructed memory representation. This part of the experiment also had a hidden subject screening purpose. If the subjects were incapable of drawing or describing from memory, or if their perceptions were of the free association type (i.e. "looks like my aunt's house"), then they could be eliminated. However, all of the initial four subjects were sufficiently successful.

The verbal descriptions given were recorded on audio tape, and the accompanying drawing sequences were recorded on video tape to be used as data as will be discussed in the analysis section to follow.

The instructions to Part II of Experiment I are given in Figure II.3 ; and an example of the stimuli employed is given in Figure II.4. This part was conducted in anticipation of Part III. From several pre-experimental trials of Part III, it was found that a vocabulary problem could exist, resulting in the use of hand waving to describe intermediate object constructs. This part was thus primarily a screening part in purpose, though the verbal results obtained and recorded on audio tape were considered to be valuable data as discussed later.

Instructions

You will now be presented with a series of drawings. You should consider each to be a small part of a larger whole. To one of those seen in the previous part. Looking at such a drawing through a viewing device, you may view of the drawing at any time. You do as follows:

1. describe what the partial drawing shows.
2. give any ideas you have as to what the complete given partial picture could be.

I will do one example for clarification.

26a

EXPERIMENT I

PART II

Figure 11.3

Instructions

You will now be presented with a series of partial line drawings. You should consider each to be a small part of a larger line drawing similar to one of those seen in the previous part of the experiment. as if one were looking at such a drawing through a tube which allowed only a small, partial view of the drawing at any time. While viewing the picture you are to orally do as follows:

1. describe what the partial picture is itself
2. give any ideas you have about the whole, of which the
given partial picture could be part

I will do one example for clarification.

Figure 11.4



The instructions to Part III are given in Figure II.5. The experimental procedure and task of Part III is an extension of a method noted by Hochberg (1968). In those preliminary studies, subjects were presented with sequences of partial views enclosed within a circle representative of the viewing hole. The experimenter remained in control of the succession and presentation time of the discrete, partial views. In Part III here the subject has complete control over the hole positioning including the direct movement of the hole. In the Hochberg tasks, sequences of partial views are presented without direct movement control or occurrence. Hole movement in those cases had to be totally inferred or was implied by directional arrows interspersed between views.

Part III then is the significant part of this experiment. The behavior obtained here is to be modeled and explained by the program implemented perceptual system. The "think out loud" verbalizations obtained here are not introspective accounts by the subject concerning his behavior. They are primarily short verbal cues indicative of his internal processing status.

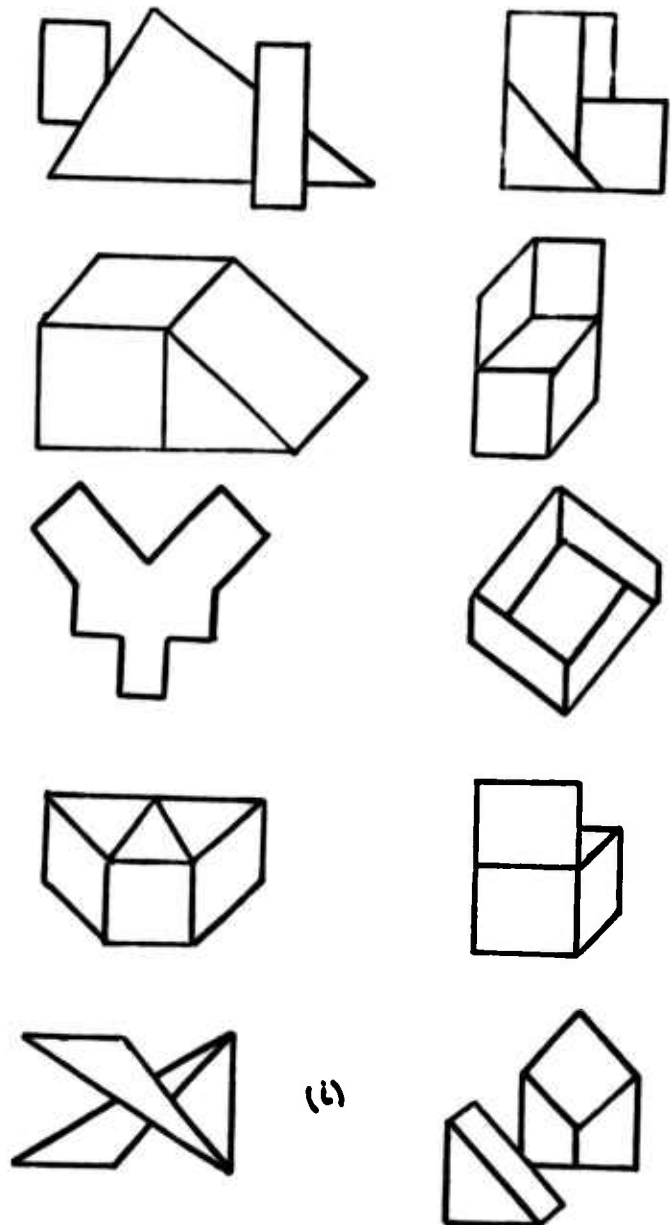
The pictures employed in this part were "similar in type" to those of Part I. No picture from Part I was employed. Figure II.6 presents pictures of the line drawing stimuli used in Part III. The hole-movement and verbalization protocol, verbal description, and drawing sequence were all recorded on video tape. I found it unnecessary to ask questions of the subjects concerning their descriptions, as I also had found in Part I.

All the subjects were capable of this task. However, all subjects found it to be a demanding task and also a tiring one. One can qualitatively note a degradation in performance and weariness of voice during the last protocols of each

EXHIBIT I
PART III

—

Before you is a cardboard and paper shield, which will easily move about the table top. Try this now.



subject. These indications are in line with the design goal that the task given force the subject to extensively utilize his perceptual ability and memories.

EXPERIMENT II

This experiment was conducted about one month after Experiment I, employing two of the previous subjects (P and S). This experiment was conducted upon eye movement tracking equipment designed and constructed by John Gould [Gould and Schaffer, 1965]. It mixes a television camera input image of the presented visual stimuli (line drawing) with a camera input of a beam of light reflected off the corneal lens of the subject's left eye. This mixing produces a video tape protocol of a light spot moving about the presented visual scene. Due to the bulging of the eye at the lens region, and the geometrical optics of reflection, this light spot, within an error of 3 degrees, approximates the current eye fixation point upon the presented line drawing.

This experiment had two parts, the instructions to Part I being given in Figure II.7. Figure II.8 presents pictures of the stimuli used in this part. Each picture was flashed for approximately one-quarter to one-third second, this time being that found to be the basic (average) fixation time of the eye during the scanning of a presented visual stimuli (reference).

The verbal description given by the subject was recorded and, together with the drawing done, was kept as important data, as discussed later. The eye movement video tape record was also kept, to be utilized as a check to insure only a single eye fixation during the drawing presentation.

The instructions for Part II are given in Figure II.9. Figure II.10 presents pictures

28a

EXPERIMENT II

PART I

Figure II.7

This part, and Part II to follow, will be conducted on this eye movement tracking apparatus.

Instructions

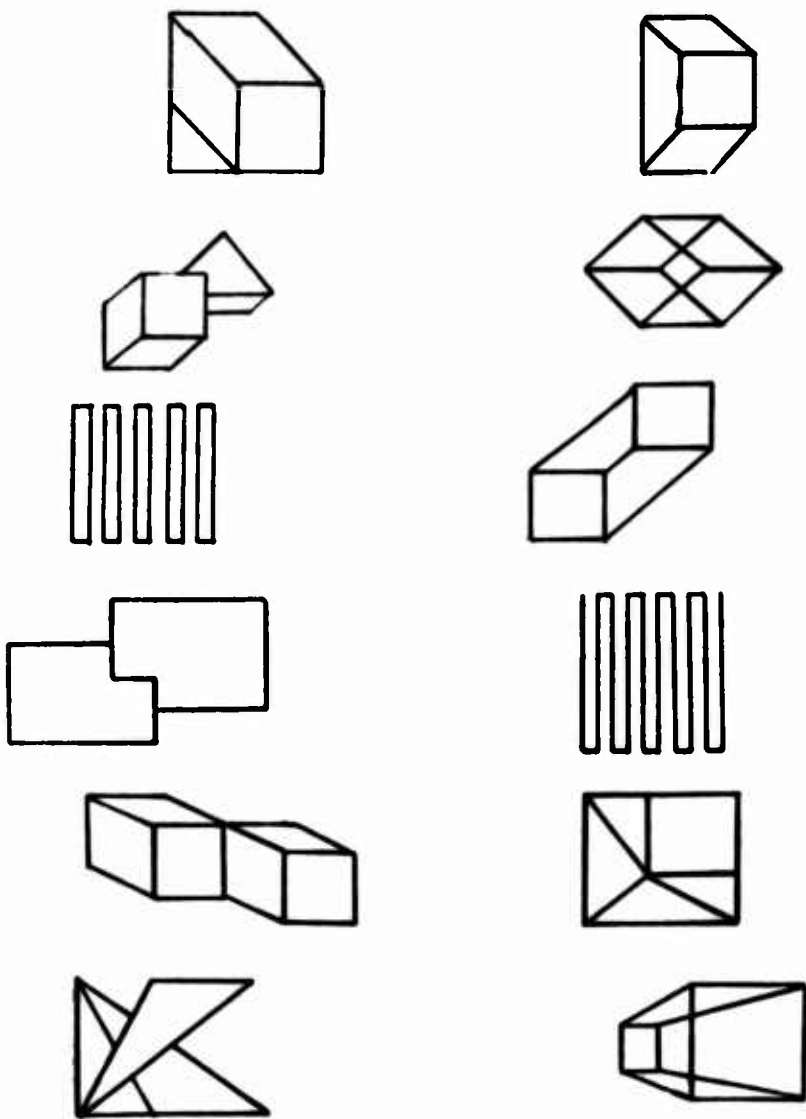
In this part you will be presented with a series of stimuli similar in type to those presented in the earlier experiments. Before each line of drawing is presented, a small cross will be drawn upon the screen in front of you. Once upon the bite plate, you are to look at (fixate) the intersection point of the cross. I will then make a final zeroing adjustment of the eye tracking device. Then while you continue to fixate the cross intersection, the line drawing stimulus will be presented upon the screen. You are to constantly fixate the cross intersection throughout the picture presentation. Never move your eyes. The eye tracking device will note any movement, thus invalidating that trial.

As soon as the picture is removed you are to get off the bite plate and describe orally what you recall about the presented drawing, while sketching those parts that you can.

Each picture will be presented for a timed interval. Times will vary.

28b

Figure II.8



28c

EXPERIMENT II
PART II

Figure II.9

Instructions

You will now be presented with a series of partial line drawings. You should consider each to be a small part of a larger line drawing similar to one of those seen in the previous part of the experiment, as if one were looking at such a drawing through a tube which allowed only a small, partial view of the drawing at any time. While viewing the picture you are to orally do as follows:

1. describe what the partial picture is itself
2. give any ideas you have about the whole, of which the given partial picture could be part

I will do one example for clarification.

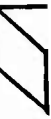
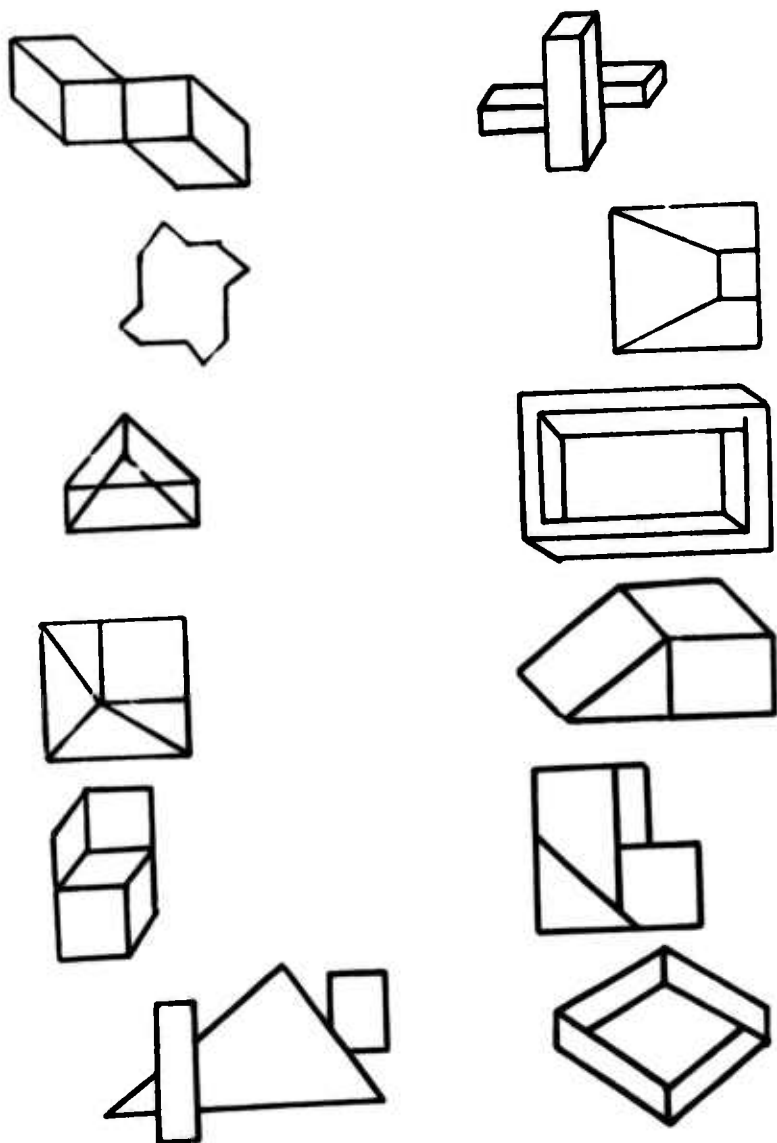


Figure 11.10



of those line drawing stimuli used. Note that several of these are repeats from Part III of the prior experiment. The given description of the presented picture stimuli was audio-taped and the resulting drawing was kept. The eye-movement video tape protocol was kept, also. There were ten pictures presented in this part to each subject.

STIMULI CHARACTERISTICS

A note should be made here concerning stimuli characteristics for both experiments I and II, and the subsequent effects of these characteristics upon the analysis procedure and the proposed image memory. The two important stimuli characteristics to be discussed are

1. the overall line drawing size
2. the line orientations within the drawings

All of the line drawings presented as stimuli covered comparable areas within the subject's visual field. At the distances presented, the line drawings in both experiments spanned twenty-five to thirty-five degrees of visual angle between their extreme points. This insured the necessary and useful employment of eye movements in Experiment II. The inter-vertex distances (contained line segment lengths) of all the presented pictures were within common bounds, and thus of comparable lengths. In Part III, of Experiment I, only one vertex could be seen at any one time through the hole (which spanned approximately 3 degrees of visual angle in diameter).

The directly obtained experimental data indicates that the consistency of overall picture size resulted in the subjects' use of a constant absolute positional reference frame, as had been expected. The bounds on the contained line segment's length and

their regularity in length also enabled the subjects to employ a constant absolute distance reference set also. Both of these absolute reference frames have a characteristic "fuzzy" correspondence to regions or neighborhoods of pictorial positions and lengths (distances). (This will be further discussed and clarified in Chapter III.)

The second important stimulus characteristic is that only four basic line orientations were utilized, these being: vertical, horizontal, upper-left to lower-right diagonal, and lower-left to upper-right diagonal. With few exceptions, no line drawing contained two or more diagonals of the same basic orientation with differing slopes.

The relevant experimental data indicates that this stimulus characteristic led the subjects to classify (perceive), and subsequently remember, line orientations in terms of just these four basic orientations, yielding eight possible line traversing directions (up - UP, down - DO, left - LE, right - RT, down-left - DL, up-left - UL, down-right - DR, up-right - UR).

The exceptions (See Figures II 2-11 and II 6-7) provide anecdotal evidence in favor of these inferred direction encodings. The subjects encountered, and verbally noted, some processing difficulty, due to the unusual existence of the different-sloped, same-directed diagonals within a single presented stimulus picture. Also subjects never mentioned verbally any degree of slope for the lines, normally using one of the eight proposed directions when verbally noting movement direction in protocol verbalizations or line orientation within description.

Previous research (statistical research) has shown that slope judgments are not distributed uniformly, but tend to cluster about several common reference orientations (references). This seems to indicate that humans in normal visual perception situations

do remember directions in terms of a relatively small, finite set of values (though, in more general environments, possibly more than the eight proposed). This clustering of human orientation perceptions has been similarly noted as partial justification for the representation of a block world with shadows in terms of the same eight basic directions (Waltz, 1972). Waltz does not claim a psychological model otherwise.

2. Analysis

DATA TRANSCRIPTION

The data, which was acquired from the experiments discussed above, consists of audio tapes of vertex, partial-picture, and complete-picture verbal descriptions, video-tapes, which included simultaneous audio recording of verbalizations, of the subjects' performances of the hole moving task, video tapes of subjects' drawing sequences of recalled line drawings, video-tapes of eye movement scanning sequences, plus actual hard-copy drawings produced by the subjects' through recall of presented stimuli. This is not the usual statistical data gathered in many psychological experiments. Thus, how this data has been used to generate the proposed visual perception theory and the program-implemented visual information processing system requires explanation.

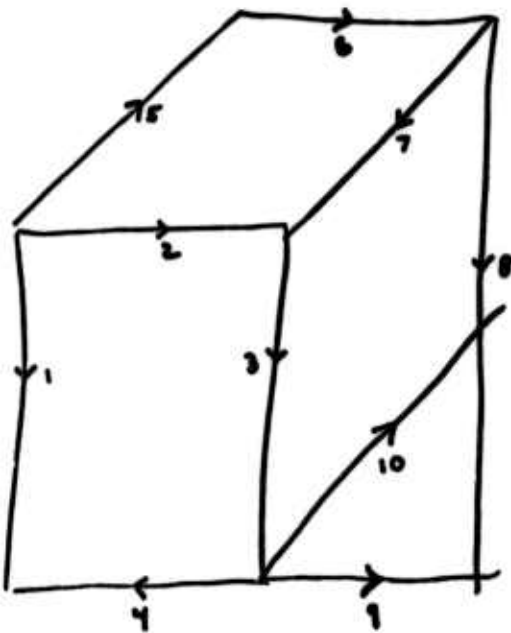
The audio and video tape recordings were first transcribed into suitable forms for analysis. These written or graphic transcriptions became the data actually considered as the bases for system specification.

The data of Experiment I, Part I, was audio recordings of complete line drawing descriptions and video recordings of the drawing activity of subjects reproducing the line drawings. The transcribed form of this data is illustrated by Figure II.11. The

Figure II.11

I.12.5
VTIME=5 sec.

okay
now there's a uh
it looks
basically the shape of a
parallelo
rectangualr paralleliped except in
down in the lower right hand corner
an extra
an extra triangle has been
has been drawn on



11b

Figure II.12

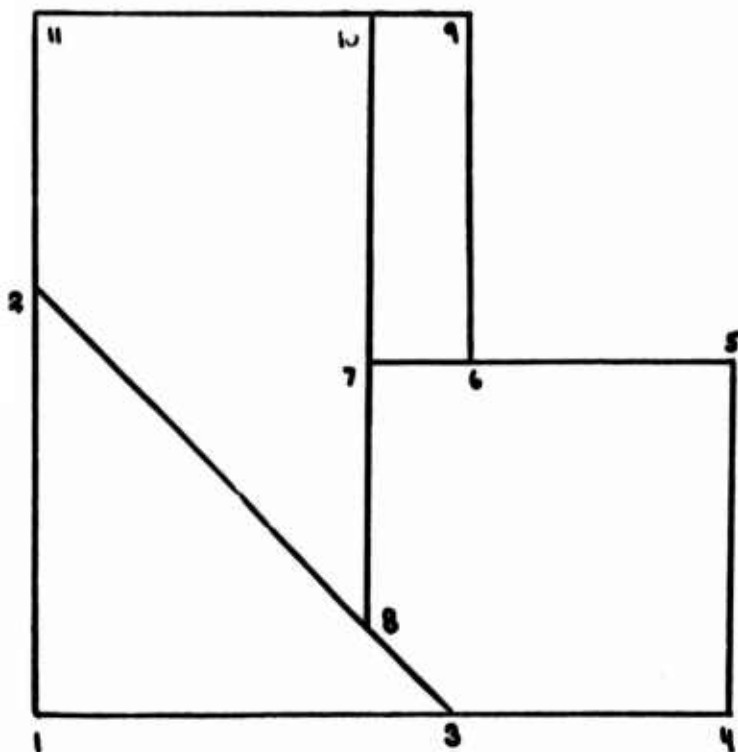
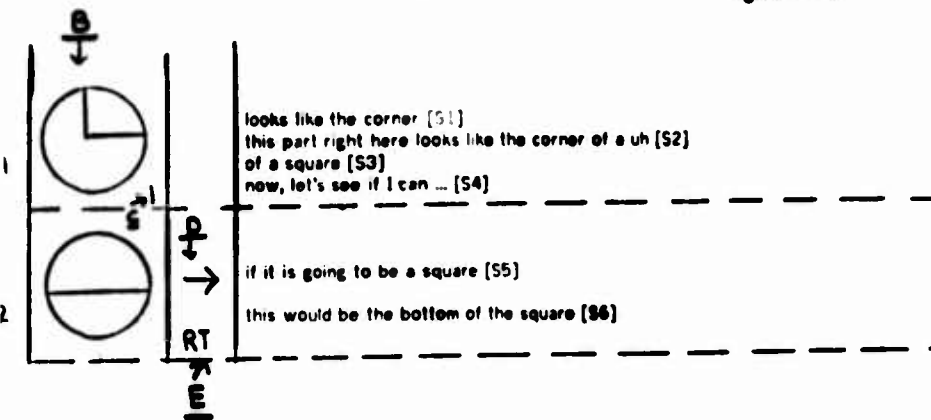
II.8.5

it's a
it's a right angle with a line bisecting it
and
uh
could be part of a
two-dimensional projection of a cube
or could just be
um
a line bisecting a square or
um
a diagonal of a square



31c

Figure II.13



audio data is transcribed in a straight-forward manner. Each hesitation in speech is noted by the commencement of a new line of transcription. Any significant pause (over 2 sec.) is noted by a "(pause)". VTIME is the approximate viewing time in seconds for that subject on that picture.

The actual, hard-copy line drawing reproduction executed by the subject served as a framework for the transcription of the drawing behavior. Numbers and arrows were appropriately added to indicate the order and direction of the drawing strokes observed. Various features of this data contributed to the specification of the image representation of VIPS. The visual concepts and interrelationships between them which occurred in the verbal descriptions were most important. Interactions between the verbal and drawing behaviors were others.

Part II of Experiment I was primarily of a training and screening nature. The data obtained were verbal reports based upon partial (single vertex) picture information. These verbalizations were transcribed in a manner analogous to that of Part I, as illustrated in Figure II.12. The data was marginally valuable for both image and processor proposals. The verbalizations provided indications of subjects' vertex classifications and the perceptual inferences made from such vertex classes.

The data of Experiment I, Part III, was in the form of video-tapes of subject hole-moving behavior, which included accompanying audio recording of subject verbalizations. Selected protocols have been transcribed as a series of informational frames. Two such frames are illustrated in Figure II.13. The element A (see Figure II.13) is the ordinal number of that frame in the protocol, in the form V_i , i an integer. Element B is the partial view which the subject sees through the hole at that point in the protocol, the circle being representative of the hole opening. Element C is

included only when the subject is at a vertex of the line drawing. It indicates which vertex is in view according to a numbering system defined by the first page of the transcription (see bottom of Figure II.13).

The second field of a transcription frame may be empty (as in V1) or may have elements D and E as specified contents. D is an arrow representing the direction of the hole movement occurring at that point in the protocol. E is the symbolic code used in VIPS to represent that direction. When no movement is occurring this field is empty. The third field is used for the transcription of verbalizations occurring in conjunction with the hole position and movement state defined by the first fields. These verbalizations are again transcribed analogously to those of Part I. Furthermore, each speech segment is labelled at its end by a statement number of the form "[Si]" (i an integer), facilitating its reference by the text. The complete verbal descriptions and reproductive drawings obtained were transcribed as for Part I, also. Figure II.14 presents a sequence of nine protocol frames to illustrate more fully the nature of the transcription and behavior.

The data from Experiment II, Part I, were audio recordings of partial picture descriptions and associated partial drawings. These resulted from the short, single fixation presentation of line drawing stimuli. The verbalizations were transcribed as usual, and the hard copy drawings paired to these. See Figure II.15 for an example of the transcribed data.

The data from Part II of Experiment II consisted of video tapes of eye movement protocols, audio tapes of subsequent verbal descriptions, and drawn reproductions. The eye movement sequences have been transcribed as in Figure II.16. Element A is the sequential fixation number, and B is the fixation duration in sixtieths of a second. The initial sequence segment is connected by dotted lines.

Figure II.14

			looks like the corner [S1] this part right here looks like the corner of a uh [S2] of a square [S3] now, let's see if I can ... [S4]
	→ RT		if it is going to be a square [S5] this would be the bottom of the square [S6]
			aha! [S7] looks like we might have a triangular object [S8] instead of a square [S9] this would be
	↗ UL		the second side of the triangle [S10] (slowly)
	↗ UL		(through vertex B)
	↗ UL		
			
	↓ DL		and this is the third side [S11]
			okay [S12]

11b

Figure 11.15

um
the figure as a whole was like
a square
it could be three dimensional
um (pause)
there was
a line coming from the right corner
going up pretty far
like a diagonal across it
and then up here there was a very
(laugh)
there was a bunch of lines
but I
can't recall which way they were going
or what types of shapes they formed
I think there was a triangle
or something up there

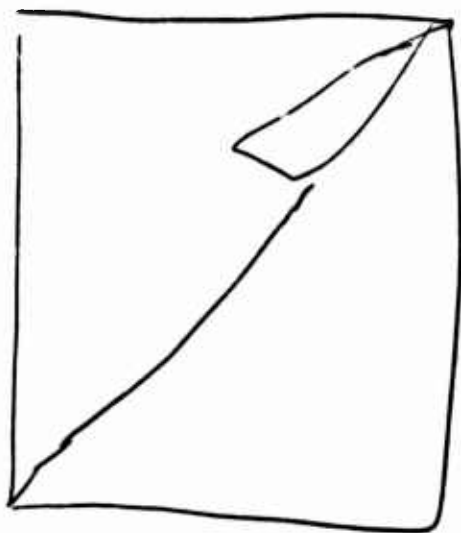
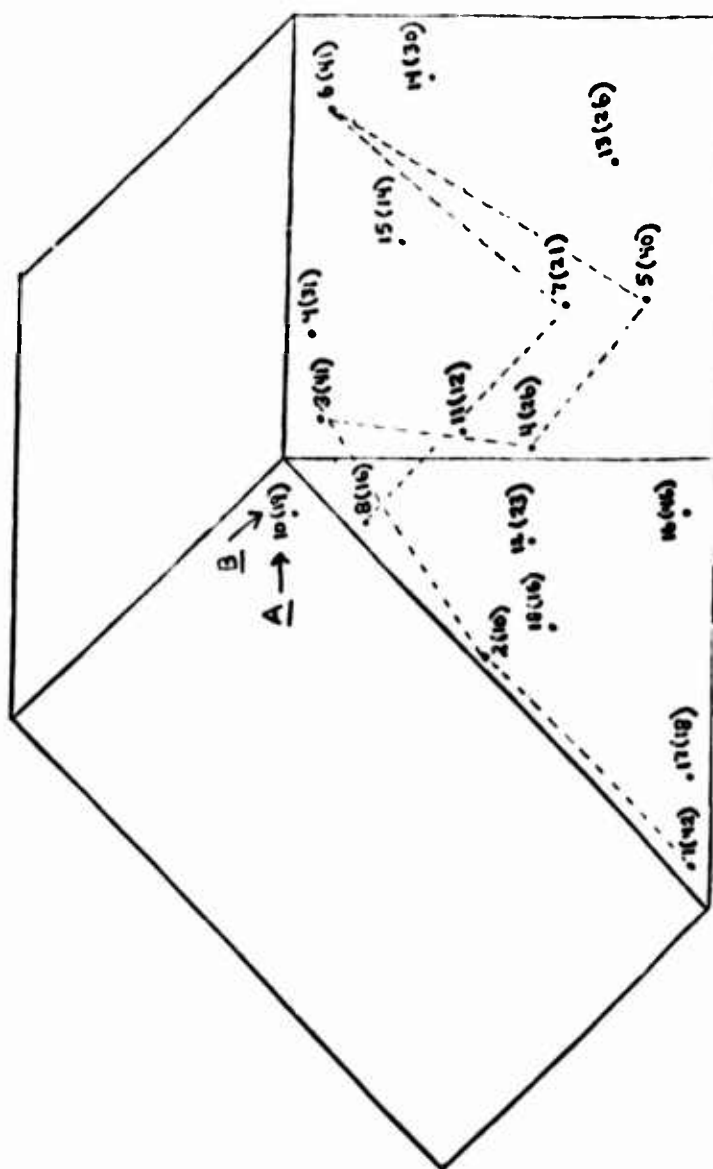


Figure II.16



STEVE (S)

$\sim 1^\circ$

S4

TRANSCRIPTION USE

The verbal descriptions from memory of completely perceived line drawings that were obtained through Experiment I, parts I and III, and Experiment II, Part II, served as a universe of information for the inference of the image memory yielding such descriptions. The types of things (concepts) remembered, the relationships used between remembered picture specifics and concepts, picture characteristics not noted, the types of errors occurring, and the sequence of the verbal descriptions were all used as inferential indications of the content and structure of the underlying image representation.

Something must be said concerning the relevance of the verbal data to the specification of the visual image and perception system. The protocol verbalizations indicate internal states (goals, confirmations) and as such do form a reasonable basis for process rule specification. These verbalizations are not introspective assertions as to the nature and course of activity. The visual image which is constructed by the perceptual activity serves as the semantic basis for the subsequent verbal description. As such, the verbal description indicates information that must be contained in or readily derived from that visual image.

These same experiment parts also provided a corresponding group of line drawings performed by recall, plus video tape recordings of the drawing sequences. The drawing sequence itself, drawing hesitations and false pen positionings, and the final drawings in terms of errors, consistent alterations, and correctness all were again considered effective bases for inference concerning the information and structure of the complete image memory representation of line drawings.

The behavior protocols obtained in Experiment I, Part III also provided information relevant to the image representation design. The chosen representation must naturally facilitate realization of the perceptual inferences, image alterations, and input integrations which occur in the observed behavior.

The behavior modelled by VIPS is that of the hole-moving protocols from Experiment I, Part III. As such, the transcribed video tape protocols served as the primary basis for the specific definition and implementation of the perceptual processes of VIPS. The implementation reflects the desired extendability of both the proposed processes and image representation to a more general domain of visual scene input and the desired corelation (extendability) to the normal eye movement situation (including peripheral vision considerations).

All of the protocols obtained from the experiment were considered at a general, qualitative, anecdotal level first. This analysis gave rise to some general characteristics of the perceptual processing system, plus the information relevant to image memory design as discussed above. This overall consideration was instrumental in the proposal of four component processes (see Chapter 4). Then three protocols for each subject were completely transcribed and analyzed in depth to further instantiate the constituent perceptual processes in terms of control and operations necessary. Finally, processes were implemented to account specifically for four of these. Figure II-17 indicates the degree of analytic consideration given to each obtained protocol.

The protocol data was considered to be two mutually interacting groups of information. One group is the hole-movement sequence itself, in terms of location (view), movement, and general speed and time factors. The other body of information

Part III, Experiment I Data Analysis

Stimulus Subject										
B	T	-	T	T	N	N	N	N	N	N
G	I	N	T	T	N	N	N	N	N	N
P	-	I	T	T	N	N	N	N	N	N
S	I	I	-	T	N	N	N	N	N	N

15a

Figure II.17

I ~ implemented, thoroughly analyzed N ~ not transcribed, considered generally
T ~ transcribed, thoroughly analyzed - ~ behavior protocol not taken

was the verbalizations provided at various occasions by the subjects during the hole movement task performance. The verbal statements were used as indications of currently active perceptual goals, realized recognitions, and attended image memory information. The hole-movement sequences were analyzed to relate movement subsequences to these inferred strategic perceptual goals. Then each subsequence was further analyzed to infer its associated underlying control tactics. In the absence of verbalization, perceptual goals were inversely inferred from observed movement subsequences with control tactics which had elsewhere been related to a verbalized goal.

The analysis of the data obtained from Experiment II upon the eye-movement tracking device was mostly carried out at a qualitative level with the exception of a quantitative analysis of fixation durations. This data was analyzed in light of the perceptual system and image representation developed from the analysis of the hole-moving protocols. The data served as the basis for the extension of the implemented proposal, also.

Part I of Experiment II was designed to yield insight into the amount, content, and structure of visual information entered into image memory from one initial brief fixation upon a presented line drawing. It is assumed that the primary informational difference between the positioning of the hole and the positioning of the eye at a picture location is the absence of peripheral visual information in the hole positioning case. It is noted that a completely different motor system for positioning is involved in each case. These are taken to be functionally equivalent.

The eye movement sequence protocols obtained from Part II of Experiment II (two subjects) were all studied at a general level to note any regularities of behavior

possibl
being
been th
of fixa
duratio
each fi
of pro
Un
consid
inform
the th
sizes,
guides
produ
the pr
Thus,
simula
simula
proce
Ti
and a
theor

possibly appearing. Eight eye movement protocols were then fully transcribed, these being upon four pictures for which hole-movement protocols by the same subjects had been thoroughly transcribed and analyzed. Each sequence is transcribed as a sequence of fixations. Each fixation is represented in terms of location and duration. Fixation duration is assumed to be directly related to the amount of processing occurring at each fixation. The sequence of fixation locations provide indications for the inference of processing strategies and states.

Underlying the development of the proposed processing model was a general consideration and an overall amalgamation of as much relevant psychological information as possible. Some of these results are appropriately referenced throughout the thesis. Findings concerning short term and long term memories, their existences, sizes, overall retrieval structures, and operating characteristics were used as design guides. So were ideas concerning the representation of cognitive processes in a production system form [Newell and Simon, 1972]. Together all of this external input to the proposal furnished a valuable and necessary basis for overall system character. Thus, the implemented system is not merely a computer program which accomplishes a simulation of a body of observed human behavior. The system further realizes this simulation through the embodiment and utilization of the functional units of the human processor, as proposed from the results of recent research in cognitive psychology.

This chapter has provided some idea of what observed behavior, transcribed data, and analysis procedure has led to the proposed imagery and perceptual system and theory. The following chapters are an exposition of that proposal.

Chapter 13 — The Image Representation

This chapter is an exposition of the symbolic image representation used in VIPS to embody visual information. Criteria which served to determine the form of the representation are discussed. This is followed by a description of the representation. Finally, how the image representation appears in the memories of the cognitive (perceptual) processor is briefly defined.

1. Image Representation Criteria

Throughout the development of the proposed image memory representation several guiding design criteria were applied. Possession of these critical characteristic abilities by the proposal were deemed to be necessary for an adequate definition. These necessary properties are:

1. VERBAL DESCRIPTION ADEQUACY - The ability to realize representations which not only contain the information necessary to yield the complete picture descriptions obtained as data in Experiments I and II, but also embody this information in symbolic structures bearing straight-forward correspondence to those perceptual units (i.e. objects, other concepts) and relationships used in those descriptions.

2. DRAWING BEHAVIOR ADEQUACY - The ability to realize representations which are sufficient as a source of (control) information for the task of drawing from memory, and which can be so used in a straight-forward manner to yield the complete picture drawing sequences obtained as data in the experiments;

3. PERCEPTUAL ACTIVITY ADEQUACY - The ability to realize partial picture representations which can, in a straight-forward manner, be dynamically (in time) augmented (added to) or altered, and which furthermore can serve as a source of information for object recognition, goal determination, and hole (eye) movement control throughout perceptual behavior;

4. SEMANTIC CHARACTER - The ability to allow for the coexistence of the operationally temporal (sequential) and representationally spatial characteristics of visual imagery (memory) in a straight-forward manner;

5. INTERNAL MODALITY CHARACTER - The ability to be accessed by means bearing straight-forward (isomorphic) correspondence to that used to access external visual information;

6. EXTENDABILITY - The ability to be extended to realize representations of all visual perceptions of form in a straight-forward manner.

The word "straight-forward" (the phrase "in a straight-forward manner") has been consistently used above. The word (phrase) is to be understood to mean, "with no, or a minimum of, additional processing"; where processing could be derivation, calculation or re-representation.

Satisfaction of these design criteria insures a visual image representation which is consistent with the image definition of Chapter 11. The first three criteria correspond to the image's definitional ability to serve as guide for visually-related behavior. The fourth criteria reflects the necessary semantic character of the visual image's information representation. Satisfaction of this requirement insures adequate independence of image content from the sequential course of human behavior, yet allows the image to be a participating factor in that behavior.

The fifth criteria is taken directly from that previous image definition. This requirement represents the internal modality characteristic of the visual image. It influences both representation structure and the processes which access this structure. The last necessary ability is the result of the specific nature of the image representation developed here. It is limited to the visual environment of straight line drawings. The afore, the proposal clearly must be readily extendable to be of value.

These criteria for design of the visual image representation reflect the necessary and functional explanatory adequacy required of the image as a part of a scientific theory of perception and imagery. The representation must be a functional component of a system capable of providing mechanistic explanations of empirical evidence. VIPS employs the image representation in the reproduction (simulation) of obtained protocol (behavior) sequences.

In the above criteria there is no mention of introspective suggestion or confirmation as being a necessary characteristic of the proposed image representation. No aspect need be suggested or confirmed by consideration of conscious experience. Pylyshyn cautions that:

"What is available to conscious inspection may not be what plays the important causal role in psychological process." (Pylyshyn, 1973)

Furthermore, conscious experience is not a direct reflection of the content of attended memory. It is rather an internal phenomenon created by the application of cognitive processes to the content of attended memory. The internal modality criterion is an indirect acknowledgement of the sensory - like nature of the conscious experience of (visual) imagery.

2. The Visual Image Representation

SYMBOLIC REPRESENTATION

The visual image representation of VIPS is symbolic in nature. A visual image is a structure of interrelated symbols. That information is represented symbolically by the human is not a new hypothesis in psychology, as indicated by the following statement:

"My hypothesis is then that thought models or parallels reality -- that its essential feature is not 'the mind', 'the self', 'sense data', nor propositions, but symbolism." (Craig, 1943, p 57)

The hypothesis that a visual image consists of symbols and relations implies the primacy of the abstract in perception and cognition. Without an existing symbolic (abstract) framework for representation, there is only and always the everchanging flux, the "blooming, buzzing confusion" of William James. Cognition knows all only in terms of abstract symbols and relations. The primacy of the abstract (symbolic) is indicated by our failure to find a physical measure which is directly related to any domain of perceptual qualities (symbols) and by the relativity of all perceptual qualities which implicitly makes such a finding impossible.

A computer may be described as a symbol processor. As such, it provides a means for testing and evaluating the symbolic representation hypothesis. VIPS investigates the use of symbolic representation and processing as an explanation for visual imagery and perception. Recent research has shown the efficacy of symbolic representation in the computer - implemented explanations of human behavior on tasks involving visual information (Baylor, 1972; Newell, 1972; Moran, 1973). Moran presents an extensive

survey of recent proposals for the form and content of the symbolic representation of visual information.

IMAGE CHUNKS

The basic unit of visually meaningful image information within VIPS is the chunk. A visual image is an interrelated group (structure) of image chunks. A chunk is an interrelated group of symbolic information elements. Access to one element of a chunk implies the immediate accessibility of all other contained elements and the information they embody. There is considerable experimental evidence which is favorable to the proposal of information chunking in human memory (Miller, 1956; Newell and Simon, 1972). Transcribed data from the thesis experiments provide directly acquired, inferential evidence for the chunking of visual information. The relevant data factor is the consistent use by the subjects of a small (finite) set of visual concept types. Each concept type has an associated class of symbolic information structures which embody the representative spatial configurations of visual features.

This concept-based chunking of visual information influences all of the observed forms of visually-related behavior. During perceptual processing, it influences the determination of goals and expectations, is the unit of visual hypotheses, and accordingly affects the scanning (eye or hole) of the line drawing stimulus. The use of visual concepts as the basis for partitioning (chunking) the representation of the image information is most directly evident by their use in the verbal recall descriptions and protocol verbalizations. The influence extends to the reproductive drawing behavior observed, being a primary factor in determining the sequence of drawing strokes used by the subjects. Bartlett (1917) reports similar results from a recall and reproduction task.

The role which image chunking plays in the perceptual processing is discussed and illustrated in Chapter 14. Appendix A Concept now presents three groups of data from Part I of Experiment 1. These illustrate the effect which information chunking within the visual image has upon the observed verbal and drawing recall behavior. They also illustrate the use of visual concepts as the basis for image chunking. Two visual concept types are also introduced.

Group 11 introduces the use of the object type visual concept (OBJECT chunk). Both subjects use the "cross" and "triangle" object concepts to embody the visual information as shown by the verbal descriptions. The drawing sequences further reflect the use of these visual concepts. Subsequences of drawing strokes correspond to the partitionings of the image chunks. Both drawn reproductions contain errors, however. This indicates the inability of object type concepts alone to represent a picture consisting of several spatially interrelated (interacting) objects.

Group 12 again illustrates the use of the object type image concept. Furthermore, 12 P and 12 S introduce the use of the line type visual image concept (LINE chunk). These three examples of Group 12 provide more evidence of chunking's effect upon the reproductive drawing behavior. Each verbal description reflects a different conceptual partitioning of the visual information. The differences are apparent in the drawing behavior, as subsequences of strokes show correspondence to the visual concepts of the related verbal description.

Group 13 illustrates the important point that to chunk visual information according to an object type concept does not necessarily imply the naming of that object. Neither subject has a name for the object type image concept used to embody the visual information appearing on either side of the center rectangle. The primary

function of the choice of concept type used to chunk any visual information is to determine basic characteristics of the symbolic structure which is to represent that information. This operation is more fundamental than (not equivalent to) verbally labelling (naming) that visual information.

Other experimental data indicate three other visual concept types which are used as chunk types in VIPS to create the visual images of straight line drawings. There are the vertex, side, and face (in conjunction with three-dimensional projection images) type visual concepts. Thus, there are five basic (corresponding) image chunk types used in VIPS, these being VERTEX, OBJECT, LINE, SIDE, and FACE.

IMAGE CHUNK CHARACTERISTICS

There are three classes of elements which are found in every image chunk of VIPS. There is one Chunk Header element, several Image Body elements, and one or more Position elements (see Figure III.1). Each symbolic element is a group of properties which interrelate the elements themselves and also relate the element to the appropriate visually meaningful symbols.

The Chunk Header element serves to indicate the type of visual concept embodied by the image chunk. This type is indicated by the value of the element's associated TYPE property (see Figure III.2). Another associated property is CR (Current Reference) which serves to reference (afford access to) one Image Body element of the chunk's image structure. Depending upon the chunk type, other properties are associated with the Chunk Header element, such as the NAME property of an OBJECT type chunk's Chunk Header element (see Figure III.2).

The Image Body elements are the constituent symbolic elements of the semantic

VIPS

The pos.

CHUN

IMAC

POS1

Figure III.1

A VIPS Image Chunk consists of:

- One CHUNK HEADER element
- Several IMAGE BODY elements
- One or more POSITION element(s)

The possible symbols for each type are:

CHUNK HEADER := V1, O1, L1, S1, F1
IMAGE BODY := X1, A1, I1, E1, Q1
POSITION := P1

Chunk Header Elements

VERTEX <V1>		OBJECT <O1>		LINE <L1>		SIDE <S1>		FACE <F1>	
TYPE	VERTEX	TYPE	OBJECT	TYPE	LINE	TYPE	SIDE	TYPE	FACE
SPEC	<S>	NAME	<N>	(or ILINE)		SPEC	<S> <S>	CR	<F>
VSPEC	<VS>	DIM	<D>	SPEC	"(L1) <L1>"	CR	<F>	OBJECT	<O1>
CR	<CR>	CR	<CR>	CR	<CR>				
		NUMS	<N>						

- <S> := { V2, TE, MU, RT, KR }
- <F> := { an image body element }
- <L> := { UP, UR, RT, OR, DO, DL, LE, UL }
- <S> := { TWO, THREE }
- <N> := { TRAPEZOID, SQUARE-OR-RECTANGLE, SQUARE, RECTANGLE, TRIANGLE }
- <VS> := { vertex specification varies with the value of <S> <SPEC> }
- <L> := { TWO, THREE, FOUR, FIVE, SIX, SEVEN, EIGHT }

Figure III.2

structure embodied by the image chunk. This symbolic structure is called the "image body" of an image chunk. As noted above, each chunk type has a characteristic class of image body structures associated with it. Five types of Image Body elements are used within VIPS to realize these differing classes of information structures. These types are XIT, ANGLE, INTERNAL, END, and QUICKSEE. Each element type is distinguished by the properties which may be associated with it (see Figure III.3). These properties determine a different role for each element type within the image body information structure.

Each type of Image Body element has both mandatory and optional properties associated with it. The mandatory properties are specified for every given instance of that element type within an image body. The optional properties are indicated by a star, "*", in Figure III.3. Which further optional properties are specified for a given element instance determines the value of that instance's mandatory TYPE property. For example, an instance of a XIT type element may have TYPE value UEX (Unknown EXit) when only the mandatory properties are specified, or may have TYPE value KEX (Known EXit) when both optional properties RVX (Range Vertex exit) and LVX (Link Vertex exit) are specified. See appendix A.LTM for the expansion of the mnemonics.

Before further describing the symbolic image structures which VIPS creates with these different Image Body elements, the last class of image chunk elements, the Position type element, requires definition. The experimental results yield anecdotal evidence for the proposal that the positions of picture (visual) features are remembered in terms of an imprecise ("fuzzy") absolute locational coordinate scale which becomes associated with the portion of the visual field of interest (in this case, the whole line drawing stimulus). Subjects consistently use generalized positional

Figure III.3

TYPE	XIT <Xi>	ANGLE <Ai>	INTERNAL <Ii>	END <Ei>	QUICKSEE <Qi>
A	TYPE <Xi>	TYPE <Ai>	TYPE <Ii>	TYPE <Ei>	TYPE QUICKSEE
S	DIX <Xi>	ANG <Ai>	DIX <Ii>	LS1 <Ei>	NUMX <Qi>
S	DIV1 <Xi>	DIV1 <Ai>	RIX <Ii>	LS2 <Ei>	TD1 <Qi>
O	DIV2 <Xi>	D <Ai>	LIX <Ii>	VAL(LS2) <Ei>	TD2 <Qi>
A	VAL(DIV2) <Xi>	VAL(DIV2) <Ai>	DIV2 <Ii>	VAL(LS2) <Ei>	VAL(TD2) <Qi>
T	VAL(DIV2) <Xi>	VAL(DIV2) <Ai>	VAL(DIV2) <Ii>	LXL * <Ei>	VAL(TD2) <Qi>
E	LIX * <Xi>	EAL * <Ai>	ESL * <Ii>	LXX * <Ei>	PNUM <Qi>
P	RIX * <Xi>	PNUM <Ai>	LEL * <Ii>	SEL * <Ei>	
R	DAL * <Xi>		OCL * <Ii>	PNUM <Ei>	
O	ALL * <Ei>		EFL * <Ei>		
P	LXL * <Ei>		EOL * <Ei>		
E	PNUM <Ei>		PNUM <Ei>		

<Xi> := {UEX, ABX, KEX, XUX, KUX}

<Ai> := {KIA, KEA, KRA}

<Ii> := {INT, ILE, INS, IOC, INO, INF}

<Ei> := {SID, LEI, LEX}

<Qi> := {UP, UR, RT, DR, DO, DL, LE, UL}

<Xi> := {SH, ME, LO, LS, LM, LL}

<Ai> := {ACA, RTA, UBA, STA, SAA, SBA, SOA}

<Ii> := {ONE, TWO, THREE, FOUR, FIVE, SIX}

<Ei> := Position Type element

terms in both the verbal descriptions and the protocol verbalizations. The ability of the subjects to readily hypothesize line segment linkages between picture vertices provides further evidence. Moran (1973) has developed an image representation which does not include absolute positional information. He has subsequently noted an inherent difficulty in achieving similar spatially determined hypotheses with that representation.

A Position element embodies the "fuzzy" location of each of the Image Body elements which reference it by the PLOCM property. A Position element has two associated properties, LRP (Left - Right Position) and UDP (Up - Down Position). LRP and UDP each have seven possible symbolic values. The attended portion of the visual field is thus broken into forty-nine positional areas, as shown in Figure III-4. The size consistency of the line drawing stimuli has enabled the constant application of the locational areas to the same actual absolute picture areas in the implementation of VIPS. It is proposed that this grid of forty-nine picture areas may vary in size and symbolically, and be able to align with the properties of a scene (the visual field) to give a different strategy that the researcher may wish to apply to the positional grid itself. It may be applied to any content which is of interest.

Every Image Body type element references a Position type element in a visual field. In conjunction with a Position type element, and in the processing of the feature of the image representation given by the Position type element, the conscious experience of the contents of a visual field is somewhere taking position. The visual image is bound to locations within the perceptual space, which is embodied by the Position elements of the image representation. Perceptual space is not bound directly to the retina, but is bound to the area of interest in the environment. This indication

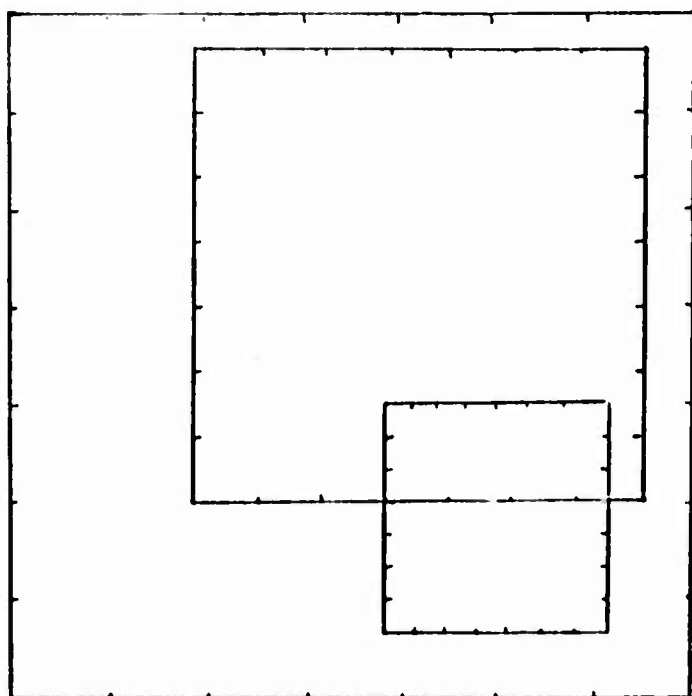
Positional Areas

Figure III.4

LLO (LRF) ULO (UDF)	LME ULO	LSH ULO	ORI ULO	RSH ULO	RME ULO	RLO ULO
LLO UME						
LLO USH						
LLO ORI			ORI ORI			
LLO DSH						
LLO DME						
LLO OLO						

46b

Figure III.5



is a basic factor in the maintenance of a stable visual world in spite of the varying retinal indications which occur during constructive perception.

In conjunction with this seven-by-seven reference grid of spatial position areas, VIPS employs six range or length codes. These are, in increasing order : SH (Short), ME (Medium), LO (Long), LS (Long - Short), LM (Long - Medium), or LL (Long - Long). The length of a line segment (or hole movement) is determined according to the number of positional areas traversed. For example, a line segment between adjacent areas is of length SH, while one from a corner area to another corner area is of length LL (very long).

VERTEX CHUNK

Each chunk type and its associated class of image body structures is best understood by the discussion of a representative example. As such, an example of a VERTEX chunk is presented in Figure III.6. As shown at the bottom of the figure, at one level the image chunk is simply a linear list of elements. The Chunk Header element (V1) occurs first in this linear list, facilitating access to the element by any attending process. The order of the remaining elements at this linear list level has no meaningful significance.

The image body of a VERTEX chunk consists of an equal number of interrelated XIT and ANGLE type elements. Each XIT element represents one line segment exit direction of the vertex (by the value of the DVX property). In Figure III.6, X2 represents the UP exit. Each XIT element is interrelated with two ANGLE elements by the values of its associated internal vertex direction properties DIV1 and DIV2. These properties have opposite directional values which are perpendicular to the value of

cx



✓1 X4 A3 P6 A5 X2

DX In Figure III.6, DIV1 of X2 is RT and DIV2 of X2 is LE. These two directional values themselves become properties of the XIT element, and each references an ANGLE element. In Figure III.6, X2 has properties RT and LE with RT referencing A3 and LE referencing A5.

An ANGLE element embodies the visual feature (angle) existing between two adjacent, different directed exit directions of a vertex. Chapter 12 has indicated that there are eight directional values (symbols) used in VIPS. Accordingly, the ANG property of an ANGLE element can have one of the seven possible angle symbols as its value. These angle symbols are ACA- acute angle, RTA- right angle, OBA- obtuse angle, STA- straight (180 degree) angle, SAA- straight acute angle, SRA- straight right (270 degree) angle, and SOA- straight obtuse angle.

An ANGLE element has two internal vertex directions associated with it, DIV1 and DIV2, which are inherited from the two XIT elements' properties which reference it. This is illustrated in Figure III.6. DIV1 of X2 is RT and RT references A3. DIV1 of X4 is UP and UP references A3. Accordingly, DIV1 of A3 is UP and UP references X2. DIV2 of A3 is RT and RT references X4. There is no significance to the choice of assignment of the two internal vertex directions to DIV1 and DIV2. The convention for the assignment of the internal vertex directions of an ANGLE element facilitates "traversal" of the image body structure.

IMAGE BODY ACCESS - "TRAVERSAL"

To access the symbolic structure of an image chunk's image body, a process must first gain access to the image chunk as a whole and then to one element of the image body (the means for this are discussed in Chapter 14). The process may then access

the contents of the image body structure by sequentially "traversing" it (from the initial element). Image body elements are accessed by traversal of available interrelating properties. The image body elements of a chunk are immediately accessible (chunk definition) but not accessed in parallel. An image body element is said to be "traversed" when it is so accessed by the active process.

With reference to the VERTEX chunk, a process may traverse the structure from one X1T element to another by using the same interrelating directional property twice. In Figure III.6, RT of X2 is A3 and RT of A3 is X4 or equivalently (RT of (RT of X2)) is X4. Note that a VERTEX chunk is completely traversable in two sequential courses (either X2 to A3 to X4 to A5 to X2 or X2 to A5 to X4 to A3 to X2). As such, the structure is a semantically determined, doubly-linked (doubly-directed) circular list. This type of structure will here be called a doubly-directed ring of elements.

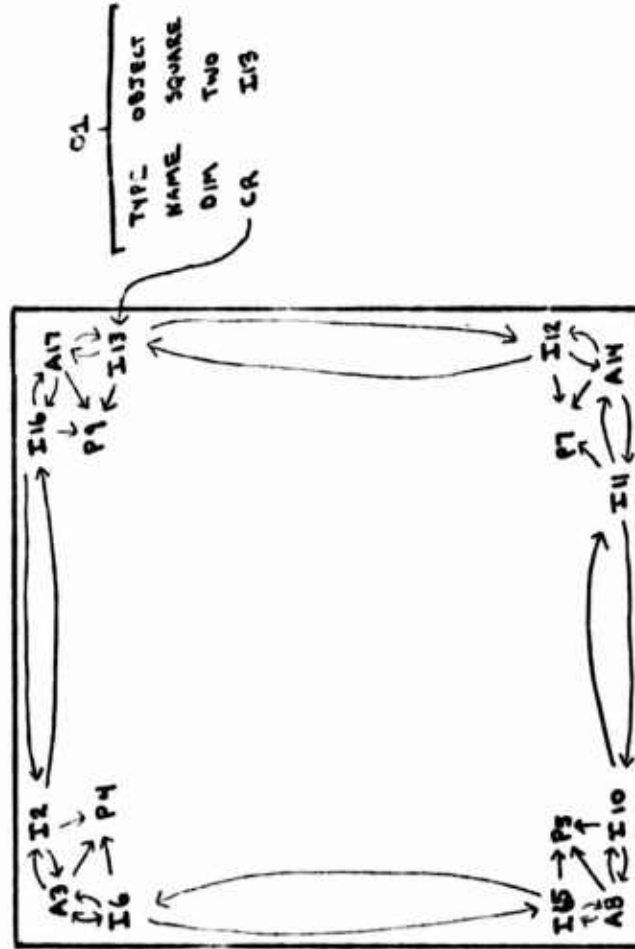
The image body structure of the VERTEX chunk represents the visual feature configuration at one spatial location. Thus, there is only one Position type element in the VERTEX type chunk. Each image body element references that element by its associated PNUM property (not shown in Figure III.6). A Position element differs from an image body element in that it is not (along with the Chunk Header element) a traversable element of the image structure of the chunk. A Position element is indirectly accessible from any image body element which references it by the mandatory PNUM property.

OBJECT CHUNK

An OBJECT type image chunk is schematically illustrated by Figure III.7. The image body of an OBJECT chunk is again a doubly - directed ring of elements. This

Figure III.7

AN OBJECT chunk



O1 I2 A3 P4 P5 I6 I10 I11 P7 P9 A8 I15 A4 I16 I12 I13 A17

time the elements are of types INTERNAL and ANGLE. The common configuration (non-hierarchical sub-structure) of image body elements which is used as the building block of the OBJECT chunk's image body is the "corner" configuration. The lower left "corner" of the Figure III. 7., is shown in detail by Figure III. 8. This configuration represents an object corner, but a corner is not an explicit, hierarchical sub-component of the OBJECT chunk image body structure.

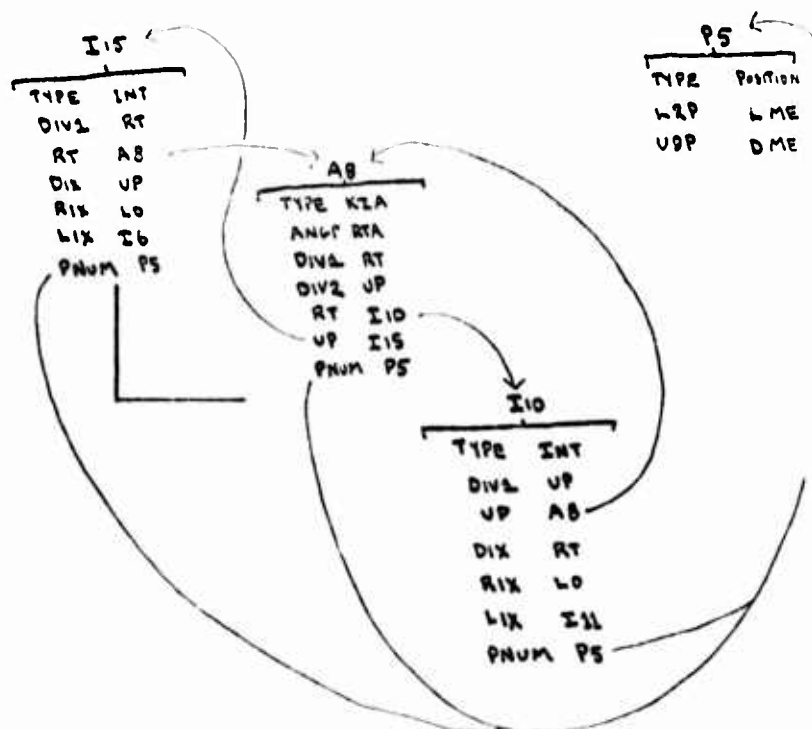
The corner configuration differs significantly from the image body of a VERTEX chunk described above. The angle exterior to the object's corner is not included in the corner's representation. An INTERNAL element has only one internal direction property (DIV1). A process traversing the OBJECT chunk's image body structure can not "look outside or around the corner" without accessing another image chunk. This characteristic of the image representation of an object causes a mode of visual image accessing which provides a basis for the experience of figure-ground phenomena (Koffka, 1935).

Each corner configuration represents visual features at one spatial location. Thus, an object chunk has as many partition elements as it has corner configurations in its image body structure (see Figure III.7). The ANGLE element inherits the value of its internal traversal directions (DIV1 and DIV2) from the INTERNAL elements of a corner in the same manner as from the XIT elements of a VERTEX chunk.

Two interrelated INTERNAL elements span the spatial separation of two adjacent object "corners" and as such are representative of an object "side" (which also is not an explicit image component). In Figure III.7, 16 references 115 by its LIX property. The direction (DIX of 16) is down (DO) and the range or length (RIX of 16) is long (LO). Inversely, 115 references 16 by its LIX (Link Internal eXit) property, with DIX (Direction Internal eXit) being UP and RIX (Range Internal eXit) being LO.

Figure III.8

"CORNER" CONFIGURATION (DETAIL)



The isomorphic relationship of a process's traversal of image body elements to the operation of external visual scanning and information accessing (criterion #5) becomes clear with the introduction of the OBJECT chunk. The contents of an object's image representation are accessed by the sequential traversal of the image body elements in a manner which is semantically equivalent to the observed hole movement behavior. The image representation and this operational relationship are straight-forwardly extendable to accommodate normal eye movement behavior (Chapter 15).

Furthermore, this traversal of the image body structures shows a close behavioral relationship to the activity involved in the generation of drawn reproductions. The non-hierarchical nature of each image body representation and the spatial meaning of the interrelational properties of the elements are two basic representation characteristics which make possible (1) straight-forward application of this behaviorally isomorphic traversal operation for image accessing. The definition of symbolic elements as property groups also facilitates traversal and significantly effects system operation by eliminating meaning from chunk surface structure (order of elements).

LINE CHUNK

A LINE type chunk is illustrated by Figure III-9. It is again a doubly directed ring of image body elements (a circular list structure). The LINE structure also introduces the use of the END and QUICKSEE types of Image Body elements. The body of a LINE type chunk includes two END type elements. Each END element references an appropriate POSITION element which serves to spatially locate that line end. Each END element has an associated property which serves to link that line end to the visual

A LINE chunk

51a

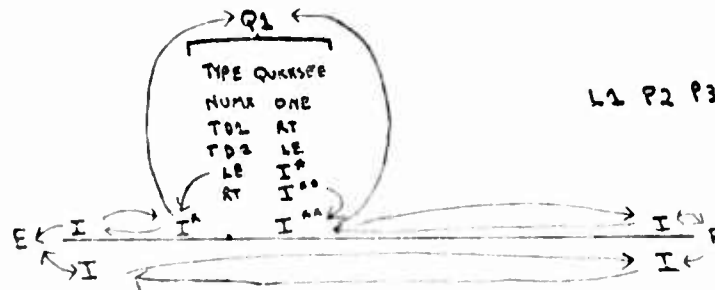
Figure III.9



L1 P2 P3 P4



L1 P2 P3 P4



context embodied in another image chunk. The TYPE property's value varies with that external context.

Each END element has two internal traversal directions associated with it (LS1 and LS2) showing traversal of each "side" of the line. Each LINE chunk "side" employs interrelated INTERNAL type elements to span spatial locations as in the OBJECT type chunk. At appropriate locations, vertex-like configurations are represented by either by interrelated XIT and ANGLE and INTERNAL elements (as shown in Figure III.9 (a)) or by a single QUICKSEE type element, as shown in Figure III.9 (b). A QUICKSEE element only partially represents the vertex information existing on that line "side". It only embodies the number of vertex exits leaving the line at that location (from that side). A QUICKSEE element does reference a position element by its associated PNUM property. Two traversal directions are also associated (VU1 and TD2), which are inherited from its neighboring INTERNAL elements.

The QUICKSEE element is introduced to represent the incomplete information acquired by a subject in the case that the hole is moved through a vertex on a line without stopping. This incompleteness of visual feature representation also typifies peripheral vision input information. As such, an extension (generalization) of the QUICKSEE element is later proposed as a means of symbolically representing such peripherally accessed visual information (Chapter 15).

ILINE

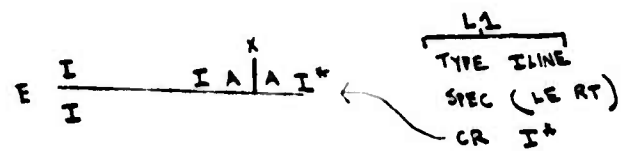
The ILINE chunk is an Incomplete LINE type chunk which may occur as part of a partial visual image existing at some intermediate point of the visual perception process. A chunk of this type is never part of the finalized visual image (perception).

THE ILINE CHUNK

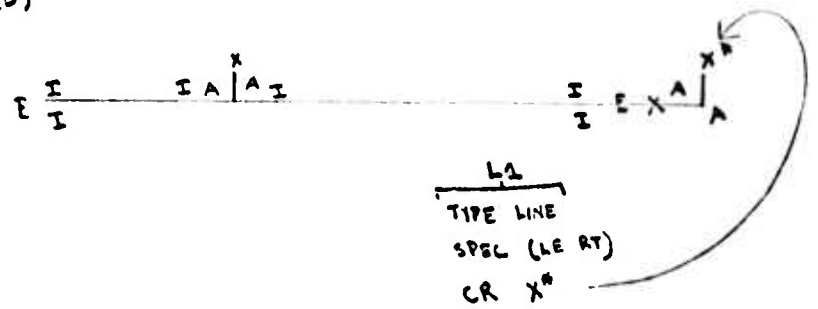
Figure III.10



(a)



(b)



which is the result of that perceptual activity. As such, it is not one of the five basic chunk types as listed earlier, but it is used by VIPS to accomplish perception.

Figure III 10(a) presents an illustration of an ILINE type chunk. This chunk type makes a definite specific use of the CR property of its Chunk Header element. CR indicates the INTERNAL element which is to be linked to elements representing the imminent visual information during its incorporation into and the completion of the image body (of the line). Figure III 10(b) illustrates the addition of further information completing the line and changing the ILINE chunk to a LINE chunk. Note that the center-like configuration is part of the resultant LINE type chunk.

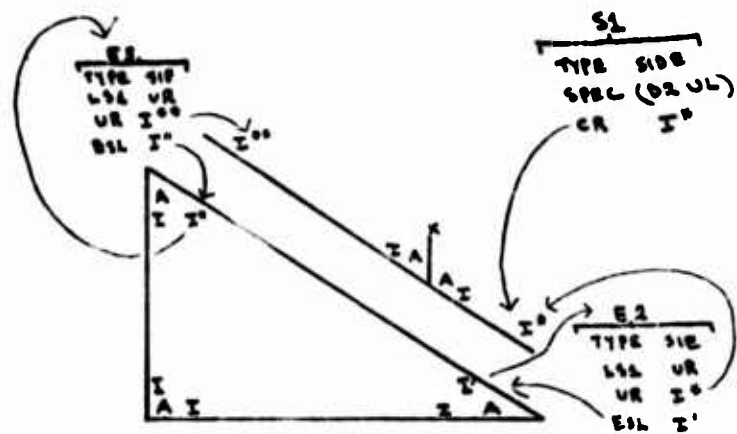
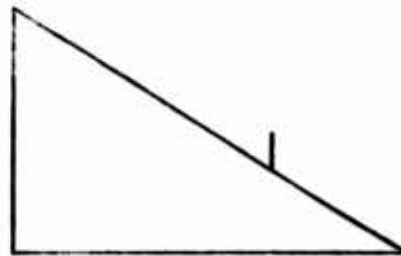
The CR property of the Chunk Header element is similarly used by the visual perception process in the construction of OBJECT type chunks. Incomplete OBJECT chunks are not made into such a special image type (eg. IOBJECT) though, as the object incompleteness is always embodied within the active perceptual goal. An ILINE chunk may occur in conjunction with goals which do not specifically note its existence.

SIDE CHUNK

A SIDE type chunk is used (occurs) only in conjunction with an associated OBJECT or FACE chunk. The relationship is illustrated in the schematic sketch of Figure III 11. This figure introduces a new pictorial representation convention. Each image body element is only specified by one of five letters indicating its type (I-INTERNAL, X-EXIT, A-ANGLE, E-END, Q-QUICKSEE). This abbreviated pictorial representation is used throughout the rest of the thesis. The "usual" intrachunk links (those defined previously) are implied to exist. Only links between image chunks are indicated (by the arrows). No Position elements are included in these drawn representations either.

Figure III.11

A SIDE
chunk



The END elements, when used in a SIDE chunk image body, each have only one intrachunk link available for traversal (LS1). Thus, a SIDE chunk allows traversal over only one "side" of the line-like image structure it represents. As with LINE chunks, a QUICKSEE element may occur in a SIDE chunk to partially represent an included (assimilated) vertex. It is the only image body structure which is not a doubly-directed list (ring) of elements. The other "side" (of the object's side) is traversable only within the related OBJECT chunk. The SIDE may represent visual information to the outside of the object as shown by Figure III.11. Or it may represent information occurring within the object contour (as shown by Figure III.19). Only a representation of the object contour is in an OBJECT chunk. Any contextual interaction with an object "side" is represented in a SIDE chunk.

FACE CHUNK

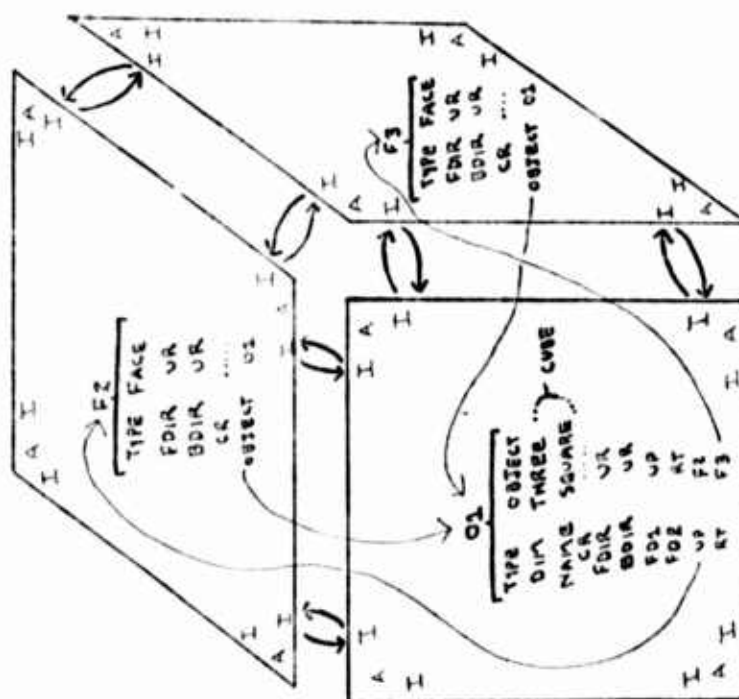
The FACE type image chunk has image chunk characteristics equivalent to the use of an OBJECT chunk. Chapter 1.5 (on extensions to the implemented VIPS) states that a three-dimensional projection of a prism-type object is visually perceived as consisting of a "prime face" and zero or more related "secondary" faces. This prime face is represented by an OBJECT type chunk, with the DIM (dimension) property of the Chunk Header specified as THREE. All secondary faces included in the image are represented in related FACE chunks. The Chunk Header element of a FACE chunk does not have a NAME property associated with it. The name of a three-dimensional object is determined by the name of the prime face chunk. Figure III.12 illustrates the use of FACE type chunks in the context of a cubic image.

The Chunk Header element of a FACE chunk and of a three-dimension OBJECT

THE FACE
CHUNK and
THREE DIM.
REPRESENTATION

544

Figure III.12



chunk has two special associated properties, BDIR and FDIR. FDIR specifies the facial extension direction of the prismatic projection and BDIR specifies the depth direction(s) associated with the object's perception (FDIR and BDIR are the same, except when FDIR is down - see Chapter 15). The image body of a FACE chunk is straight - forwardly generated from a specification of FDIR and the orientation of the line segment (of the prime face) from which it extends.

Perceptually created depth relations bring with them a third dimension for location (position) specification. As such, referenced Position type elements are modified by the addition of three new properties: PLRP, PUOP, PFBP (Perceived Front-Back Position). PLRP and PUOP assume values from the same class as LRP and UOP. Similar to the binding of a seven-by-seven two dimensional grid to the visual field of interest, the perceptual process binds a seven valued depth scale to the visual field (FLO (Front Long), FME, FSH, ORI, BSH (Back Short), BME, BLO). These symbols constitute the class of possible values for PFBP. The addition of the three new position element properties means that a group of visual features of a picture (i.e. vertex) is perceived to be at two interrelated locations. One location is in terms of the picture plane (LRP and UOP) and one is in terms of the perceptually-created three-dimensional space (PLRP, PUOP, PFBP).

The binding of this depth scale adds the third dimensional axis to the dimensional grid defined by the LRP and UOP binding. The resultant projectional plane (space) differs from the usual (mathematically defined) one in that the origin of the depth axis need not coincide with the two dimensional origin. Figure III.13 illustrates this idea, the depth origin being bound to the (LSH, USH) planar location. The back direction (BDIR) of that axis in Figure III.13 is up - right (UR). This direction plays a special role

55a

Figure III

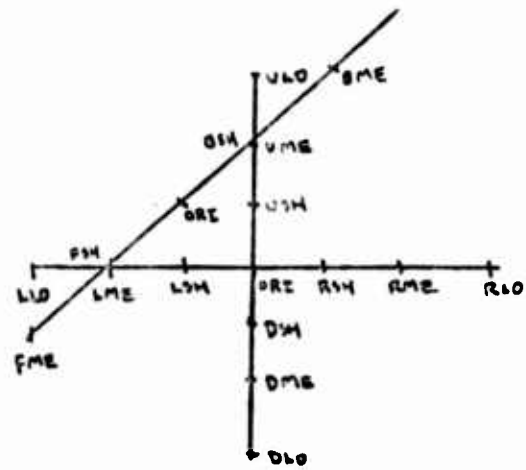
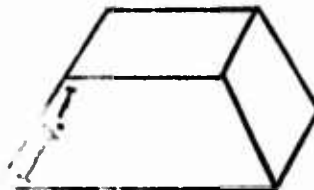


Figure III.14



in the perceptual assignment of locations (PLRP, PUOP, PFBP). Only line segments in the same orientation as the depth axis (UR - DL) may (not all are) be assigned depth meaning and accordingly affect perceived location. This necessarily adds a dimension of perceptual ambiguity (duality) to the picture plane.

Figure III.14 illustrates that not all line segments in the depth axis orientation are assigned depth meaning. Line segment "a" is bound to the prime (trapezoidal) face, which is perceived to be parallel to the picture plane. Figure III.15 illustrates the locational ambiguity of the perceptual projectional plane. Start at point "o". Moving to point "A" on line segment a (as part of the cube) is perceptually interpreted as a movement in depth (back). Moving from "o" to point "A" on side b then on side c (as part of the triangle) is a movement on the picture plane. Subjects did perceive this picture as interrelated cube and triangle (see Figure III.16). Their descriptions and drawing behavior indicate a duality of locational specification (PLRP, PUOP, PFBP and LRP, UOP).

Anecdotal experimental results led to and thus support these proposed property additions and their meaning. Three dimensional location specifications were used (i.e. "back there", "in front of that"). When traversing lines in the BDIR direction, subjects often would say "going back" or "now back here". Yet, perceptual inferences were made which indicated the maintenance of the picture plane location features in conjunction with the three-dimensional features used in the protocols verbalizations. Also, the verbal descriptions employed three-dimensional terms while the drawing behavior indicated picture plane positional knowledge. Several picture perceptions interspersed three and two dimensional objects, again interrelating two and three dimensional picture features (see Figure IV.15). Chapter 1.5 discusses three dimensional picture perception and relevant protocols.

563

Figure III.15

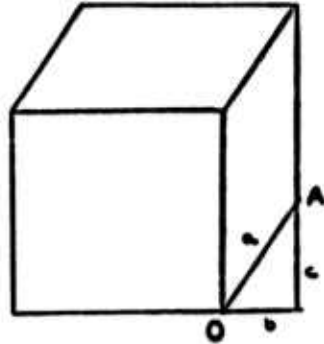


Figure III.16

I.12.S

VTIME=5 sec.

okay
now there's a uh
it looks
basically the shape of a
parallelo
rectangualr parallelpiped except in
down in the lower right hand corner
an extra
an extra triangle has been
has been drawn on

I.12.G

VTIME=

it's a rectangle in perspective and the bottom
of
bottom part on the right
is made into a triangle

INTERCHUNK RELATIONSHIPS

A visual image (in VIPS) is an interrelated group (structure) of image chunks. The chunk types have been defined. It remains to specify the interrelating property connections which are used to tie the chunks together as a complete visual image. Two basic types of links are used, equivalency links and space traversing links. These links meaningfully interrelate two image body elements, one in each of two image chunks.

Two varieties of equivalency links have been introduced (though not discussed) in the definitions of SIDE and FACE chunks above. Figure III.11 illustrates how an INTERNAL element of an OBJECT chunk may be equivalently linked to an END element of a SIDE chunk. Figure III.12 illustrates the use of equivalency links between INTERNALS of differing chunks of three-dimensional object image. These links serve to "strap" the image together.

Equivalency links serve to relate two image body elements of different chunks that represent the same physical picture component. This one physical (external) visual component is represented as being part of two different image concepts (chunks). The links of III.12 link INTERNALS which represent one line segment that is shared by two facial images (EFL - Equivalent Facial Link). Figure III.17 illustrates similar links (links a and b) between object images which share a common picture line (EOL - Equivalent OBJECT Link).

Another visual (picture) feature which may be represented within two image chunks is the angle feature. Equivalent ANGLE elements of two chunks may be linked by a pair of associated, inversely specified (directed) EAL (Equivalent Angle Link).

58a

Figure III.18

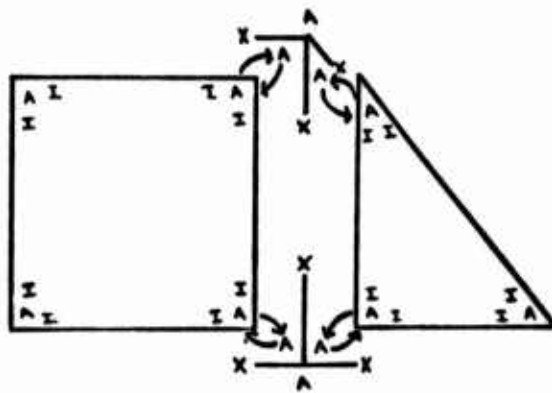
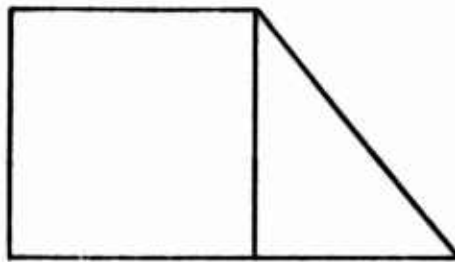
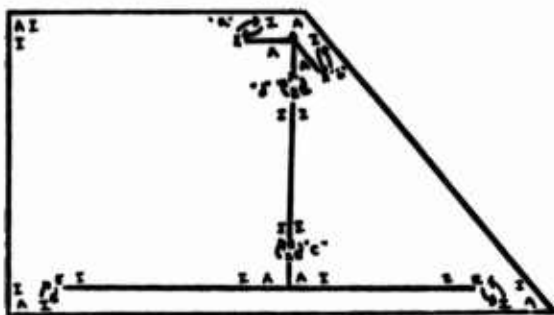
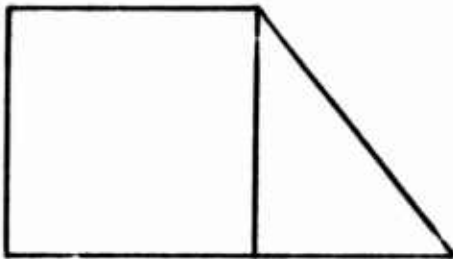
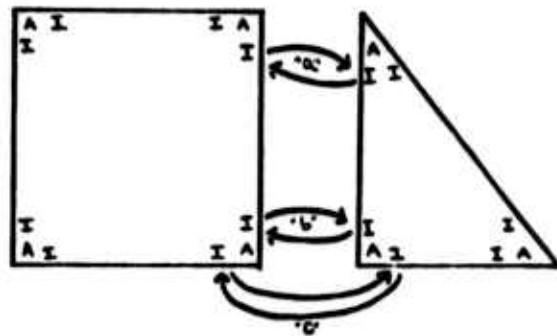
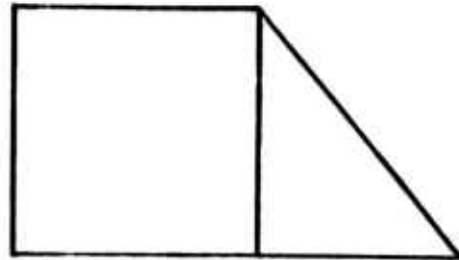


Figure III.19





properties. Figure III.18 illustrates the use of this interchunk relation. An ANGLE element of a VERTEX chunk is linked to one in an OBJECT chunk as a means of representing the context of an object "corner".

The two INTERNAL elements of an OBJECT chunk "corner" configuration may be equivalently linked to two XIT elements of a related VERTEX chunk. This occurs in the case of an object being perceived as containing a line (inside the contour) originating at that object corner. Figure III.19 illustrates the use of these links (links a and b). This figure illustrates the use of a SIDE chunk to represent the "inside" edge of an object's side. This figure also illustrates the equivalency link (links c and d) which is used between an XIT element and an END (of a LINE type chunk) element. The external link of an END element of a LINE chunk is an equivalency link.

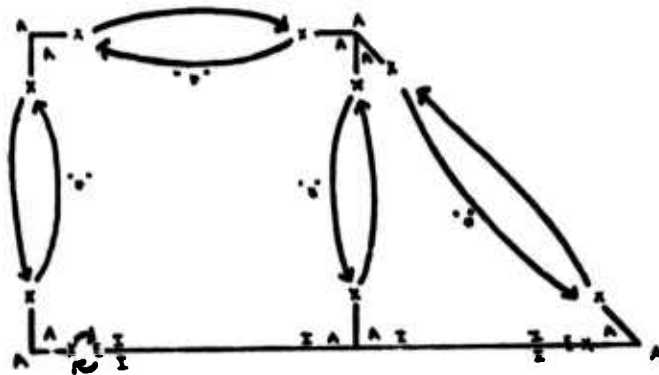
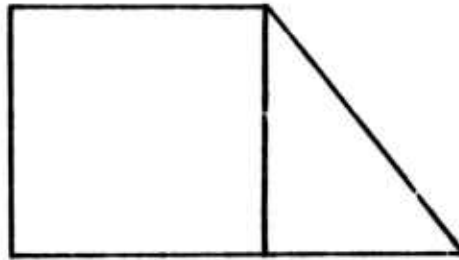
The other type of inter-chunk links are space-traversing in meaning. These links occur between two XIT elements of differing LINE, SIDE, or VERTEX chunks. Figure III.20 illustrates the use of these links ("s") between a LINE and several VERTEX chunks. The three properties DVX (Direction of Vertex eXit), LVX (Link of Vertex eXit), and RVX (Range of Vertex eXit) serve to specify these links.

EQUIVALENCY AND REDUNDANCY

The use of equivalency links to interrelate image chunks introduces a necessary degree of redundancy into the visual image representation. This redundancy is a fundamental aspect of the semantic nature of the visual image. A line segment between two objects is meaningfully part of both image concepts. Figure - ground phenomenon considerations require that this part be included in each of the two image chunks. Similarly, angles are seen as part of an object and as part of a complex

58c

Figure III.20



vertex w
essential
assimilate
equivalen
between
the perce

IMAGE R

To s
flexibility
one singl
diversity
single line

The
represent
descriptio
line draw

The n
nature of
element i
of the de
their asso

vertex where several object corners may meet. The use of equivalency links are an essential factor in the realization of input assimilation. A complex vertex is readily assimilated as object environment (context) during perceptual processing. These equivalency links also provide the necessary means of image traversal (passing) between chunks (concepts). This traversal allows the behavior necessary to explain the perceptual (verbal) description and drawing activities.

IMAGE REPRESENTATION FLEXIBILITY

To satisfy the design criteria 1, 2, and 3, a most necessary characteristic was flexibility of image representation. "Flexibility of image representation" means that one single line drawing must be representable in more than one way, reflecting the diversity of subject behaviors and resultant perceptions observed with regards to a single line drawing.

The Figures III.17 - III.20 illustrate the flexibility of the proposed image representation. Figures III.17(a) - III.20(a) exemplify the general class of verbal description which is associated with each respective image representation of that one line drawing.

The redundancy and flexibility of the image representation, the non-hierarchical nature of the image body structures, and the spatial and equivalency meaning of the element interrelations are the main design features which allow adequate satisfaction of the design criteria specified. The proposal of the various image body elements and their associated interrelationships yield these representation capabilities.



Figure III.17(a)

There is a square next to a triangle
the square is to the left of the triangle

Figure III.18(a)

There is a square
which meets a triangle to the right
meeting at the top and bottom corners of that side

Figure III.19(a)

There is a trapezoid
which has a line from its obtuse corner
to the bottom

Figure III.20(a)

There is a right angle in the lower left
above that is another right angle
then to the right is a three-exit vertex
lines go down and down-right from there to hit the bottom

3. The Image in VIPS

Chapter 1.4 will discuss how the visual image serves to guide the perceptual behavior of VIPS as chunks STM, and how, as each STM chunk is visually confirmed and meaningfully integrated into the image, these chunks are transferred to ITM. ITM is the residence of the complete finalized visual image (perception) resulting from the activity of VIPS.

The image chunks of STM are interrelated by equivalency and space traversing links, as discussed above. When these chunks are incorporated into ITM further relational links are added. These new links are between the Chunk Header elements of related image chunks. The incorporation of image chunks is sequential as indicated. The order of object image chunk incorporation is represented explicitly by the placing of a (NEXT-OB) relational property on the currently last ITM object which is set to reference the new object image chunk upon its incorporation. Furthermore, an object image is incorporated together with its immediate environment as determined by the existing equivalency links of its image body elements. A property (ENV) is added to an object's Chunk Header during incorporation. This property has as value a list of those image body elements which have an inter-chunk equivalency link. These elements themselves become properties of the object's Chunk Header. Each has as value the Chunk Header of the image chunk which contains the image body element to which it is linked by equivalency.

One perceptual strategy (see Chapter II.1) involves the perception of the picture outline. As chunks representing parts of this outline are incorporated environmental links are added between Chunk Headers. These involve the use of the ENV property

and the image body properties as with object incorporation but the inter-chunk links in this case may be space-traversing also.

Figure 21(a), (b), (c) and Figure 22(a), (b), (c) illustrate two image representations produced by VIPS during the explanation of two protocols upon the same line drawing. The component (a) of each figure presents the subject's verbal description. Component (b) presents the image's pictorial representation as shown before. Component (c) presents the Chunk Header structure produced by chunk incorporation into ITM. The doubly directed links are environmental links. The uni-directional dotted links are the NEXT-OB links mentioned above. The figure also indicates the order that the visual concepts (chunks) of the image are noted in the subject's description.

The flexibility of the representation is again illustrated by these two examples. Furthermore, the use of both the equivalency and spatial links between image bodies and the relational links between Chunk Headers appears to be basic to an explanation of the verbal and drawing descriptive behavior.

Figure 21(c) includes a verbalization by VIPS which is based upon NEXT-OB link traversal only. Figure 21(a) shows that these object concepts occur in the same order (by (i), (ii), (iii) and (iv)). Section A of the verbalization is derivable from the Chunk Header (environmental) links. Section B indicates image body traversal of the rectangle (0184) image and of the equivalency link between that image and that of S101. Section C again is derivable solely from the environmental links of the Chunk Header.

The verbalization of Figure 22(a) employs the image body links to a greater degree. The first section (A) is derivable from Chunk Header links. All of the rest (B) of that verbal description is the verbalization of a traversal of the image chunks according to the inter-chunk image body links joining them.

The Chunk Header relations in ITM add a dimension of hierarchy to the image representation beyond that implied by the chunking of the visual information. The final form of the visual image in VIPS (the perception found in ITM) is a heterarchical symbol structure.

Figure III.21(a)

- okay
 - um
 - I have a three-sided triangle (i)
- A which intersects a square on the right (ii)
 - there is a rectangle (iii)
 - on top of the square on the right
 - um
 - this rectangle
- um
 - has one side
 - that is the same side as the square
 - that's on the left of the square
- B and
 - then the rectangle
 - goes to the right
 - and comes down
 - and intersects
 - the square
- um
 - on top of the square
 - and then coming off this rectangle
- C is a four-sided object (iv)
 - which pulls the whole drawing together
 - the four-sided figure has sides
 - on the rectangle
 - the triangle
- and the square

62a

Figure III.21(b)

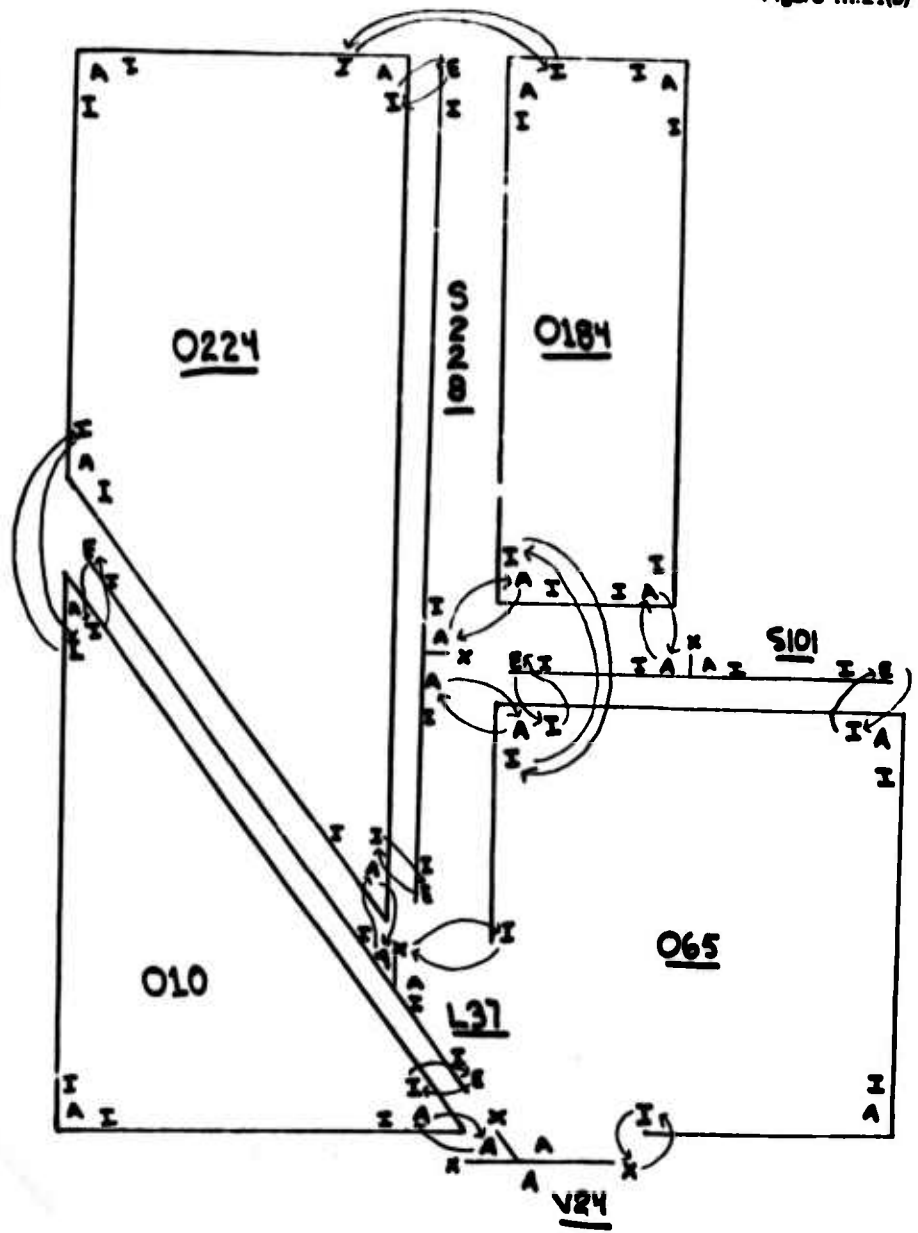
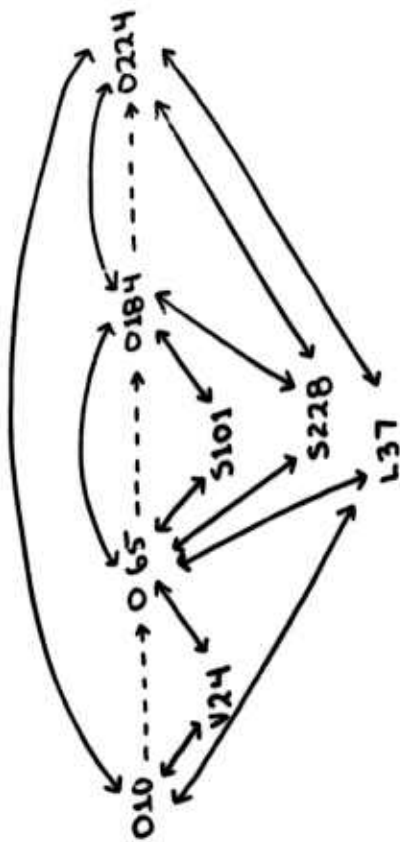


Figure III.21(c)



VIP3
VERBALIZATION
SELECTION

Q NOW THERE IS A TRIANGLE
A1 AND THEN AN OVERLAIN SQUARE-OR-RECTANGLE
A2 AND THEN A RECTANGLE
A3 AND THEN A FOUR SIDED OBJECT

order of verbalized concepts :

O10, O65, O184, O224, S101, S228, O184, O65, O184, O224, O184, O10, O65

Figure III.22(e)

okay
the whole thing
we have
uh
we have a triangle in the lower left hand corner
and we have a vertical coming up to that
and it hits a horizontal over there on the top
and
(pause)
a vertical goes down to the right
a vertical goes down
and it hits a horizontal
you go over to the right and there is a right angle
you move down to the left there's a right
you move down there's
a right angle
and there is a bottom

Figure III.22(b)

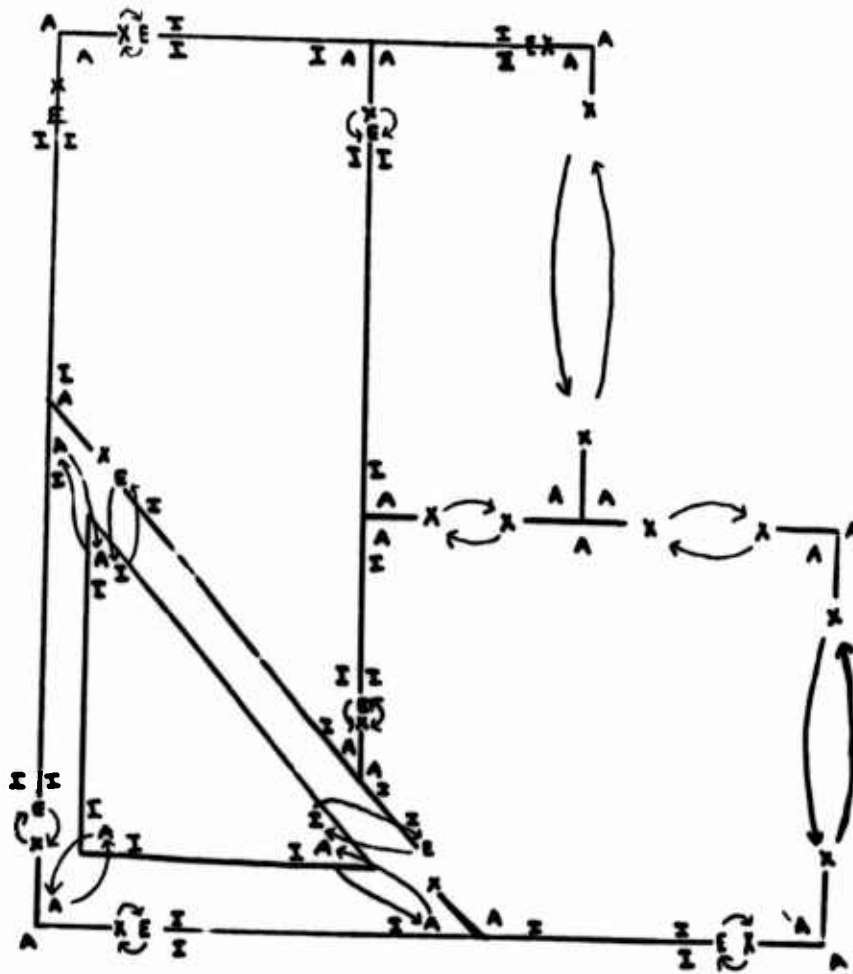
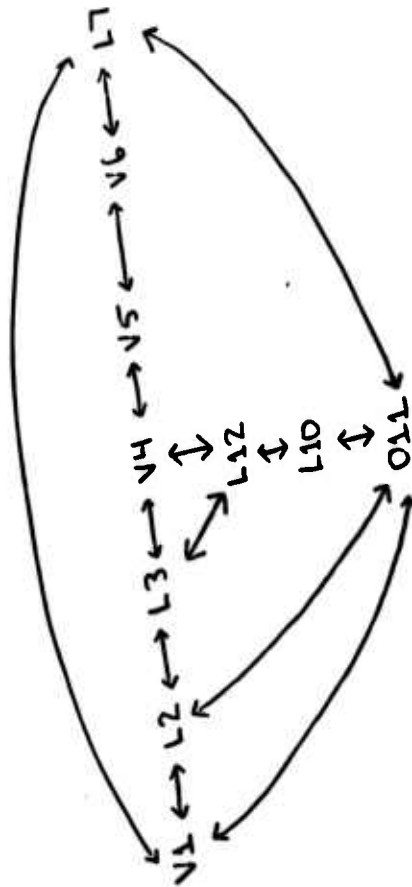


Figure III.22(c)



ORDER OF VERBALIZED CONCEPTS:
 O11, L10, L12, L3, V4, V5, V6, L7

Chapter L4 The Form and Activity of VIPS

The purpose of this chapter is two-fold. It first presents a specification of the memories and processes which are coordinated participants in visual perception. It then illustrates these components in action by presenting an example segment of a trace of VIPS activity. The image representation is also demonstrated as a guide for this activity.

1. System Overview

Throughout the development of VIPS the proliferation of memories and processes was accepted, indeed considered valuable, when such proliferation expedited system design, function, and description. This proliferation was tempered by the requirement that favorable experimental indications for any memory proposal exist. As a result, the proposed system is composed of six memories and four processes, named as follows:

Memories:

1. Picture (PIC) (an external memory)
2. Visual Register (VR)
3. Short Term Memory (STM)
4. Long Term Memory (LTM)
5. Recognition Long Term Memory (RLTM) (long term memory components)
6. Intermediate Term Memory (ITM)

Processes:

1. Assimilation - Accomodation process (AA-SYS)
2. Visual Input process (VI-SYS)
3. Recognition process (REC-SYS)
4. Incorporation process (INC-SYS)

ORDER OF VERBALIZED CONCEPTS:

O11, L10, L12, L3, V4, V5, V6, L7

THE MEMORIES

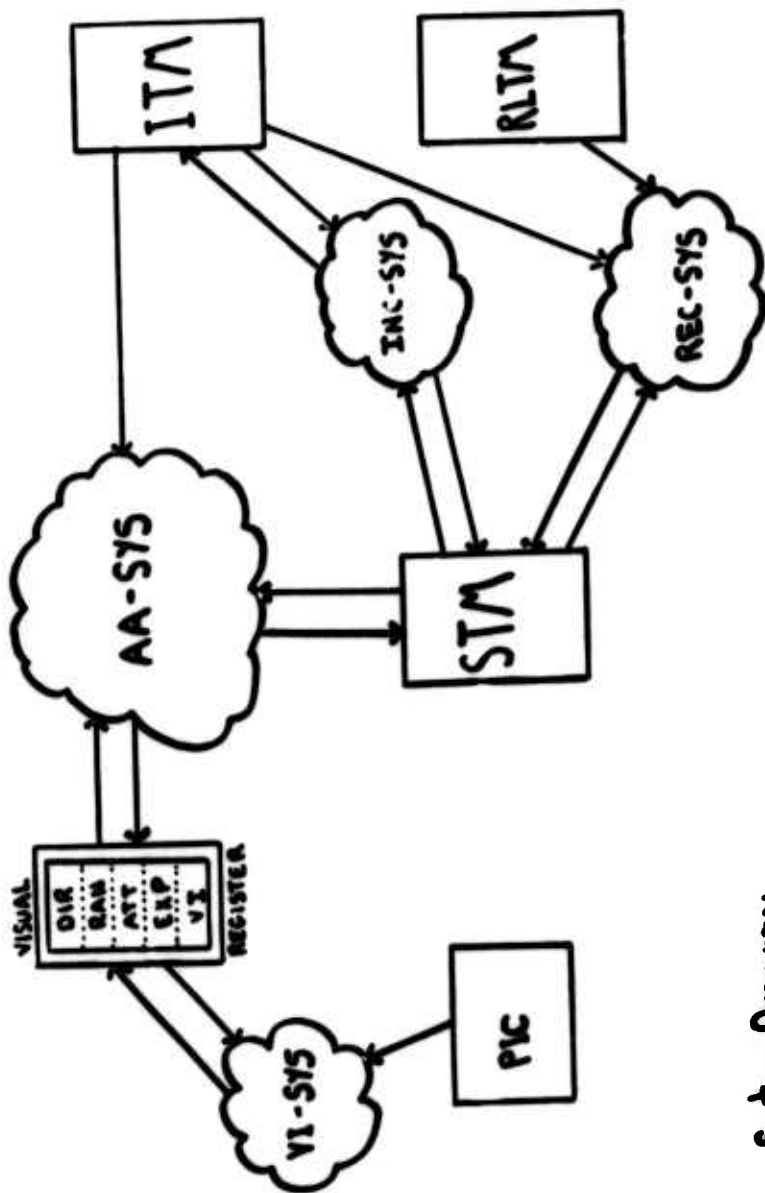
The memories of VIPS differ in several primary characteristics, which serve to distinguish and define each memory. These characteristics are:

1. the associated overall (supra-elemental) information structure;
2. the internal characteristics and structure of the contained memory elements;
3. the variability (in terms of addition of, alteration of, and loss of contained elements) through the performance of a perceptual task;
4. the allowed means of information accessing;
5. the associated processes (those which can access the memory, and how each can effect the memory).

The memories are not separate storage compartments of the cognitive processor, though computer implementation fosters this interpretation. Each active memory is defined by the functional role which its contents play in the perceptual activity undertaken by the system. A symbol or symbol structure may be said to be in one of the active memories (STM or a cell of VR) when it plays the associated processing role.

Figure IV.1 indicates the accessibility relationships between the constituent memories and processes of VIPS. An arrow from a memory to a process indicates that the process can access information from that memory by that memory's associated accessing means. An arrow from a process to a memory indicates that the process can alter that memory by entering new elements or changing existing ones. (Note that LTM is not shown, its elements permeating (being) all memories and processes.)

Figure IV.2 shows the possible inter-memory transfers of information within the



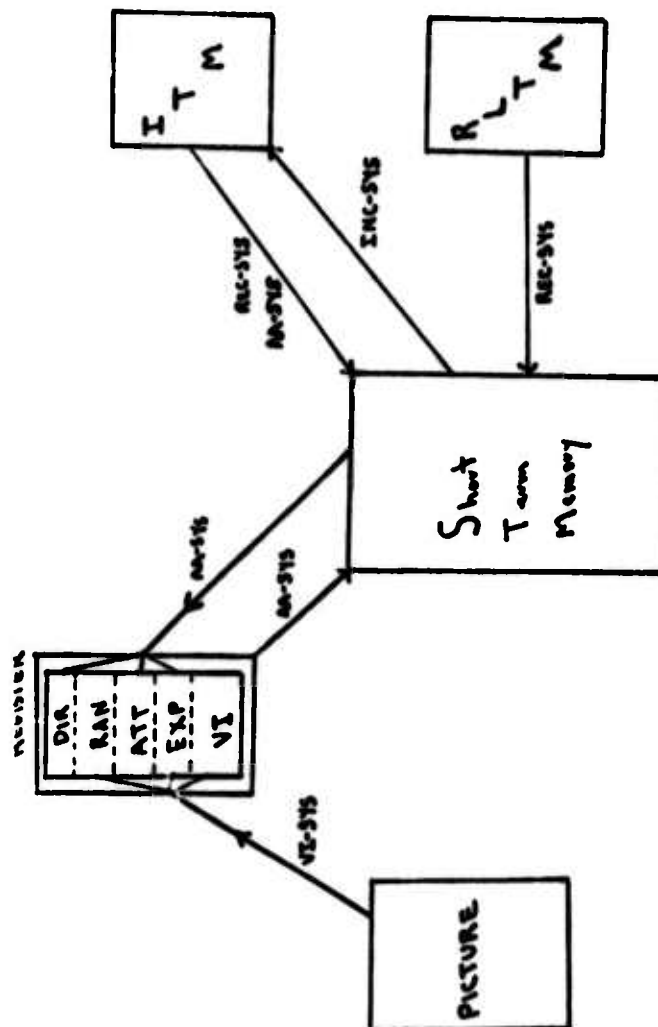
64a

Figure IV.1

System Overview:

Process [O] - Memory [O] Structure

Figure IV 2



System Overview: Inter-memory Information Flow

system.
directed
within
1. P
PI
informa
the sy
behavio
E
organis
organis
called
direct,
to bel
imping
indeter
in
visual
"recep
environ
that ar
than p
Sim

system. Each arrow is labelled with the process capable of performing the indicated, directed information transfer. This figure indicates the flow of visual information within VIPS. It is redundant with Figure IV.1 but emphasizes a different system aspect.

1. PIC (PICTure)

PIC is an external memory (Newell and Simon, 1972) and the source of visual input information in VIPS. Due to the nature of perceptual processing, which VIPS models, the system must represent both subject and environment in order to explain the behavior. As such, PIC is that part of VIPS which represents the environment.

Elsewhere I have defined perception to be the active process by which an organism interprets the states of a subset of its cells (Farley, 1971). Given sufficient organism complexity, these cells are subsets of an organism's nervous system, being called sensory receptors. The significant feature of sensory receptor cells is their direct, or effective, link with the organism's surface. This feature leads the organism to believe that the sensory receptor states indicate conditions of a surrounding, impinging environment. That this belief is reality or illusion appears phenomenally indeterminate.

In the human organism, the receptor cell states which are interpreted during visual perception are those of the rod and cone cells of the eye as organized into "receptive fields" at the retina (Leibovic, 1972). These cells (states) form the visual environment. This is the environment represented by PIC. This is the most external that any memory can be said to be. It is the environment in terms of sensory rather than physical qualities.

Since involvement with actual camera input was to be avoided as necessarily

posing additional problems outside the realm of the thesis, an intermediate descriptive representation of the picture was needed. The representation chosen for this description is a list of vertex descriptions, each vertex description being a linear list of vertex features. This vertex representation has the characteristic of being consistent with feature analysis and extraction proposals for the role of the eye in visual perception. Various neurological and psychophysical studies have indicated the apparent feature extraction capabilities of the eye sub-system of the visual processor (for references see Chapter 11). These studies indicate that this analysis may be effected by a high-order perceptron-like process (Muniz, 1971). Uhr (1973) discusses means for the computer implementation of this "front-end" of visual form perception activity. Each vertex representation of PIC is proposed as being analogous to a list of features which the eye could retrieve at that vertex location. The list of vertex descriptions comprising PIC remains constant throughout the perceptual activity of VIPS.

Gibson (1966) has described the senses as the active systems which register the "invariant structures of available stimulation furnished at the receptors". The theory of VIPS agrees that structure does exist in visual stimulation and that it is registered during perceptual activity. But it argues that this is only an initial level of processing. Structure exists as potential information in the environment. The initial level of perceptual processing transforms that structure into symbols and relations which are known by the organism (or, neurological information). The results of this direct analysis are found in PIC, and possibly VI as the icon is automatically constructed. The visual image produced by visual perception in STM and ITM is the result of constructive activity which is influenced by goals that are beyond the mere registration of stimulation structure (resonance).

The process VI-SYS is the only process with access to the Picture memory. This access is accomplished by means of the Picture memory's associated reference pointer, the Current Picture Pointer (CPP). This pointer always references one of the vertex description lists, being analogous to the current hole position of the primary experimental task situation. Each vertex description list also contains the necessary spatial information to facilitate the repositioning of CPP in a way analogous to that of hole movement. That is by specification of a direction of hole movement. (See Appendix APIC2 for an example of a PIC representation and details of the vertex features and spatial interrelationships used.)

2. VR (Visual Register)

The visual register consists of five independent cells of information, these being:

- A. DIR - direction (of eye-hole movement)
- B. RAN - range (of eye-hole movement)
- C. EXP - expectation
- D. ATT - attention
- E. VI - visual information.

This memory is utilized in VIPS as a communication register between the VI-SYS process and the AA-SYS process and, as such, is highly variable throughout the performance of a perceptual task, with both processes having the ability to inspect and alter its component cells. DIR and ATT may be null, or contain a directionally meaningful image symbol. RAN may be null or contain an image symbol representing distance. EXP may be null or contain an image angle symbol or list of vertex specification symbols. VI may likewise be null or contain a partial or complete image structure having its origin in VI-SYS process operation.

VI

The VI cell is proposed as the equivalent of the so-called Sperling (Sperling, 1960) or iconic (Neisser, 1967) visual memory. Upon VI-SYS activation VI is always empty (set to a null value). At the time of VI-SYS deactivation, the cell contains an image representation (chunk) of the new visual information acquired by the changing of fixation (hole) location. This iconic image in VI is the result of the first (two) of several active stages of information processing and abstraction (Posner, 1969). Sternberg (1967) observed reaction time differences indicating the presence of such a stage of visual representation prior to the application of recognition operations.

Visual information, as it appears in VI, has not yet been spatially integrated into or in any meaningful way associated with the current perceptual image contents of STM. This is one basic characteristic of the visual icon which has been inferred from experimental results (Haber ; 1971). Another is the independence of its contents from the influence of any active perceptual goal. Although perceptual goals play a prime role in VIPS, the functioning of VI-SYS is only dependent upon the contents of PIC and the Visual Register. This provides for the desired goal independence of the resultant VI cell contents. Given visual register specifications for the movement of CPP (the hole), the encoding of information into VI (the icon) is "automatic" (Shiffrin and Geisler , 1973). Finally, the results of visual masking experimentation indicate that the icon is susceptible to erasure from overwriting. VI also has this characteristic as it is set to null then reconstructed with each activation of VI-SYS. The representation of visual information in VI in the present implementation is that of an image chunk. A possibly better choice is that of just the vertex feature list itself, allowing greater flexibility of STM representation.

EXP and ATT

The EXP and ATT cells provide the information necessary for the operation of "preattentive functions", which the perceptual system (through VI-SYS) finds advantageous to apply. The "preattentive functions" of VIPS differ from the wholistic type described by Neisser (1967), but they serve similar purposes.

The EXP cell allows the VI-SYS process to apply a function capable of immediately determining whether a specified visual feature or feature configuration is found within the newly acquired information. The function application is immediate operating in conjunction with iconic image construction. It precedes any effect by the new input upon the current image contents of STM. The contents of VI are not affected by the operation of an expectation function. The use of expectation operations in visual perception has been indicated by various experimental results (Adelman and Smith, 1971). In VIPS, the EXP cell is loaded with the expected visual feature value(s) prior to VI-SYS activation. An appropriate preattentive function is applied by VI-SYS whenever this cell has been prespecified (by AA-SYS). The results of the function activity (YES, NO, or CONTAINED) is placed in EXP prior to deactivation of the VI-SYS process. In some cases the returned value of EXP is sufficient to determine subsequent perceptual behavior (by AA-SYS). In such cases, VI is not accessed nor incorporated into STM. The extent of evaluation of newly arriving sensory input depends upon the predisposition of the system.

The ATT cell allows the VI-SYS process to bypass vertices of negligible interest to the present perceptual goal when moving to a new fixation location (vertex). The

eye movement results of Yarbus (1967) illustrate this ability of the visual system. In ViPS, the ATT cell may be specified with a direction value perpendicular to that of the specified movement direction (DIR) prior to Vi-SYS activation. As such, ATT specifies a side of the line segment which is being traversed. Any vertex encountered during hole movement which is straight on the side of the line specified by ATT is scanned through (no halt) on the way to a further new fixation location (vertex).

DIR and RAN

The DIR and RAN cells are those information units which provide for the necessary communication between the visual information processor (visual cortex) and the motor control processor (cerebellum or basal ganglia) (E. V. Evars, 1973), which controls the specified body, head, or eye (hole) movement activity. These movements are specified in terms of direction (DIR) and range (RAN) at this control level. Neurophysiological research (Pribram, 1971) has indicated such an abstract representation of motor functions at the cortex level.

DIR and RAN indicate the intrinsic role of motor (efferent) activity and information in visual perception. Movement is necessary for the fixation of new external information. But these cells do not only provide information for Vi-SYS to effect such movement. Their spatially meaningful values (symbols) are incorporated into the visual image structure itself by AA-SYS. RAN is a value returned from the motor system in the hole-movement task (so also in ViPS). Neurophysiological results have indicated that the cortex can not be partitioned solely on the basis of sensory or efferent nerve input connections. The interpretation of sensory nerve states (perception) involves the interpretation of associated efferent nerve states (Festinger et al, 1967). Eye

movement research has indicated the effects of eye movements upon perceived position and motion.

3. STM (Short Term Memory)

The Short Term Memory of VIPS corresponds to the notion of short term memory under discussion in cognitive psychology today (Miller, 1956; Newell and Simon, 1972). It represents that limited part of human memory which is active or has recently been active and is still readily accessible. As such, STM is an ordered list of a constant number of independent chunks. A chunk is a collection of symbolic information such that access to any constituent element of a chunk implies the immediate accessibility of all other constituent elements. A chunk is not an independent unit of memory, though. It is rather a pseudo self-contained symbol structure. It is a selective activation of the symbols and relations of LTM. Each selected symbol maintains the potential use (activation) of all of its LTM associations by the active process. The number of chunks in STM has been set at nine in VIPS. This number is within the limits of current estimates of STM chunk capacity. Chapter II.2 shows that the number nine is necessary to explain the observed behavior with the present VIPS process implementations. The three processes (AA-SYS, REC-SYS, INC-SYS) other than VI-SYS all have access to the chunks of STM.

The order of chunks in STM is the result of the action of two basic STM operations. The first is the insertion of a new chunk at the head of STM. This chunk may be created by the currently active process or may be transferred from ITM or the VI cell of the Visual Register. This insertion results in the "pushing" of the existing STM chunks to the back and causing the last chunk to be lost, or "forgotten". This

keeps the number of STM chunks constant. The second operation is the movement of a current chunk of STM to the front of STM. This results from the "attendance" (accessing) of the chunk by the active process (either AA-SYS, REC-SYS, or INC-SYS).

Chunks are only lost from STM by displacement. There is no explicit decay over time, though the oldest, non-attended chunk is the one to be lost and forgotten. This proposal is consistent with those of Waugh and Norman (1968) and Newell and Simon (1972). Reitman (1971) and Shiffrin (1972) have experimentally shown that loss of information from STM is not merely a function of elapsed time. Reitman says:

"Forgetting is produced by characteristics of the events of filled time, not by time itself." (Reitman, 1971, p194)

The order of STM not only is the prime determinant of chunk loss from STM, but also plays a role in the accessing of STM. Chunks in STM are accessed by an operation which sequentially searches (scans) STM for the first chunk satisfying a given condition. Access (attendance) is then accomplished for that chunk. This search traverses STM from front to back. This gives front chunks priority, possibly screening less recently attended chunks in STM which satisfy the same accessing condition. The conditions used to access chunks in VIPS are defined later in this chapter as common features of the three processes which may access STM.

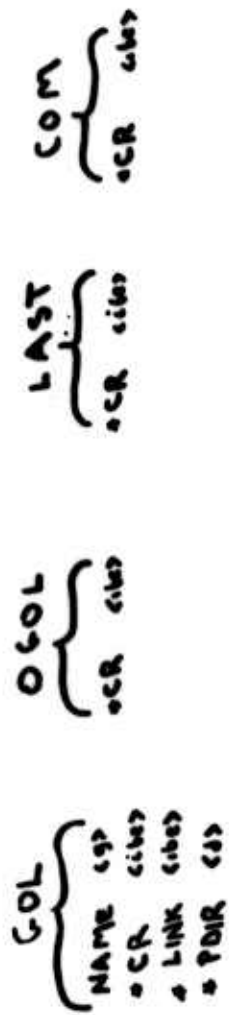
It remains only to discuss the internal characteristics of an STM chunk to complete the specification of Short Term Memory. A visually-meaningful chunk of STM is either just an image chunk (as defined in Chapter 1.3) or an image chunk with additional perceptual control information included. This control information is embodied within a Special Type element, which may be appended to the head (linear list level) of an

Figure IV.3

$\langle \text{STM chunk} \rangle ::= \langle \text{Image chunk} \rangle$ or $\langle \text{Special Element} \rangle \langle \text{Image chunk} \rangle$

$\langle \text{Special Element} \rangle ::= \{ \text{GOL}, \text{LAST}, \text{COM}, \text{OGOL} \}$

72a



$\langle \text{ib} \rangle ::=$ an image body element of the STM chunk
 $\langle \text{g} \rangle ::=$ a gol
 $\langle \text{ib} \rangle ::= \{ \text{up, ur, rt, dr, oo, ol, le, ul} \}$

image chunk of STM (see Figure IV.3). There are four such Special Type elements used within VIPS, these being GOL, COM, LAST, and OGOL. Each will have relevant control properties associated with it. For example, the GOL element holds the name of the active perceptual goal in its associated NAME property. A Special Type element is placed at the head of a chunk with special relevance to the present perceptual goal, facilitating the future access (attendance) of that chunk. Though these Special Type elements are incorporated directly into an image chunk, their constituency is not permanent and is strictly for control purposes. These elements are never part of the final image representation (perception).

Components of Long Term Memory

Long term memory serves the activity of visual perception both as a source of relevant representational knowledge and processes and as an available storage location for selected meaningful perceptual results. The LTM, RLTM, and ITM memories of VIPS are all proposed as being differing, necessary components of this multi-purpose long term memory. LTM is a set of interrelated perceptually-meaningful symbols. RLTM is a specially structured component of long term memory facilitating object recognitions and expectations. ITM is an extension of STM which serves as the available long term storage area for selected, resultant image chunks of the current perceptual activity.

4. LTM (Long Term Memory)

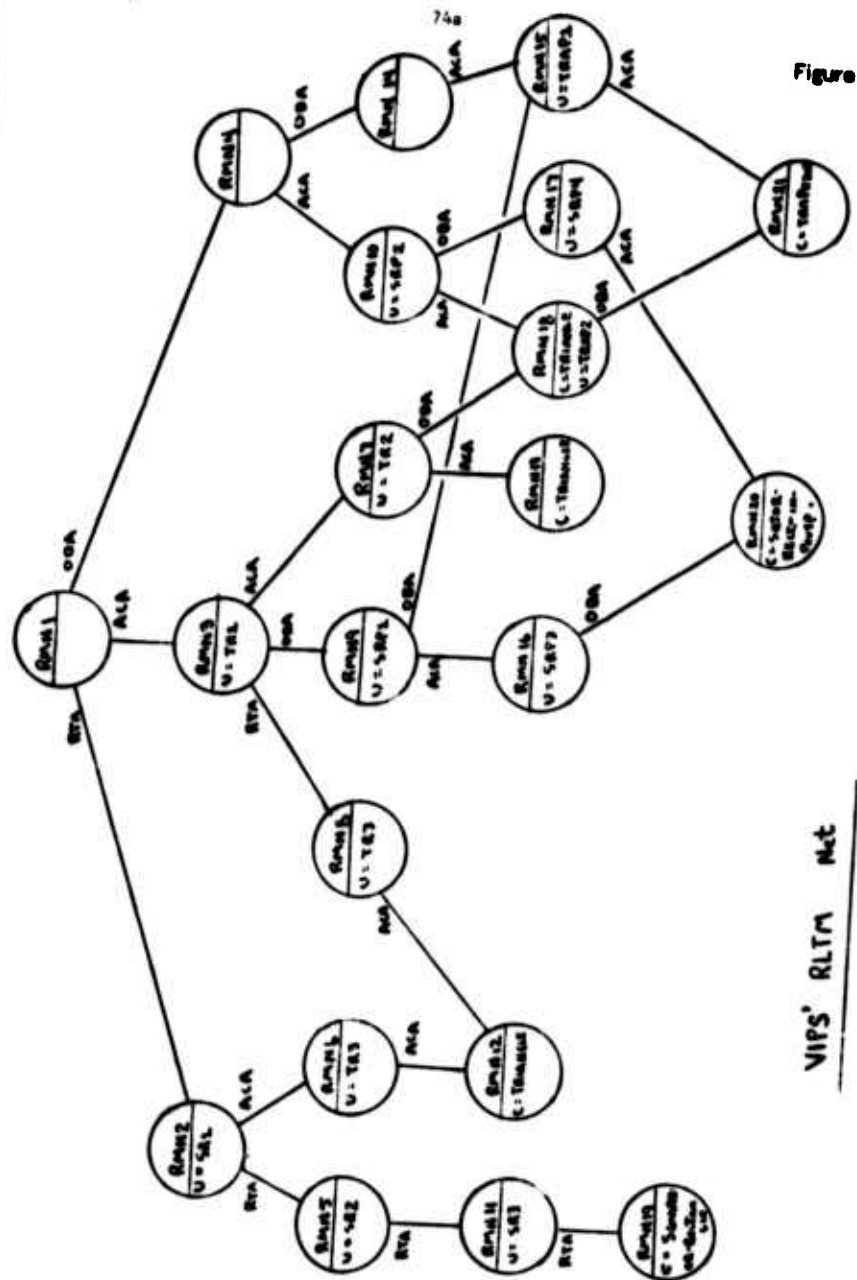
The Long Term Memory of VIPS has no overall information structure associated

with it. Though implemented in the program as a linear list, it is ideally to be considered as the set of symbols usable throughout the system by all processes and in all memories. This includes those symbols introduced in Chapter 13 which defined the image representation. Thus, in one sense LTM does not exist as a separate entity at all, but permeates all the other processes and memories. The use of a symbol is an activation of that element of LTM (Anderson and Bower, 1973). The set of LTM symbols does not change within the course of one perceptual task performance, though theoretically it could be altered by learning processes using experience over performances.

The visually-meaningful symbols of LTM are interconnected through the placement of properties representing spatial and featural interrelationships. The resulting information structure is a form of semantic net (Quillian, 1968). The existence of these interrelationships affords a valuable means of indirectly accessing LTM elements. For example, the opposite property ("OPPO") interrelates directional symbols, with "TOPPO of LE" being (accessing) RT. (See Appendix A.LTM for exposition of the symbols and the associated semantic interrelationships.)

5. R.LTM (Recognition Long Term Memory)

The recognition long term memory is structured as a non-binary (n-ary) discrimination net with elements being accessible through traversal from an initial element RMN] (see Figure IV.4). This memory does not change throughout a perceptual performance by VIPS. As noted with LTM, theoretically it is alterable by learning processes. The basic information structure, operation, and possible learning (and forgetting) processes of R.LTM have been investigated previously as EPAM (Feigenbaum and Simon, 1964).



VIPS'	RLTM	Net
-------	------	-----

RLTM has associated with it a reference pointer, the current recognition memory pointer (CRMP). It is used to access the memory by the REC-SYS process, the only process capable of accessing RLTM. The elements of RLTM are called recognition memory nodes (named RMN_i, i an integer). RMN₁ is the initial node of the decision network, CRMP being set to it upon activation of the REC-SYS process. As REC-SYS acquires new relevant angle feature information from STM, CRMP is made to traverse RLTM accordingly.

There are two special properties (COM and UQX), which may be associated with every recognition memory node, exclusive of that node's position (interconnections) within the recognition network. Either or both of these special properties may be specified for any node. If REC-SYS is successful in constructing a complete object image in STM, the COM (Complete) property of the last RLTM node referenced by CRMP indicates the name of that object. COM is specified only if a name is known for the traversed image object. If, on the other hand, only partial object information is available to REC-SYS in STM, the UQX property of the last node referenced by CRMP yields a process which will complete the image of the suggested, known object. UQX is specified only if a known object is suggested by the partial information in memory.

There is one node acting as a sink node in RLTM (CRMP is set to a null value). This node is characterized by its complete lack of any exiting traversal links and by the absence of both COM and UQX properties. The node is reached in traversal when the current object angle information found by REC-SYS is not compatible with any known object, nor is part of any known object. (See Appendix A.RLTM for the definition of the RLTM used in the implemented VIPS.) Figure IV.4 is a graphic representation of that appendix.

image chunk of STM (see Figure IV.3). There are four such Special Type elements used within ViPS, these being GOL, COM, LAST, and OGOL. Each will have relevant control properties associated with it. For example, the GOL element holds the name of the active perceptual goal in its associated NAME property. A Special Type element is placed at the head of a chunk with special relevance to the present perceptual goal, facilitating the future access (attendance) of that chunk. Though these Special Type elements are incorporated directly into an image chunk, their constituency is not permanent and is strictly for control purposes. These elements are never part of the final image representation (perception).

Components of Long Term Memory

Long term memory serves the activity of visual perception both as a source of relevant representational knowledge and processes and as an available storage location for selected meaningful perceptual results. The LTM, RLTM, and ITM memories of ViPS are all proposed as being differing, necessary components of this multi-purpose long term memory. LTM is a set of interrelated perceptually-meaningful symbols. RLTM is a specially structured component of long term memory facilitating object recognitions and expectations. ITM is an extension of STM which serves as the available long term storage area for selected, resultant image chunks of the current perceptual activity.

4. LTM (Long Term Memory)

The Long Term Memory of ViPS has no overall information structure associated

6. ITM (Intermediate Term Memory)

The elements of ITM are copies (from STM) of selected image chunks. At the time of initial ViPS activation, ITM is empty (set to a null value). ITM is unbounded in size and grows, with the addition of new element chunks, throughout the perceptual performance. ITM holds the complete perception (visual image) at the conclusion of the perceptual activity. Elements are entered into ITM from STM by the INC-SYS process. Any special Type element (i.e. GOL) is removed from a selected chunk prior to being copied into ITM.

The activity of INC-SYS is equivalent to a "rehearsal" operation (Waugh and Norman, 1965). Actually ViPS provides two forms of rehearsal operations. The temporary means of chunk attendance maintains an image chunk in STM by moving it to the head of that ordered memory. The more permanent means is the copying of selected chunks into ITM as performed by INC-SYS.

IN ViPS, four reference pointers are associated with ITM and are utilized by INC-SYS to access and update ITM contents. FIRST-ITM references the first image chunk entered into ITM by the current perceptual activity. LAST-ITM references the last chunk of ITM. FIRST-OB points to the first element (chunk) in ITM representing an object (OBJECT type chunk). LAST-OB references the last chunk currently in ITM which represents an object.

This ITM proposal is in concert with the "assumption of unreliable STM", which is discussed by Newell (1973). STM, as it has been proposed, is a highly volatile, transient storage area. It receives content input from both internal memories and processes and sensory-derived (iconic) sources. It will also contain incorrect or irrelevant debris from the active processing of such inputs. ITM is a cache-like

component of LTM. ITM is a readily accessible extension of STM, consisting of selected, meaningful image chunks. The processes which can access STM can also access ITM by use of the NEED operation. The NEED operation extends the search for a chunk satisfying a specified condition to the contents of ITM. If a satisfactory chunk is found in ITM, the NEED operation transfers a copy to STM. All image chunk construction or alteration is carried out in STM. Selected results are copied to ITM by INC-SYS as new current context information.

Hunt (1971) proposes a somewhat similar operational definition of ITM with regard to language comprehension. He proposes the contents of ITM to be meaningful information "molecules", as opposed to the "atoms" of STM. In VIPS all goals and image structures must at some point reside in STM in order to actively influence or determine behavior. This differs from the Hunt proposal which allows complex structures in ITM to control behavior.

The "dual trace" theory of memory postulates a short term memory which is reverberatory in nature and a consolidated long term memory trace which is stable and constructed as an effect of the reverberatory activity. This theory has been accommodated to become a "multi-trace" theory as investigations into the effects of electroconvulsive shock and certain drugs upon memory consolidation indicate the probable existence of an intermediate holding stage between the two extremes of the dual trace theory (See John, 1971).

Wickelgren (1972) develops a multi-trace theory of memory based on the accessibility (strength) decay of information with elapsed time. Though elapsed time is not a factor in the definition of VIPS memory components, an ITM with a time factor of several minutes, as is proposed by Wickelgren, is consistent with the time

requirements for information maintenance which must be imposed upon the ITM of VIPS to explain the observed behavior.

THE PROCESSES

Each of the constituent processes of VIPS (AA-SYS, REC-SYS, INC-SYS, VI-SYS) is defined in the form of a production system, a means of cognitive psychological modeling, which is adequately defined, discussed, and illustrated elsewhere (Newell and Simon, 1972; Newell, 1973). A production system is comprised of an ordered, finite set of production rules, each having a condition half and an action half. With each system operation cycle, the ordered set of rules of the currently active process is searched according to the defined order until a production rule is found whose condition half is fully satisfied by the current state of the relevant memory (memories). That rule is then said to fire; the operations of the rule's action half are executed in light of the memory information accessed during satisfaction of the rule's condition half. The necessary functions which serve to imbed the production system control structure of VIPS processes within LISP 1.6 are those implemented by Moran (1973) for his VIS system. STM management is also done by utility functions defined by Moran.

The proposed processes are distinguished by several definitional criteria, these being:

1. the memories referenced and operations employed in the condition halves of the constituent rules;
2. the memories referenced and operations employed in the action halves of the constituent rules;
3. the degree of process similarity between subjects.

One outstanding operational difference between processes is that of a process' role in the VIPS cycle of operation. AA-SYS (the assimilation - accommodation process) is considered to be the main process and all three others are supplementary processes. AA-SYS is the process initially activated, bringing life to VIPS. It can subsequently activate any of the other three supplementary processes, which upon their deactivation yield control again to AA-SYS (see Figure IV.5). Upon perception completion AA-SYS is deactivated, killing VIPS itself.

To initiate the usual cyclic course of VIPS activity, AA-SYS loads the visual register according to current STM contents and then activates VI-SYS. This is effectively a request for new visual input. VI-SYS references the external Picture memory(PIC), alters the visual register accordingly, and deactivates, returning control to AA-SYS. AA-SYS considers the resultant Visual Register cell contents in light of current, relevant STM image chunks. Either the new information can be consistently assimilated into STM or the image contents must be accommodated to realize the incorporation of the new input. AA-SYS may activate REC-SYS to perform object recognition or INC-SYS to incorporate selected STM chunks into ITM. It then reloads the Visual Register and again activates VI-SYS. This continues until a satisfactory perception has been constructed as a visual image in ITM.

1. VI-SYS (Visual Input Process)

The VI-SYS process has several distinguishing characteristics. The most notable of these is that it is proposed to be exactly the same for all subjects. This is the consequent of several factors the main one being a result of a stated goal of this thesis. This goal is the development of image elements and structures capable of

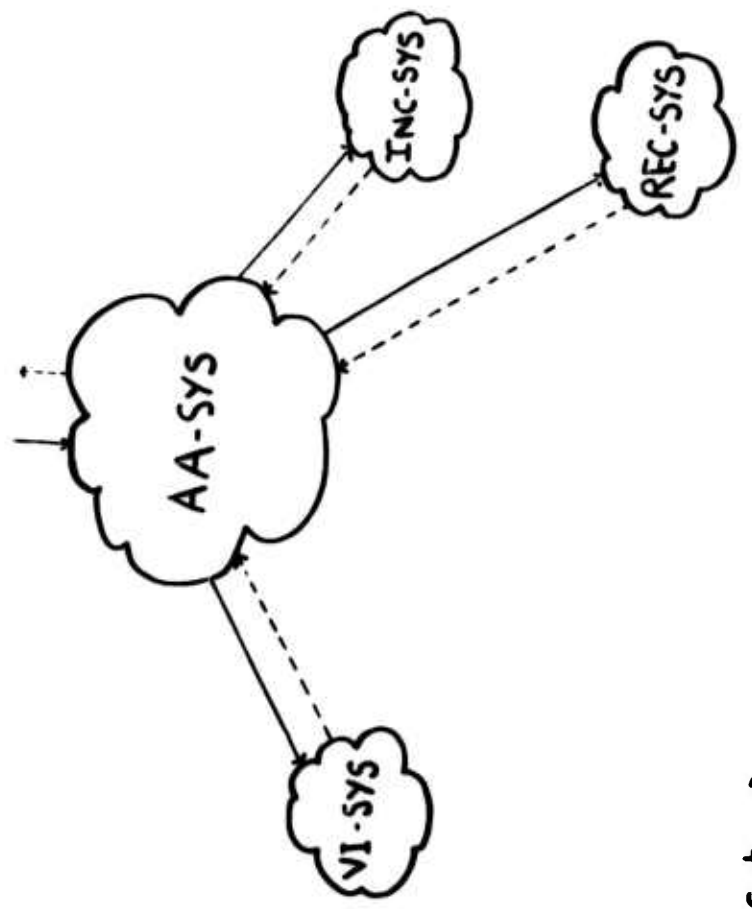


Figure IV

System Overview:

representing all subjects' complete and intermediate (partial) line drawing perceptions. Furthermore, the list of visual features obtainable at a vertex by the eyes is considered to be equivalent for all subjects for the given line drawing environment.

Another factor which leads to its unique uniform definition is the process' access of only the Visual Register and Picture memories during activation. The process is thus only indirectly affected by any active perceptual goals, which are the main source of subject behavior differences. The use to which the process is put may and does vary, but its definition does not.

VI-SYS is also special in that it is the only process with access to the Picture memory. Thus, the Visual Input process is the only interface between the environment (external visual information of PIC) and an internal memory (Visual Register) of the proposed perceptual system. The process is not directly analogous to the human eye, though. Rather, VI-SYS is a combination of the motor process controlling eye (hand) movements and the process which operates upon the eye-produced feature information to produce (construct) the visual icon.

Process Similarities

The rules of the other three processes all reference STM in their condition halves, and are all primarily dependent upon the name of the current goal for determination of the rule to be fired. Thus all rules, but for a few special exceptions, start with two basic operations in their condition halves, these being "(a1: GOL)" (or "(a1: GOL \$1)") and "((NAME OF GOL) = ??)". The "(a1: GOL)" condition operation accesses the STM chunk to which the Special Type element GOL has been appended. The local (to the rule) variable "a1:" is assigned the accessed chunk as its value (when the variable "\$1"

is added, "§1" is locally assigned the value of the Chunk Header element of the accessed chunk. The "((NAME OF GOL)=??)" operation determines if this rule is to be fired in association with the current goal (e.g. if the name of the current goal is ??). If these two condition operations are satisfied, the further condition operations of that rule are sequentially applied to STM for possible satisfaction. Such rule definition creates a valuable partitioning of the rules of a process (by goals), expediting both system operation and rule definition.

A second similarity is the means by which these processes gain access (give attention) to the chunks of information in STM. The rules of these processes achieve this access in only two basic ways. One means of access has been illustrated above by the "(el: GOL)" condition operation. This means of chunk accessing depends upon the existence of a Special Type element (ie. GOL) which has previously been appended to the appropriate chunk. This is pre-determined (premeditated) attendance of a chunk which is relevant to the current perceptual goal.

The other means of attending to a chunk in STM is through the use the "(el: §1 (HAS ??))" condition operation, §1 being an integer. The "(el: §1 (HAS ??))" operation locates and attends to the first STM chunk, if any, which contains the element ?? within its image body structure. The accessing operation only must consider the chunks of STM at the top (linear list) level to determine simple element containment. A successful search results in the local binding of "el:" to the accessed chunk and "§1" to the Chunk Header element of that chunk. The "(NE D el: (HAS ??))" condition operation extends the search for the desired chunk to ITM. The operation automatically establishes a copy of the accessed chunk at the head of ITM (if found) and locally binds the "el:" and "§1" variables as noted above.

This accessing of chunks by either image element containment or Special Type element appendment again greatly improves system operation and definition in terms of speed and simplicity. The attendance of a chunk is achieved without encountering the complex semantic structure being represented in the image body. As noted, there is no attempted syntactic matching of any element sequences at the linear list level and no attempted schema matching to the whole or any part of the image body structure to accomplish chunk access. The definition of the symbolic elements of the image chunk as property groups which embody the meaning make this possible.

Process Control Structure

The position of current goal as the primary determinant of rule firing implies a definite control structure for these production systems and thus for the cognitive processes they represent. Upon each system cycle, the GOL-appended chunk is first attended, and the name of the current goal (NAME OF GOL) is matched. Any further relevant chunks are then attended. Chunk attendance does not involve consideration of the chunk's informational content. Finally, conditions based upon the information of the accessed chunks (and the sensory register, if perception) are evaluated and a rule fires producing the "desired" activity.

The current goal is an internal (internally-derived) state which influences the perceptual activity of VIPS in three basic ways. It first determines what of the available information is to be attended during a system cycle. Secondly, by partitioning the rules it determines the class of actions possible during a system cycle. Finally, by determining actions it influences the final representation (perception) which is realized.

2. AA-SYS (Assimilation - Accomodation Process)

Piaget (1967, 1970) has defined and discussed the fundamental assimilation-accomodation activity of any living system which interacts with its environment. The activity of visual perception is a primary example of such interaction. In the interaction, the system must

"assimilate" the external world into the structures that have already been constructed and secondly to readjust these structures as a function of subtle transformations, i.e. to "accomodate" them to external objects" (Piaget, 1967, p8)

The assimilation-accomodation process is the only process with access to both the newly acquired visual information, found in the Visual Register, and the image memory currently in STM. Given this new visual information, the task of AA-SYS is either to assimilate it into the current image in STM or to accomodate this current image to accomplish integration of new conflicting visual information. The new information may be assimilated by entering the contents of VI into STM as a new image chunk. This involves the linking of the iconic image to the current image of STM by creating appropriate spatial or equivalency relations. The assimilation may be accomplished by just marking appropriate elements of an existing (hypothesized) image chunk as being confirmed. Finally, assimilation may involve no action upon the existing STM contents. When expectations concerning the new visual input are not fulfilled, the image memory of STM must be accomodated to achieve consistency with this input. Such expectations derive from a hypothesized object image, from a previously constructed image chunk thought to represent the part of the visual field currently fixated, or from other

possible perceptual inferences. The source of the incorrect expectation must be altered or another existing consistent image chunk must be found to accomodate STM to the conflicting input. REC-SYS and INC-SYS may be appropriately activated by AA-SYS to accomplish its task.

AA-SYS is the most highly variable process between subjects. It is here that the subject's perceptual strategies are primarily represented. A strategy is defined by a set of possible goals, the possible transitions between the goals, and the conditions associated with each possible transition. AA-SYS is the generator of most active perceptual goals (of their transitions), and as such, is the primary embodiment of strategy. Two general strategies have been inferred to account for the observed subjects' behavior, and are discussed in Chapter II.1

3. REC-SYS (RECOgnition Process)

When the recognition process is activated by AA-SYS, it will proceed to access and traverse various relevant image chunks of short term memory. Based upon the image information encountered, the process traverses the current recognition memory pointer (CRMP) through RLTM. It simultaneously constructs a partial or complete object image as possible. Upon completion of this image traversal an object name, a known object goal, or that insufficient information is available for either will be decided by the process, according to the COM or UQX property of the recognition memory node finally referenced by CRMP.

The basic operations which traverse the image and RLTM do not vary with distinct subjects, due to the fact that the image structures employed by all subjects are the same. The rule with the correct memory traversing operation action part is

determined by the condition part based upon the type of image structure the currently accessed chunk embodies, the type of image element referenced within that image chunk, and the current perceptual goal (prespecified by AA-SYS). The dependence of rule firing upon a current goal results in REC-SYS varying between subjects. The regularity of image body and RLTM structures between subjects restricts this variability to a level below that of AA-SYS.

4. INC-SYS (INCorporation Process)

When the incorporation process is activated by AA-SYS, its task is to enter copies of the appropriate image chunks into ITM, eliminating any outdated version of the same image structure that may exist in ITM. It is the only process which may alter the contents of ITM. (Whether a copy should be made or only a referencing pointer established to the chunk is not clear). This process may also be activated to improve upon an existing ITM representation. This may involve the realization of a more concise image or an attempt to eliminate possible inconsistencies. These functions are discussed further in the next chapter.

The basic operations for copying image chunks into ITM do not change with differing subjects. With differing perceptual strategies, different image chunk types will be marked for incorporation, though. Thus, similar to REC-SYS, INC-SYS has some limited variance between subject models.

2. The System in Action

A PROTOCOL - PROGRAM TRACE SEQUENCE

The rest of the chapter presents an illustrative example of VIPS in operation. This example corresponds to the initial protocol segment of Figure II.14. The example illustrates not only system operation but also features of protocol - program correspondence.

Figure IV.6 presents the two initial protocol frames, V1 and V2. Figure IV.7 shows the complete line drawing stimulus involved with the vertex numbering convention that is in use. Frame V1 indicates that vertex number 1 is initially in view, that being the right angle vertex in the lower left of the line drawing. The subject considers this available visual information and indicates by statement [S3] that it is sufficient to suggest (hypothesize) a square. By statements [S4], [S5], and [S6], the subject indicates the involvement of this object hypothesis in the subsequent hole movement activity of frame V2.

Figure IV.8 presents the initial segment of the corresponding trace of VIPS activity. VIPS is activated by an external activation of the AA-SYS process as indicated by the first line of the trace. Initially STM consists of nine empty cells as indicated by the line labelled by "0,0". This number is the system cycle number. It is used to label STM listings and will be underlined in the trace segments. Upon this initial activation, rule RAA00 fires which activates VI-SYS to obtain the first view of the line drawing (from PIC). In VIPS, the initial goal is NIL.

Figure IV.6

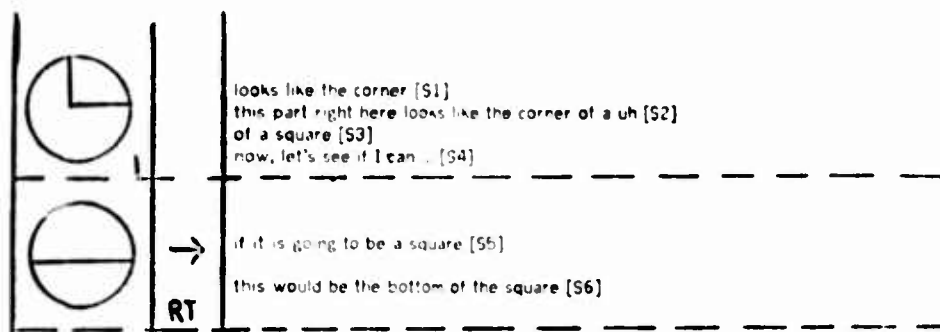
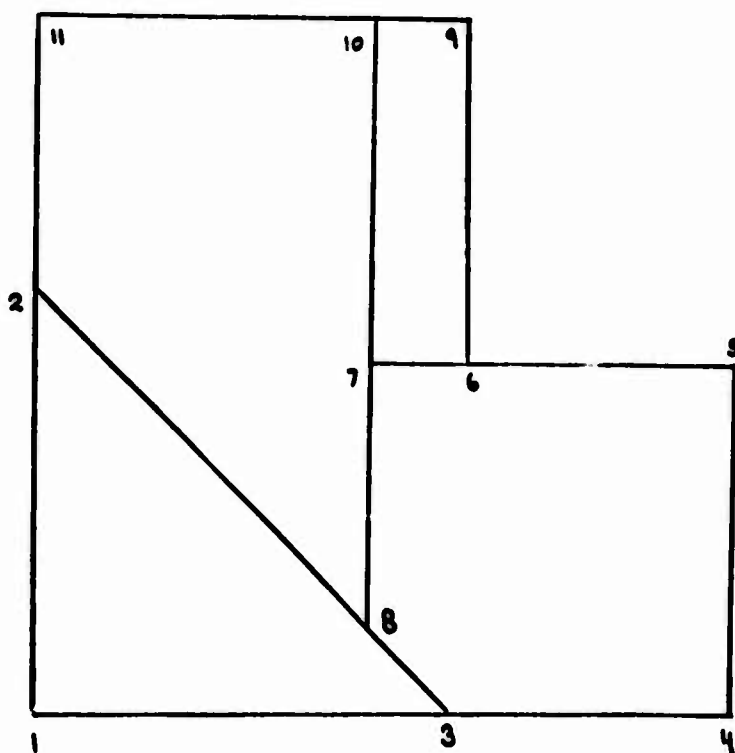


Figure IV.7



• *Source: author's calculations based on data from the 2000 Census of the United States.*

1

Rule VII of VI-SYS fires and constructs the initial iconic image in VI in accordance with Visual Register and CPP values set prior to VIPS activation. Figure IV.9 is a schematic pictorial representation of the vertex image (labelled by the Chunk Header element V3) which VI-SYS develops in VI. As indicated by the STM listing "2,0", STM is not affected by VI-SYS activity.

Upon VI-SYS deactivation, control is returned to AA-SYS and rule RAA0 fires. This rule "sees" the first view now in the VI cell and transfers this information to STM as a new image chunk. The rule also appends the Special type element LAST to this chunk and creates another STM chunk which consists solely of the Special type element GOL (see "3,0").

Two general perceptual strategies have been inferred from the available protocol data. This has resulted in the implementation of two corresponding versions of AA-SYS (see Chapter II.2). The strategy currently under discussion is that successively recognizing objects until completely representing (perceiving) the line drawing. As such, the active goal, "(NAME OF GOL)", is set to RNO and REC-SYS is activated (ACT *9).

Prior to a REC-SYS activation with goal RNO, one chunk is always marked by the element LAST to indicate the first image chunk to be traversed by REC-SYS. A chunk is also always created containing only the element GOL. That chunk is the site of the construction of the new object image.

Rule RECO (of REC-SYS) fires, initially setting CRMP to RMN1. The rule traverses the right angle of the vertex chunk V3 and accordingly sets CRMP to RMN2 (see Figure IV.4). Furthermore, the process begins construction of an object image (OIO) in the GOL - appended chunk (see "4,0"). Figure IV.10 illustrates this partial object image.

87a



Figure IV.9



Figure IV.10

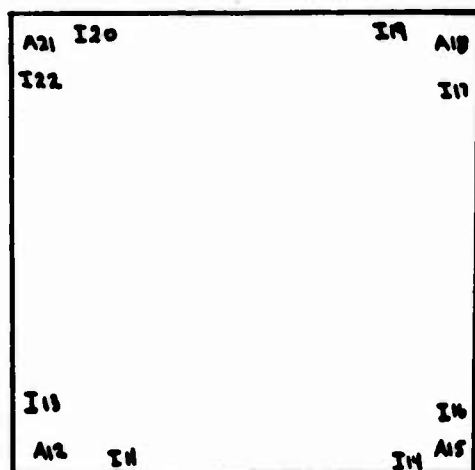


Figure IV.11

$O_{10} (5,0)$

When the active goal is *RM2*, two special properties are associated with the *GOL* element. These are *LINK* and *PDIR* (Prime DiRection). *LINK* references an image body element of a relevant image chunk. This element is the beginning point for image traversal by *REC-SYS* during a system cycle. When no further image chunks exist (in *STM* or *ITM*) which are relevant to the current object recognition, *LINK* of *GOL* is set to *UQX*. The *PDIR* property of *GOL* indicates the direction (perpendicular to the traversal direction) which is "perceived" as being to the inside of the object. The effect of *PDIR* upon *REC-SYS* activity is representative of figure-ground judgements. *PDIR* guides the traversal of image chunks and determines which encountered features (angles) are interior to the object and thus relevant to its recognition.

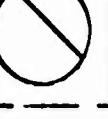
In the present case, *REC0* sets *LINK* of *GOL* to *UQX* as the vertex chunk (*V3*) is the only (relevant) image information available. *PDIR* is initially set externally but all subsequent values are determined by *VIPS*. Since *LINK* of *GOL* is *UQX*, rule *REC1* fires for the next system cycle. The *UQX* property of *RMN2* (current *CRMP* value) yields the process *SR1* which completes the hypothesized image of a square-or-rectangle. This is indicated by *STM* listing "5.0" and illustrated by Figure IV.11. The rule sets the current goal to *SK02* (Start Known Object 2-dimensional). *REC-SYS* deactivates and control again returns to *AA-SYS* (Figure IV.12).

Rule *SK021* (of *AA-SYS*) fires, changing the goal to *K02* and loading the *EXP*, *ATT*, and *DIR* cells of the Visual Register with values taken from the object image *O10* (see Figure IV.12). These values are obtained by use of the *CR* property of *GOL*. *CR* currently references image body element *I11* (type *INTERNAL*). *DIR* is set to the value of the *DIR* property of *I11*. *ATT* is set to the value of the *DIR1* property of *I11*. *EXP* is set to the value of the *ANG* property of *A15*, which is (*ANG* of (*ATT* of (*LINK* of *I11*))).

Figure IV.12

ACT #1 AA-SYS
 RULE S-001 START LOOKING FOR A DOWN OBJECT INFO
 ACT #23 VI-SYS
 RULE V19 EXP & ATT & NO STRAIGHT SIDES FOUND
 DIR IS RT
 RAN IS NIL
 EXP IS RTA
 ATT IS UP
 VI IS NIL
 AT VERTEX VER3
 DIR IS RT
 RAN IS LO
 EXP IS NO
 ATT IS UP
 VI IS (V24 P25 X26 A31 X30 A29 X28 A27)

Figure IV.13

		aha! [S7] looks like we might have a triangular object [S8] instead of a square [S9] this would be
	↑ UL	the second side of the triangle [S10] (slowly)
	↑ UL	(through vertex 8)
	↑ UL	

In the generalized rule specification, I11 is (CR of GOL). The rule then activates VI-SYS (ACT #23).

Figure IV.13 returns to the next segment of the transcribed protocol. The subject reaches vertex number 3 at protocol frame V3. He notes, by statement [S7], a discrepancy between expectations and the new visual information. The acute angle, seen as interior to the "figure", causes the subject to accommodate the previously hypothesized object (square-or-rectangle) image. The new information leads the subject to now hypothesize a triangle, as indicated by statements [S8] and [S9]. The subject moves the hole to begin confirmation of this new object image (statement [S10]). The subject moves through vertex number 8 (frame V5) without hesitation.

Correspondingly, VI-SYS rule V19 fires, which sets CPP to the vertex description list of vertex number three (VER3 of Appendix A.PIC). The rule then alters the visual register accordingly (see Figure IV.14). VI contains the image representation of the new vertex, which is illustrated by Figure IV.15. EXP is set to NO as the angle feature ACA differs from the expected RTA value. EXP is set by the pre-attentive function CAX which is applied by rule V19 of VI-SYS. VI-SYS deactivates and control returns again to AA-SYS.

Rule KO24 (of AA-SYS) fires as a result of the NO value of EXP. KO24 accommodates the object image chunk O10 in accordance with the image of VI. Image body elements I17, A18, I19, I20, A21, and I22 are removed from the chunk. The elements A15 and I16 are altered to correctly embody the iconic input in VI. The ANG property of A15 is changed to ACA and the DIX and DIV1 properties of I16 are changed to LL and DL respectively. This action is indicated by STM listing "8.0" and illustrated by Figure IV.16. The Position type element, P25, is added to the object

GOAL-IS: KO2
 7,0 STM: (GOL O10(NAME SQUARE-OR-RECTANGLE DIM TWO TYPE OBJE~
 CTP4 111 A12 113 114 A15 116 117 A18 119 120 A21 122))
 (V3(CR X5 PNUM P4 VSPEC (RTA DL) SPEC V2 TYPE VERTEX)(P4 X5 A8~
 X7 A6))
 () () () () ()

ACT *1: AA-SYS.

RULE KO24: VI IS NOT AS GOAL REQUIRES.

GOAL-IS: RK1
 8,0 STM: (V24(CR X26 PNUM P25 VSPEC (RL UL) SPEC TE TYPE VERTEX)~
 (P25 X26 A31 X30 A29 X28 A27))
 (GOL O10(NAME SQUARE-OR-RECTANGLE DIM TWO TYPE OBJECT)(P4 ~
 111 A12 113 114 A15 116 P25))
 (V3(CR X5 PNUM P4 VSPEC (RTA DL) SPEC V2 TYPE VERTEX)(P4 X5 A8~
 X7 A6))
 () () () () ()

ACT *32: REC-SYS.

RULE RK11: A KNOWN GOAL INTERRUPT CAUSES RE-RECOGNITION.

GOAL-IS: SKO2

9,0 STM: (GOL O10(NAME TRIANGLE DIM TWO TYPE OBJECT)(P4 111 ~
 A12 113 114 A15 116 P25 133 A34 135))
 (V24(CR X26 PNUM P25 VSPEC (RL UL) SPEC TE TYPE VERTEX)(P25 X2~
 6 A31 X30 A29 X28 A27))
 (V3(CR X5 PNUM P4 VSPEC (RTA DL) SPEC V2 TYPE VERTEX)(P4 X5 A8~
 X7 A6))
 () () () () ()

ACT *1: AA-SYS.

RULE SKO21: START LOOKING FOR KNOWN OBJECT INFO.

DIR IS UL

RAN IS NIL

EXP IS ACA

ATT IS DL

VI IS NIL

49b

Figure IV.15

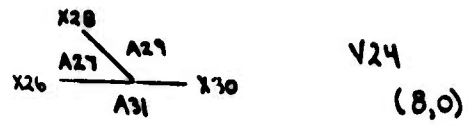
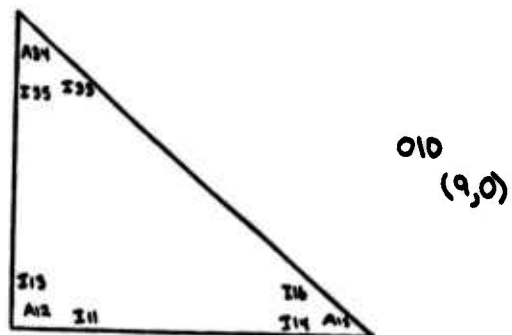


Figure IV.16



Figure IV.17



chunk, associated with the altered "corner". The vertex image, V24, is entered into STM from VI. The image body elements A15 (of O10) and A27 (of V24) are linked by equivalence property KEA (Known Equivalent ANGLE).

Furthermore, rule KO24 changes the current goal to RK1 (Rerecognize, Known Interrupted), sets CR of GOL to I16, and activates REC-SYS to aid in image accommodation. With rule RK1, rule RK11 (of REC-SYS) fires. This rule first sets CRMP to RMN1. The rule then traverses the partial object image of the GOL-appended chunk (O10) commencing at I16 (CR of GOL). CRMP accordingly is moved through the discrimination net of RLTM. In this case, RMN3 is the node referenced by CRMP. The associated UQX property yields the process TR3 which finishes a hypothesized triangle image. This is indicated by STM "9.0" (Figure IV.14) and is illustrated by Figure IV.17. Since a known object is again suggested, REC-SYS sets the goal to SKO2 and deactivates.

Rule SKO21 (of AA-SYS) fires, setting Visual Register values according to the accommodated object (triangle) image as described above, changing the goal to KO2, and activating VI-SYS. The trace of Figure IV.18 indicates that rule V18 fires as a result of the straight angle on the ATT side of vertex number 8. VI-SYS continues to move CPP in the DIR direction, reaching vertex number 2. Rule V111 fires, which completes the iconic image in VI that is illustrated by Figure IV.19. EXP is set to CON, which indicates that the expected angle feature was found but is part of a complex vertex (more than 2 exit directions). A QUICKSEE element (Q43) is used to represent vertex 8 in the iconic line image.

Figure IV.20 returns again to the next segment of protocol. The subject also reaches vertex 2 as a result of the hole movement. He indicates expectancy

90a

Figure IV.18

ACT #35: VI-SYS.

RULE VI8: ATT & EXP SPECIFIED AND STRAIGHT ATT SIDE.

AT VERTEX VER8.

MOVE THROUGH VER8.

RULE VII1: EXP & ATT & END OF LINE.

AT VERTEX VER2.

DIR IS UL

RAN IS LS

EXP IS CON

ATT IS DL

VI IS (L37 P42 P49 E38 I40 I41 Q43 I44 I39 I45 I46 E47 V48 P49 X50 A55 X54 A~
53 X52 A51)



Figure IV.19

Figure IV.20

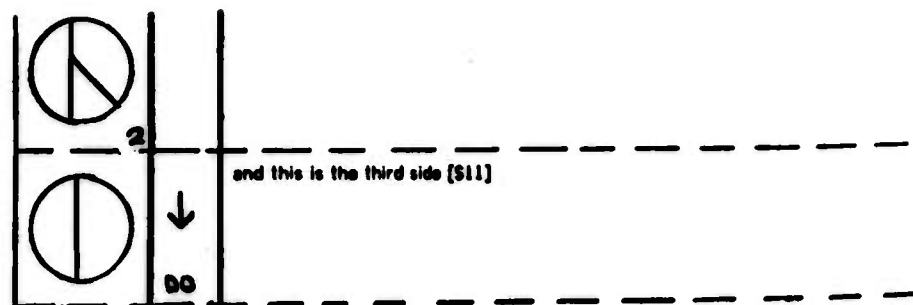


Figure IV.21

ACT #1: AA-SYS.

RULE K028: EXP IS PART OF VERTEX WITH SIDE.

GOAL-IS: K02
 14,0 STM: (V48(CR X52 PNUM P49 VSPEC (UO DR) SPEC TE TYPE VERTEX)~
 (P49 X50 A55 X54 A53 X52 A51))
 (L37(CR X52 VSPEC DIAGONAL TYPE SIDE)P49 P25 P42 P49 E38 I40 ~
 I41 Q43 I44 I45 E47))
 (GOL O10(NAME TRIANGLE DIM TWO TYPE OBJECTXP4 I11 A12 I13~
 I14 A15 I16 P25 I33 A34 I35 P49))
 (V24(CR X26 PNUM P25 VSPEC (RL UL) SPEC TE TYPE VERTEX)...)
 (V3(CR X5 PNUM P4 VSPEC (RTA DL) SPEC V2 TYPE VERTEX)...)
 () () ()

DIR IS DO

RAN IS NIL

EXP IS RTA

ATT IS RT

VI IS NIL

ACT #56: VI-SYS.

RULE V19: EXP & ATT & NO STRAIGHT SIDES FOUND.

satisfaction and continuation of the current goal (and object hypothesis) by statement [S11].

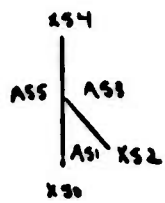
Figure IV.21 presents equivalent activity in the trace of VIPS. Rule K028 (of AA-SYS) fires due to the CON (CONTAINED) value of EXP and the LINE type of the iconic image of VI. The rule splits the image of VI into two parts, vertex V48 and line L37. The vertex image is entered into STM. This is indicated by STM listing "14,0" and illustrated by Figure IV.22. Image Body elements A34 (of O10) and A51 (of V48) being linked by the KEA equivalency property. The line image (L37) is changed into a SIDE type image chunk and entered into STM (see "14,0" and Figure IV.23). The END elements E38 and E47 are linked to INTERNAL elements I16 and I33 of O10 respectively. Position type element P49 is incorporated into O10, associated with the "corner" of I33, A34, and I35. Finally, the rule updates CR of GOL to reference I35 and loads the Visual Register according to the object image context before activating VI-SYS.

Figure IV.24 shows the last protocol segment considered here. The subject indicates completion of the visual confirmation of the proposed triangle object image by statements [S12] and [S13]. In frame V10 he moves to the right to begin another at the complex vertex (number 3) which he knows is there (VERTEX chunk V24).

Correspondingly, rule V19 (of VI-SYS) sets CPP to VER1, constructs the iconic vertex image V57 in VI, and sets EXP to YES before deactivating (see Figure IV.25). As a hypothesized object image is being visually confirmed, AA-SYS accordingly marks Image Body elements (a MARK property is set to YES). Thus, rule K027 (of AA-SYS) fires with the completion of O10 confirmation, as EXP is YES and I13, being (LIX of (CR of GOL)), has been previously marked. The rule changes the active goal to I2D

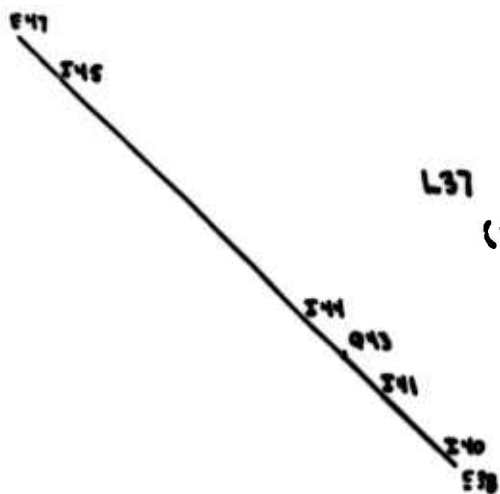
91a

Figure IV.22



V48
(14,0)

Figure IV.23



L37
(14,0)

Figure IV.24

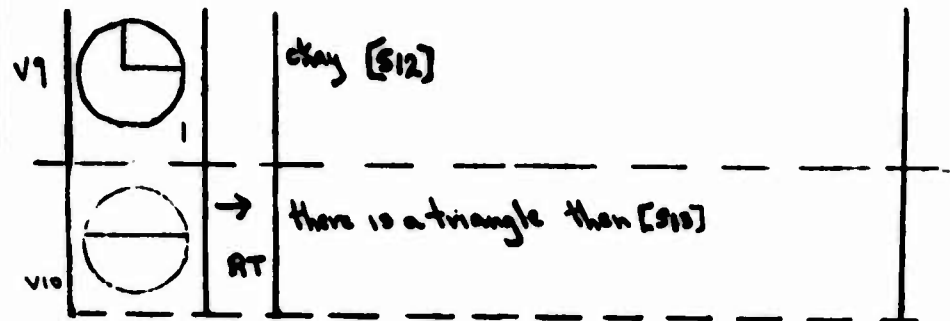


Figure IV.25

AT VERTEX VER1.

DIR IS DO

RAN IS LO

EXP IS YES

ATT IS RT

VI IS (V57 P58 X59 A62 X61 A60)

ACT #1: AA-SYS.

RULE K027: GOAL COMPLETED.

GOAL-IS: I2D
 15,0 STM: (GOL O10(NAME TRIANGLE DIM TWO TYPE OBJECTXP4 I11 ~
 A12 I13 I14 A15 I16 P25 I33 A34 I35 P49))
 (V48(CR X52 PNUM P49 VSPEC (UD DR) SPEC TE TYPE VERTEXXP49 X5~
 O A55 X54 A53 X52 A51))
 (L37(CR X52 VSPEC DIAGONAL TYPE SIDEXP49 P25 P42 P49 E38 I40 ~
 I41 Q43 I44 I45 E47))
 (V24(CR X26 PNUM P25 VSPEC (RL UL) SPEC TE TYPE VERTEX)...)
 (V3(CR X5 PNUM P4 VSPEC (RTA DL) SPEC V2 TYPE VERTEX)...)
 () () ()

(Incorporate 2 - Dimensional object) and activates INC-SYS. The contents of VI are not transferred to STM. The YES value of EXP is sufficient input in this case.

Rule 121 of INC-SYS fires to incorporate the object O10 of the GOL-appended chunk and its immediate visual context. Figure IV.26 indicates the before and after states of ITM and Figure IV.27 pictorially illustrates the new ITM contents. The incorporation is accomplished by removing GOL from the O10 chunk and copying the chunk into ITM. The image body of the chunk is then traversed, incorporating any chunk containing an Image Body element which is linked by equivalence to one in the incorporated object image. During this traversal, a new start location is also chosen. This is one of the context chunks, which has an unknown XIT element (type UEX). It is marked in STM by appending the Special element LAST to notify AA-SYS. If no chunk with an unknown link is found, the goal becomes NOMO (NO More Objects). In the current case V24 is marked with an appended LAST, and the goal becomes SNO (Start New Object) as shown by Figure IV.28.

This protocol - program trace comparison is continued in Chapter II.2 as section A.P1.

THEORETICAL IMPLICATIONS

The example protocol-program trace segment illustrates most of the theoretical implications imbedded in VIPS. The integrative and generative process of visual image construction based upon eye sensory input (perception) is illustrated. The use of current image contents to guide hole (eye) movement behavior and input interpretation is illustrated. The basic activity of assimilation-accommodation is illustrated. The proposed image representation is shown to be both adequate for and furthermore, suited to its role as a guide for perceptual activity.

Figure IV.26

ACT #63: IND SYS

RULE 121: INCORPORATE FIRST OBJECT

16.0 ITM

16.0 ITM #10048 A34 V24 A12 (TIME TRANSFER DIM TWO TYPE OBJECT)
 V24CR A33 PLUM P25 VSPED (PLUM SPEC TE TYPE VERTEX)
 L3 (CR E25 VSPED DIAGONAL TYPE SIDE)
 V24CR A30 PLUM P49 VSPED (UD CR SPEC TE TYPE VERTEX)

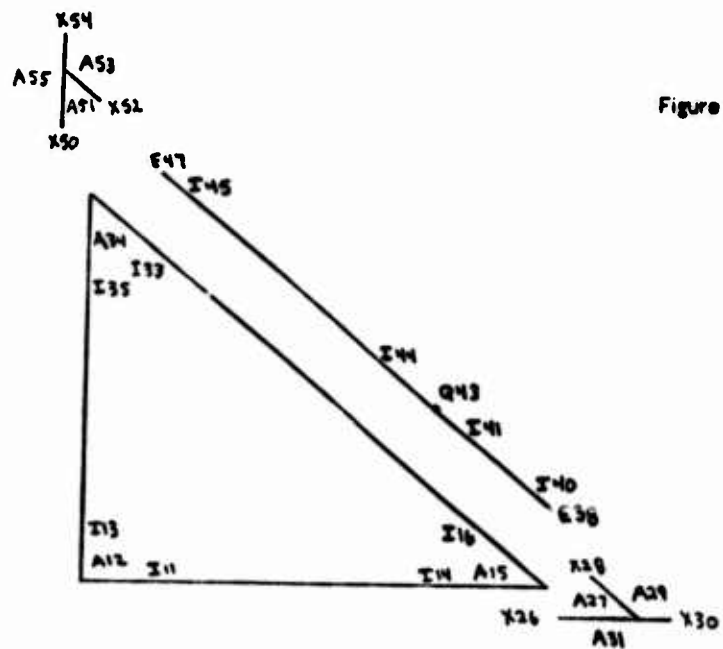


Figure IV.27

ACT #1: AA-SYS

RULE SMOO: START NEW OBJECT AT UNLINKED POSITION.

WILL START NEW OBJECT AT RSH DME.

DIRECTION IS RT

AT VERTEX VER3

Figure IV.28

GOAL RELATED EPISODES

The processes of VIPS (except VI-SYS) use the name of the current goal as the primary condition for (determinant of) rule firing. This implies an interpretation of the protocol (perceptual) behavior as consisting of a series of goal-related episodes. Figure IV.29 graphically illustrates this interpretation of the example protocol segment in the form of a "goal episode chart".

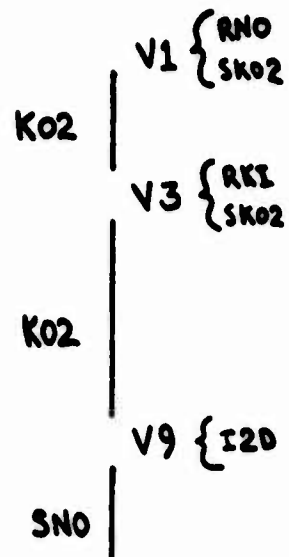
The chart indicates those protocol frames which correspond to (are concurrent with) the inferred goal transition points of the behavior (VI, V3, V9 and V11 of the figure). The vertical lines indicate a sequence of one or more hole (eye) movements. To the left of each line is the name of the current goal which is active during that hole (eye) movement sequence and associated behavior episode. To the right of each included frame number (bracketed) is listed any current goals which are active at that protocol point, but which become inactive (past) prior to commencement of the subsequent hole (eye) movement sequence.

Generally reviewing the example, goals RK0 and SK02 are associated with the initial REC-SYS activation by AA-SYS. Upon return to AA-SYS the goal is changed to K02 and hole-movement activity begins.

At frame V3 conflicting input causes AA-SYS to reactivate REC-SYS with goal RK1 in order to accommodate the object image. Goal SK02 results, which is changed by AA-SYS to K02 and hole movement again begins.

This goal remains active until protocol frame V9, when the hypothesized object has been completely confirmed. INC-SYS is activated with goal I2D. Upon deactivation of

Figure IV.29



RNO - Recognize New Object
 SKO2 - Start Known Object 2-dimensions
 KO2 - Known Object 2-dimensions
 RKI - Rerecognize, Known Interrupted
 I2D - Incorporate 2-Dimensional

INC-SYS the system moves to the new location with goal SMO. At frame V11, REC-SYS is activated to attempt recognition (suggestion) of a new object (goal RNO).

The goal episode charts are a combination of protocol information (frame numbers) and trace information (inferred current goals). These charts illustrate the inferred factorization (partitioning) of the observed behavior and, as such, an important framework for the evaluation of protocol-trace correspondence. These charts are used throughout the thesis as a condensed representation of subject and system behaviors and their correspondence. The evaluation section of Chapter II.2 uses this episodal interpretation as a basis for judging the explanatory ability of VIPS.

Chapter 1.5 VIPS Extensions

VIPS is limited in scope due to the specific, task - related behavior it represents and explains. Several memory extensions and generalizations are necessary to complete the theoretical definition of the visual (cognitive) processor. Another component of Long Term Memory is defined and the content and function of LTM of VIPS are generalized and extended. The image representation and processes involved in three dimensional visual form perception must be described. A representation generalization is necessary to define a more universal visual imagery ability. Finally, the processes of VIPS must be generalized to explain the eye-movement and peripheral vision factors of visual perception.

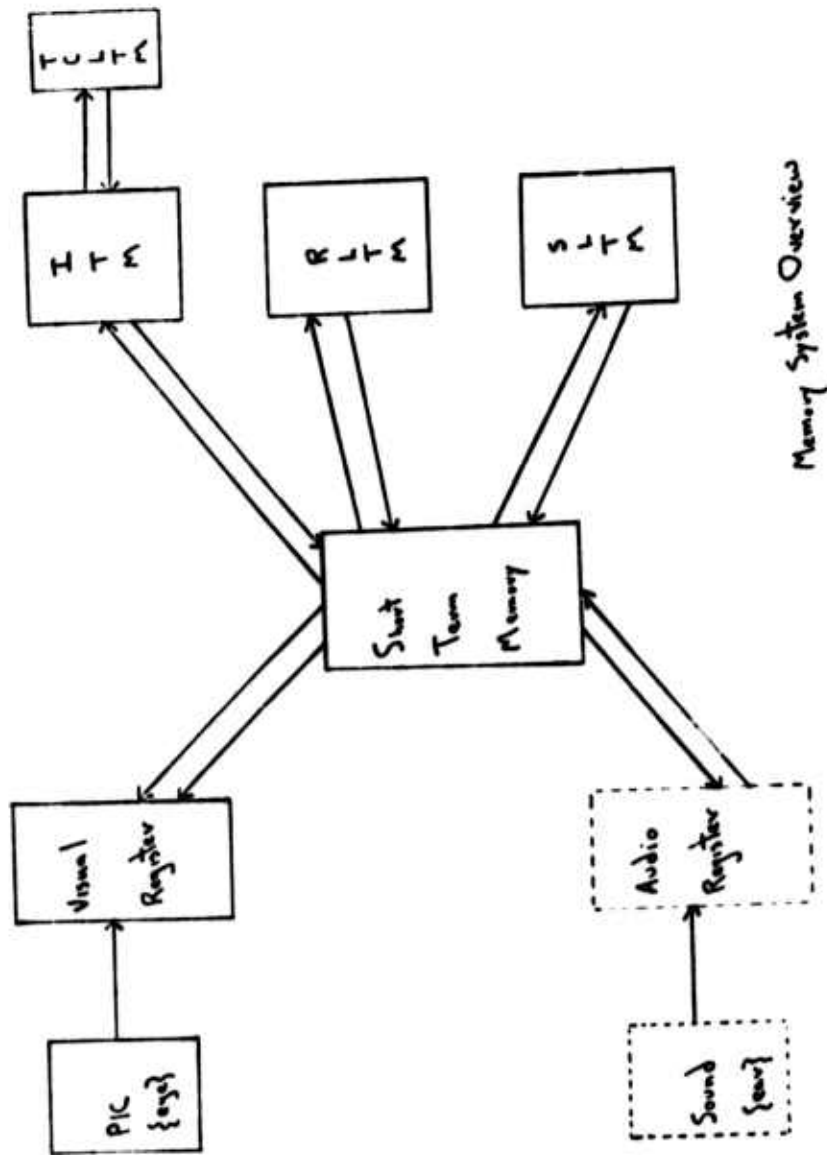
1. Memory Extensions

TCLTM (Temporal - Contextual Long Term Memory)

Temporal - Contextual Long Term Memory (TCLTM) is the component of long term memory which serves as a permanent repository for the episodic and contextual information representations (images) which are initially constructed as the contents of ITM (see figure V.1). Therefore, the elements of TCLTM are images or other integrated (interrelated) structures of selected (meaningful) past chunks of STM.

Each elemental image of TCLTM has an associated label element. This label affords

Figure V.1



the two basic means of access to the memory image's content. For one, the label element has associated properties which specify information as to the time and place associated with the initial realization of the image in ITM. These temporal-contextual information features (Shiffrin and Geisler, 1973) may correspond to properties of a later, current situation specification. This correspondence affords access to the label element. Each label element also has properties which reference content elements within the associated image structure. These references afford access to the memory image's content after access to the label element by temporal-contextual correspondence has been accomplished. The content referencing properties provide the other means of accessing the contents of the memory images of TCLTM. Just as the features of the temporal-contextual information may bear correspondence to current STM contents, content elements of the image structure may also be equivalent to elements of current STM and facilitate that image's retrieval.

In the extension of VIPS, the final image representations (perceptions) of ITM could become TCLTM elements. Prior to VIPS deactivation, by a final activation of INC-SYS, a label element can be created. The final values of the FIRST-ITM, LAST-ITM, FIRST-OB and LAST-OB reference properties of ITM could then be transferred to the label element as content reference properties. Situational properties (temporal-contextual information) could then be specified for the label element, also. Upon VIPS deactivation and the accompanying disengagement of the cognitive processor from perceptual activity, the perception (image) can become an element (the most recent) of TCLTM.

Retrieval of TCLTM image elements is accomplished by an active, searching process (Shiffrin and Geisler, 1973). The search attempts to locate an image with temporal-

contextual or content features equivalent to those provided to the process from the current contents of STM (a search-set). Though an element is entered into TCLTM permanently, forgetting may occur as a result of the failure of this retrieval (search) process.

Accessibility of an image element of TCLTM decreases as a function of element age (Shiffrin, 1970; Tulving, 1970). With the equating of image accessibility to trace strength, Wickelgren (1972) provides further support. With the passage of time, features of current temporal-contextual information change. Thus, the longer an element is a member of TCLTM, the less similar is its temporal-contextual information to that of a current search-set (Shiffrin, 1968; 1970). With regards to retrieval by content correspondence, the probability of a TCLTM element with similar (equivalent features) content being in TCLTM increases with elapsed time. The existence of such an element(s) decreases the accessibility of the initial by being a source of retrieval competition, confusion, and interference.

Recent episodal and contextual information (memory images) play an important role in many (most) experimental tests of memory. Due to the capacity limitations of STM (Miller, 1956), most typical tests, such as the recognition of a current stimulus as an element of a previously presented set, necessarily use a recalled element of TCLTM. This image is readily retrievable as it is usually a recent or the most recent entry of TCLTM, or may still reside in ITM.

When an element of TCLTM is retrieved, it becomes the current image of ITM. From ITM, the structure is recalled into STM in part or as a whole (chunk by chunk) for comparison with new current STM contents. Such comparison is template-like in nature being an attempt to directly match possibly equivalent image structure schema.

Such equivalency would depend upon correspondence of image element properties which specify visual features and spatial interrelationships. During perception if a match to partial input succeeds, the retrieved image is used to guide the perceptual behavior as are newly created (hypothesized) images.

Statement [S12] of Figure V.2 indicates one of the rare instances of this behavior within the protocols. This scarcity is due to the discrete (discontinuous) nature of the experimental visual field which consisted of a series of distinct (only generally related) line drawings. In more usual continuous context situations, REC-SYS retrieves images from TCLTM (and ITM) with greater frequency. TCLTM was not implemented in VIPS due to the rare use of its visual image elements in the protocols.

The use of retrieved TCLTM images to guide the perceptual activity can account for the results of Noton and Stark (1970). Their results indicated that subjects use similar eye-movement sequences during the recognition of a previously presented drawing. This indicates use of the visual image result (perception) of prior behavior in a subsequent perceptual situation. Furthermore, Noton and Stark propose that a perceptual (visual) image's content strictly reflect (include) its generating eye-movement behavior. This is not proposed here. It is proposed only that the sequence of input views influences image contents, as do perceptual goals. In later perceptual use, these influenced image contents inversely affect the eye movement behavior.

SLTM (Semantic Long Term Memory)

Semantic Long Term Memory (SLTM) is an extension of the LTM implemented in VIPS. It is thus not a new memory component, but replaces LTM in the final memory system proposal (see Figure V.1). SLTM is a semantic net (Quillian, 1968) of a

conceptual-relational nature (Kintsch, 1972; Bower and Anderson, 1973). SLTM holds the system's symbolic representation of abstract, generalized, factual knowledge. As such, it contains (as activatable sub-structures) associative, heterarchical representations of visual information (Winston, 1970).

Similar representations of visual information have proven adequate to realize and explain behavior on visually-related, cognitive tasks (Baylor, 1972; Moran, 1973). Those tasks are of a more abstract nature than the perceptual activity represented by VIPS. Segments of the more abstract contents of SLTM become partitioned and activated as chunks of STM (Moran, 1973). As chunks of STM, these segments of abstract "facts" determine (guide) the subsequent visually-related cognitive activity.

The contents of SLTM and RLTM are altered by learning process based upon (accommodating to) current experience (STM contents). EPAM investigations suggest RLTM alteration means. Winston (1970) investigates a means for the updating (learning) of the visually-meaningful contents of SLTM.

2. Three-Dimensional Picture Perception

Picture perception differs from other visual perception activity as the visual environment is two dimensional. A presented line drawing is known to be of two dimensions as an object itself. That some straight line drawings are perceived in three-dimensional terms (with depth) is a prime example of what E. H. Gombrich refers to as the "beholder's share" of visual picture perception in his most interesting book, *Art and Illusion* (Gombrich, 1960). The three-dimensional perception of pictures occurs in the absence of the positive efferent indications of accommodation and convergence.

What are the basic aspects of the processes and memory representations which

become involved in the development of a three-dimensional perception of a presented straight line drawing? What characteristics of a line drawing lead to its interpretation and representation in three-dimensional terms?

RELEVANT PROPOSALS

A prime target for attempts at quantifying Gestalt organizational laws has been to explain the perceived three-dimensionality of certain line drawings. Their belief is that the basic organizational laws of perception will attempt to realize a description of minimum complexity for a presented stimuli ("Pragnanz"). As such, a "simple" three-dimensional perception of a line drawing results as favored over an otherwise necessarily "complex" two-dimensional perception. Several studies by Hochberg (Hochberg and Brooks, 1960; Hochberg and McAllister, 1953) resulted in the development of a complexity measure for straight line drawing stimuli. This complexity measure demonstrated significant positive correlation with the relative perceived three-dimensionality of various presented stimuli sets. Attneave and Frost (1969) carried the investigation of "minimum principle" implications further. Their experimentation produced results which indicated that the perceived slant in depth of a tridimensional perception was highly predictable from basic simplicity criteria.

These studies all considered the "minimum principle" to apply to the presented line drawing as a whole. This was necessary as Gestalt theory offers no explanation for the integration of successive "partial wholes" of a stimulus to realize its perception. If such a "minimum principle" were the only source of three-dimensional hypotheses and the resulting perceptions, it is predicted that mention of three-dimensionality in the hole-moving task protocols will occur only after the whole line drawing has been

traversed at least once. Yet, subjects definitely mention three-dimensional ideas with only partial picture information available. This is clearly seen in the three opening protocol segments presented as Figures V.2, V.3, and V.4.

From these three initial protocol segments it is apparent that local picture information (even one vertex) is sufficient to illicit attempts to realize a three-dimensional line-drawing perception. Simon (1967) noted the ability of an attention shift to flip the perceived orientation of a presented Necker Cube stimulus. Hochberg (1968) noted the ability of local depth cues to generate proposed three-dimensional interpretations. The artists M.C. Escher and Hogarth (See Gregory, 1970) were both adept at creating "impossible" pictures through a combination of inconsistent localized depth cues (See Figure V.5 for an example). Gregory (1970) investigates several local depth cues and proposes that their perceptual effect results from the memory of past experiences involving factual - visual coordination.

A THEORY

Since a line drawing is itself a two-dimensional object, the initial perceptual set is naturally toward the realization of a planar or two-dimensional perception (visual image representation). This perceptual assumption remains in effect until a so-called "trigger" vertex (or feature configuration, in general) is encountered. A "trigger" vertex is one having three exit directions, no two being in oppositely oriented directions (in other words, "T" vertices are not included). No "trigger" vertices are encountered by subjects in the four protocols discussed in Chapter 7. No mention of any possible three-dimensional interpretations occurs in those protocols either. In all three initial protocol segments of Figures V.2, V.3, and V.4, a "trigger" vertex

101a

Figure V.2

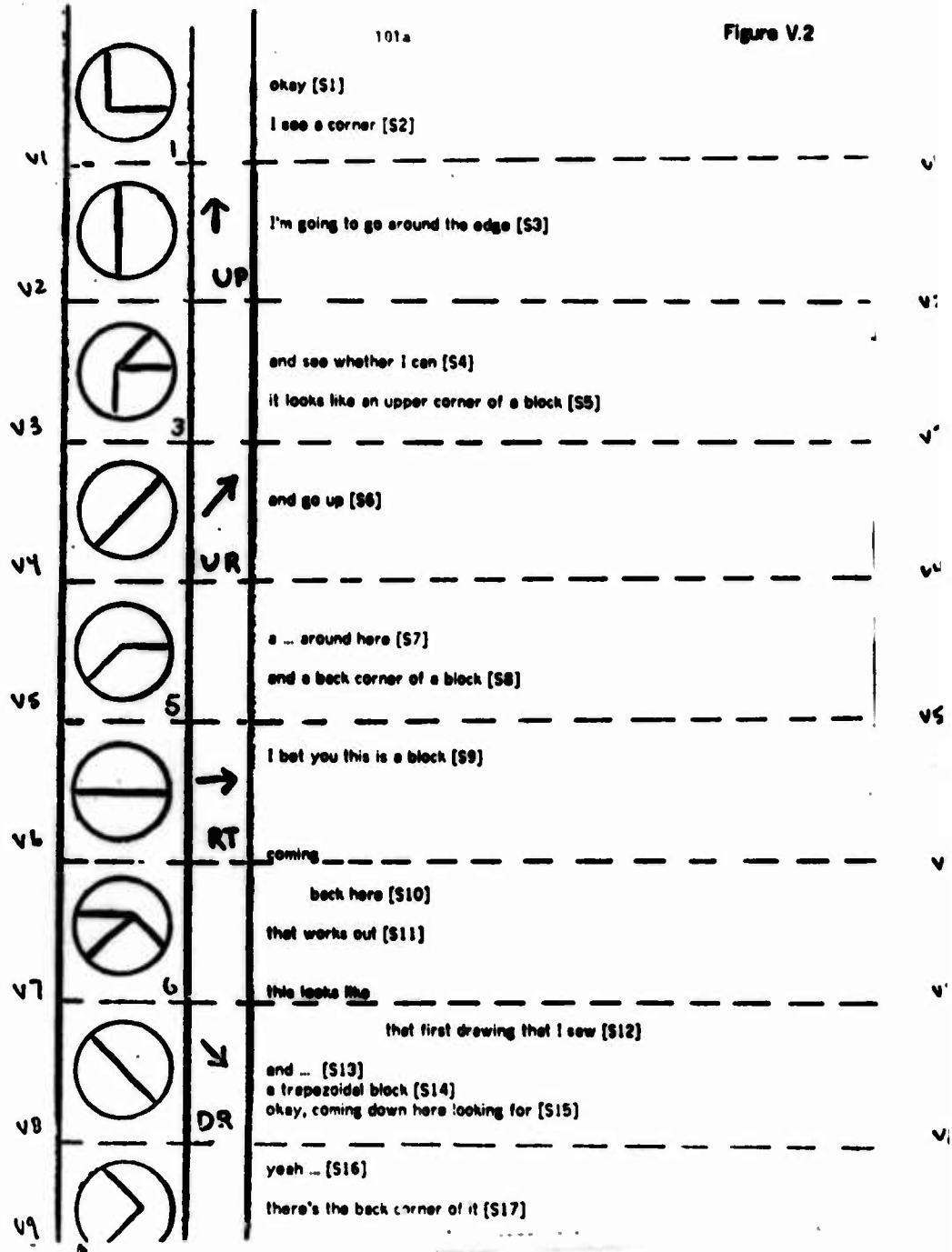
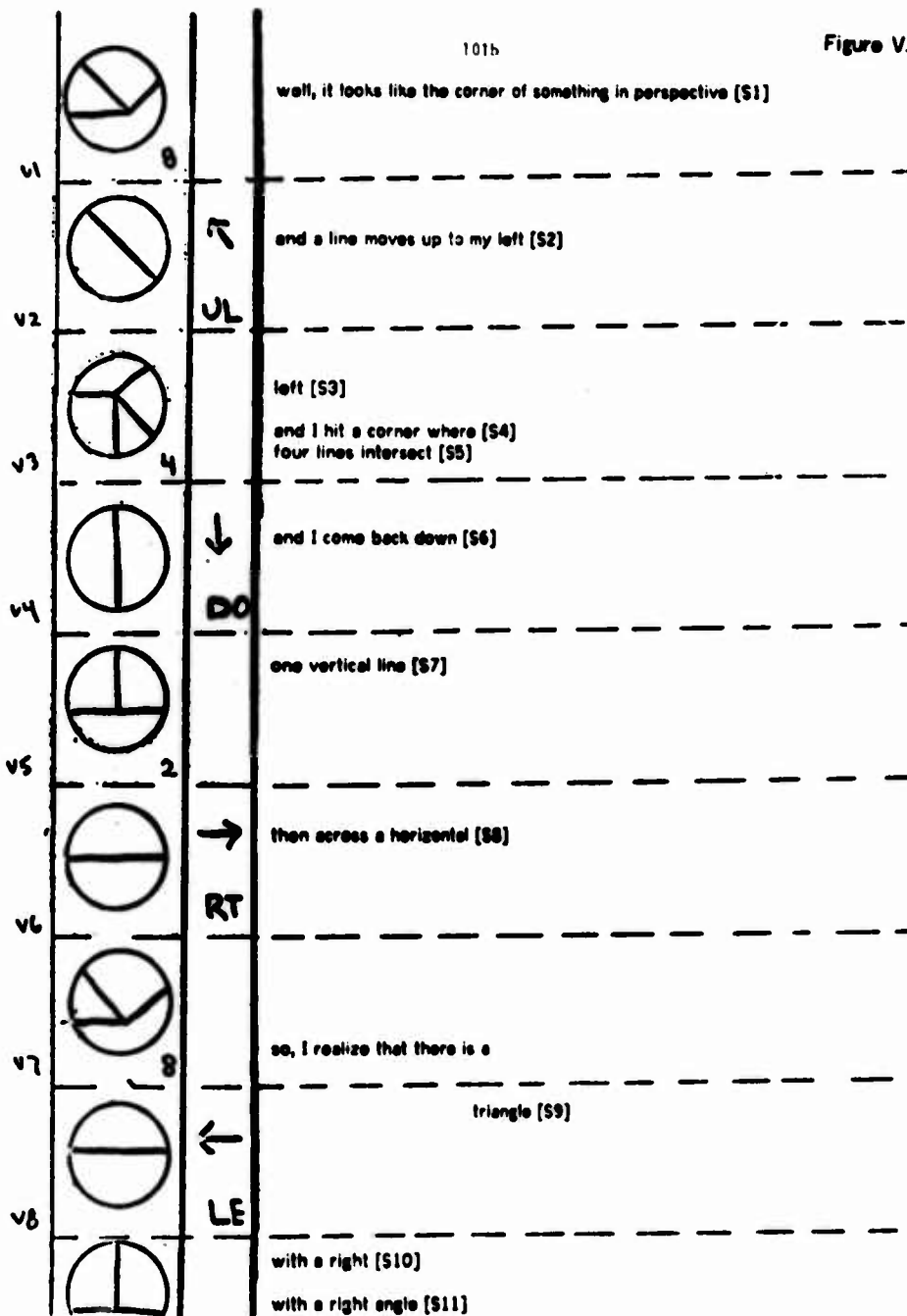


Figure V:



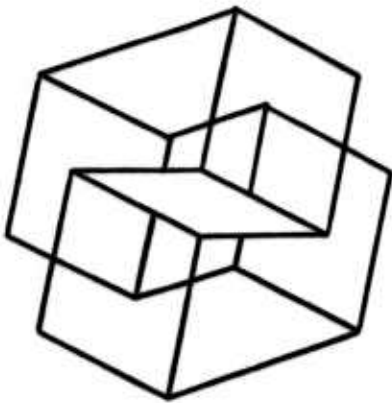
101c

Figure V.4

v1		R	oh! [S1] this looks like um ... [S2] the beginnings of a solid object [S3] could be a cube um ... [S4] let's see, which way should I go [S5]
v2		UR	I'll try going up here [S6]
v3		7	she ... [S7]
v4		UR	she ... [S8] it looks like a cube still [S9]
v5		6	
v6		DL	see which way this comes [S10]
v7		4	
v8		DL	definitely has the makings of a cube [S11]
			

101d

Figure V.5



© M. J. ALBERS

"Structural Constellation" c. 1954

consistently elicits immediate three-dimensional considerations. The perceptual effect of a "trigger" configuration is consistent with the experience-based theory (Gregory, 1970) and with the local application of the Gestalt "minimum principle".

In the usual case, two vertex exits of the trigger vertex are bound to an object face, the "prime face", which is perceived to be parallel to the picture plane. The third exit direction is perceived to be extending forward or backward in space (depth) and being the edge between two "secondary" faces. The phrase "in the usual case" was included in the description of "trigger" vertex perceptions above. Note that when the perceived depth direction is down, the usual interpretation is rejected and the other two vertex exit directions (Still of the prime face) are assigned depth meaning. This may be an inherent result of the "support hypothesis", an assumption resulting from the normal inability of an object to hang in space. Support is usually from below, denying down as a viable depth direction in physical space with gravitation. Figure V.6 shows a "solid octagon" with all eight secondary face extension directions illustrated. The starred ("a") drawing shows the faces extending down. In this case these faces are not perceived to extend in depth as usual. The octagon (prime) face is perceived as the object top and extends in depth itself.

The decision to consider a presented picture in three-dimensional terms is a primary perceptual decision. This decision results in the application of a set of perceptual goals and associated object hypotheses of currently active goals and associated object hypotheses and differing from that realized in conjunction with two dimensional considerations. Aubrey Beardsley, in his ornate art-nouveau drawings, contraposes complex border patterns and complex pictorial scenes, both often sharing equivalent components (See Figure V.7). These illustrations afford introspective

102a

Figure V.6

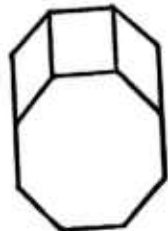
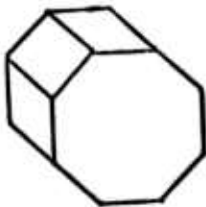
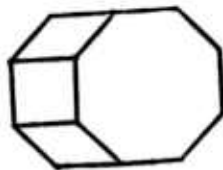
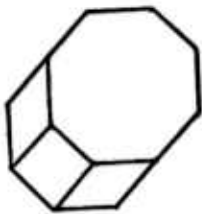
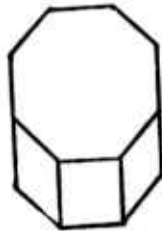
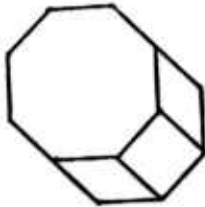
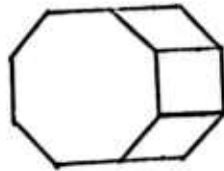
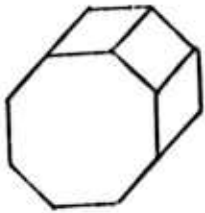


Figure V.7

102b



Merlin Taketh the Child Arthur into His Keeping.

appreciation of the effect of the pattern or three-dimensional interpretation decision upon the visual experience. Though similar visual elements occur in both border and picture, they are experienced differently due to the difference in assumed dimensionality.

Now we must consider the activity of the visual processor after encountering a "trigger" vertex. At the strategic level, the goal K03 becomes active, which will attempt to recognize a three-dimensional prism-type object. Such objects are recognized in terms of a prime face, which serves to name the object, and several associated secondary faces, which usually are perceived as extending in depth.

As with the two-dimensional case, an image representation of the object is constructed in STM. Each face is represented in a different chunk. Chapter 13 discusses the VIPS image representation of three-dimensional prism-type objects in conjunction with the definition of the FACE type chunk. These image chunks serve as the source of tactical information to generate visual feature expectations and hole (eye) movement specifications. An assimilation-accommodation process equivalent to that of the implemented system is used to either confirm expectations or alter and abort the object images (as conflicting information is encountered).

PROTOCOL INDICATIONS

The action of assimilation-accommodation is well illustrated by the initial protocol segments of Figures V.2, V.3, and V.4. In Figure V.2, a block (cube) image is hypothesized as a result of view V3 (Statement [S5]). That image is accommodated to a trapezoidal block image at V7 (Statement [S14]), due to the conflicting external visual input of that view.

In the protocol segment of Figure V.3, commencing at a different vertex location, the initial three-dimensional interpretation is dropped due to conflicting inputs encountered in subsequent vertex views. This is a most severe accommodation due to the primacy of the dimensional decision as noted.

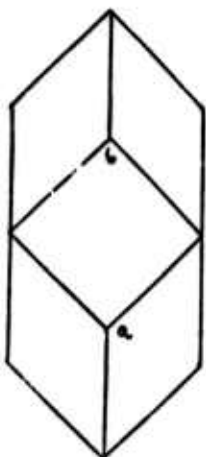
In Figure V.4, the subject scans the prime face of the proposed cubic image, so far accepting (assimilating) the extra vertex exit encountered at V7 into the three-dimensional perception.

SPATIAL CONSISTENCY

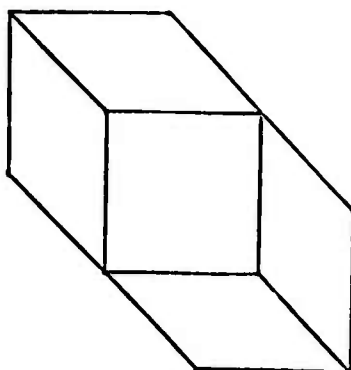
There is one completely new perceptual factor which enters into the realization of a three-dimensional image representation and perception. This new factor is that of spatial consistency. The determination of spatial consistency involves the consideration of two factors. One is the constancy of depth direction, within the image as a whole and more importantly for any given line segment within the image. The other is the consistency of the in-depth locational interrelationships between the various feature groupings (vertices) of the image. Chapter 13 discusses the inherent ambiguity of any three-dimensional interpretation of a planar surface. The perceptual system must construct (by choice) a non-ambiguous, consistent visual image.

Figure V.8 illustrates the two possible inconsistency factors in example line drawings. All of the examples presented there are inconsistent when perceived as two interacting cubes. Example ii involves only the inconsistency of perceived depth direction for the picture as a whole. Example i involves both inconsistency of depth direction for the whole picture and inconsistency of perceived vertex locations. This involves the added irregularity in the perception of prism projections when the facial

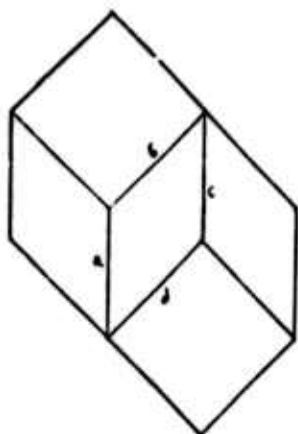
Figure V.8 104a



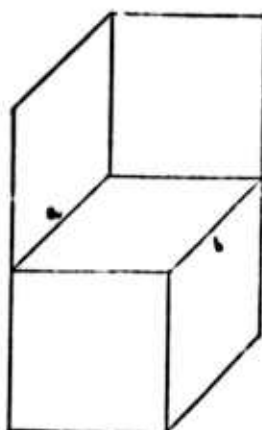
i



ii



iii



iv

extension direction is down. Thus, when considering the top cube, b is above a, while when considering the bottom cube, b is behind a. Note the introspective "feeling" of bend of the square prime face as the differing in-depth directions act to warp that face.

In examples iii and iv, a face is shared which is not the object's prime face. In figure iii, the special case with down is again illustrated. As with example i, just discussed, the edges a, b, c, d of the shared face are inconsistently perceived to be extending in-depth in one (the top) object related perception, and to be parallel to the picture plan with regards to the other cube perception. Again inconsistencies of spatial locations result. Example iv involves the inconsistencies of perceived depth directions for line segments a and b and the associated reversible depth relationships perceived between the vertices at their ends.

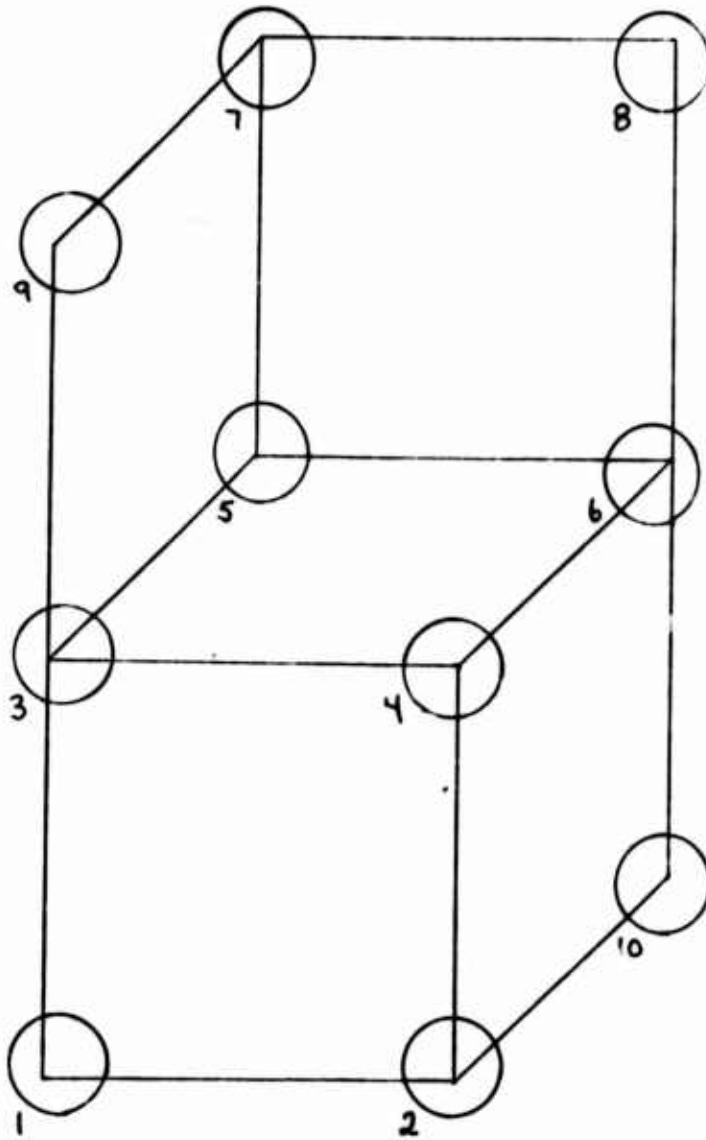
FURTHER PROTOCOL INDICATIONS

To illustrate the effects of inconsistency upon the visual process, three protocols are presented which show subject hole-movement behavior upon Example iv (see Figure V.9). Each appendix includes a fully transcribed protocol, an active-goal episode chart, and a pictorial representation of the final ITM contents as inferred. The active goal episode charts are presented as figures here and indicate the possible course of VIPS implementation to account for these protocols. The complete appendices are not included in this report but are available from the author if this detail is desired. Dots occurring in the goal episode chart indicate episodes of activity not straight-forwardly explained.

Consider protocol AP3 whose goal episode chart is Figure V.10. The subject

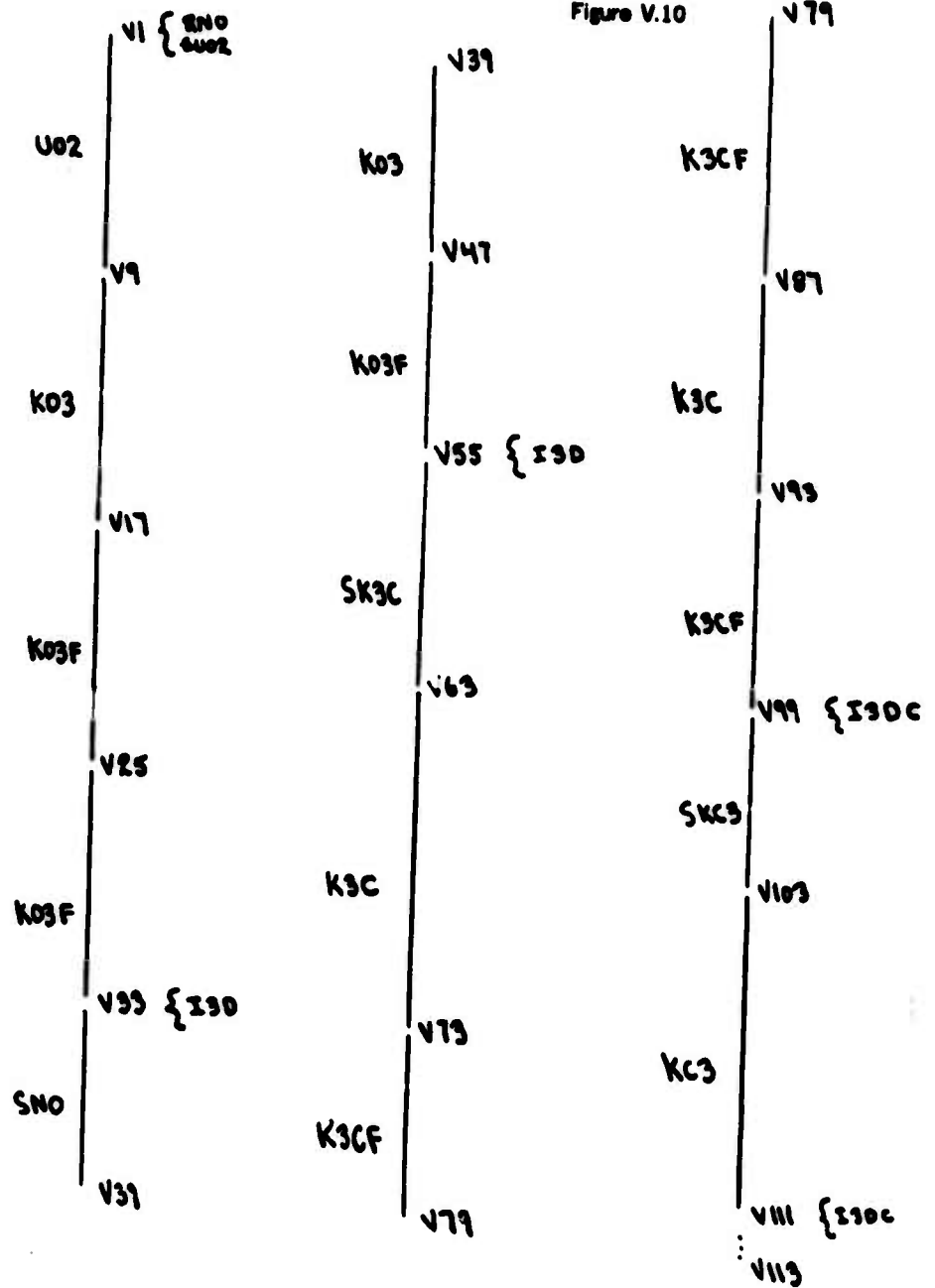
105a

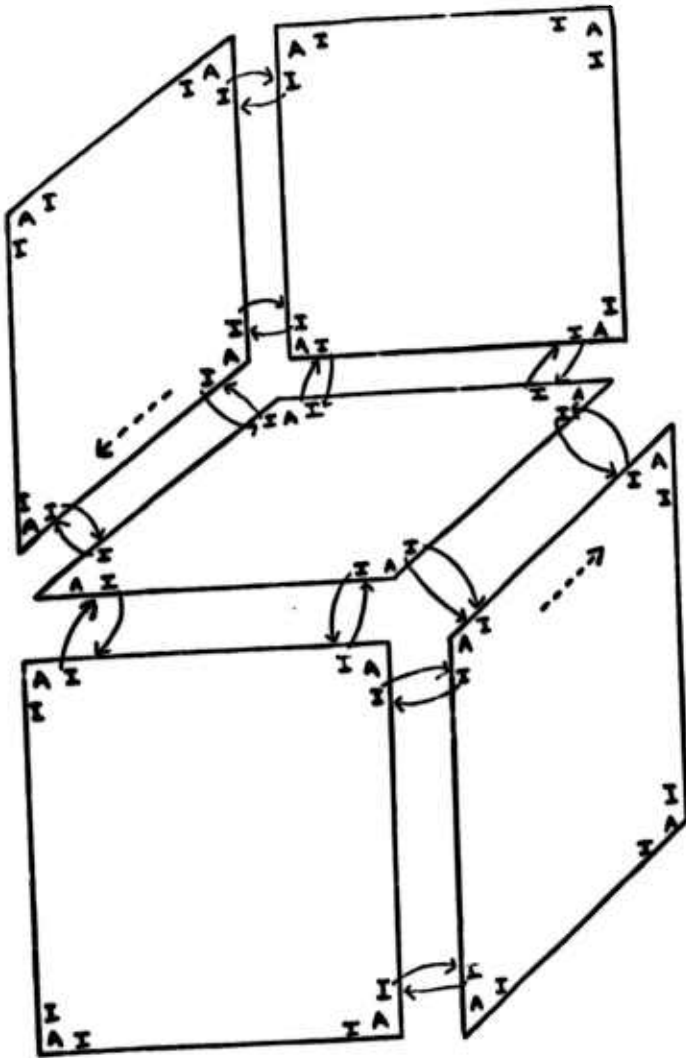
Figure V.9



105b

Figure V.10



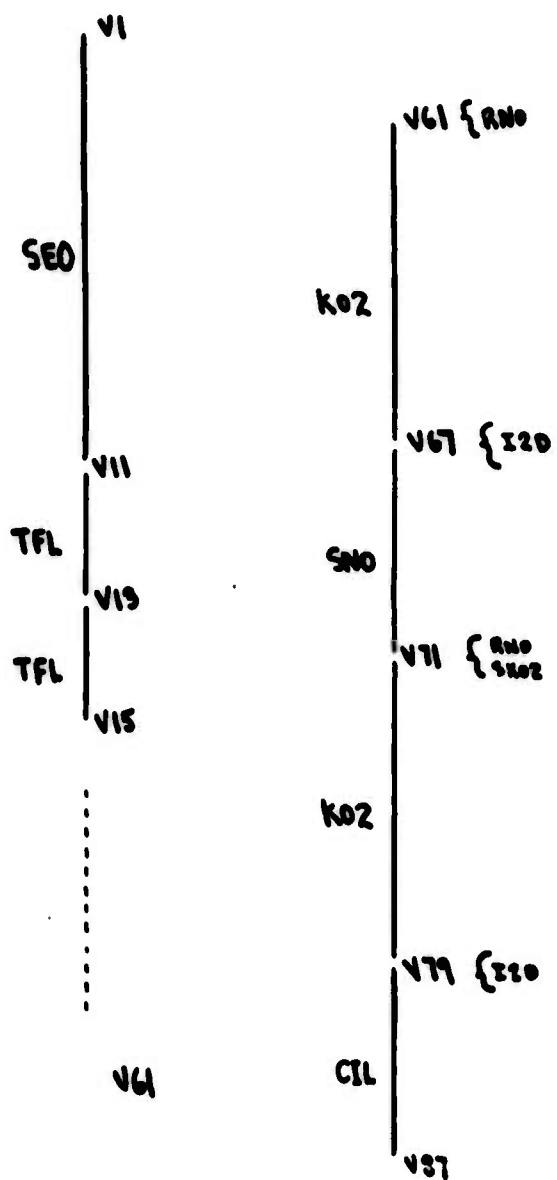


attempts to recognize a two-dimensional figure, with goal U02 (nothing is suggested by the partial information) until V9, seeing vertex 5. This "trigger" vertex generates the three-dimensional goal K03. This goal and the accompanying cube image guides behavior until V33. This results in the full external visual scan and confirmed perception of the proposed cube. At V39, by statements [S44] through [S46], the subject indicates the realized perception of the cube and statement [S49] indicates attention is now to be directed to recognizing what is below that cube.

At V41, the subject sees vertex 4, another "trigger" vertex and statement [S51] indicates that another cubic goal has been generated for consideration and confirmation. This guides behavior at the strategic level until V55, where statements [S65] and [S66] indicate that the incorporation process is having difficulty due to the inconsistencies of depth direction and spatial locations of the two cube images. The behavior from this point on is an interaction between the incorporation (INC-SYS) process and the AA-SYS process in an attempt to resolve inconsistencies. Reconfirmations of the two separate cubes are realized, but the inconsistencies are never resolved, as shown by the pictorial representation of the final ITM contents (Figure V.11). The dotted arrows indicate the conflicting depth directions perceived. This representation allows the verbalized report of two cubes, but fails to be satisfactory for reproductive drawing generation, as shown by the appendix.

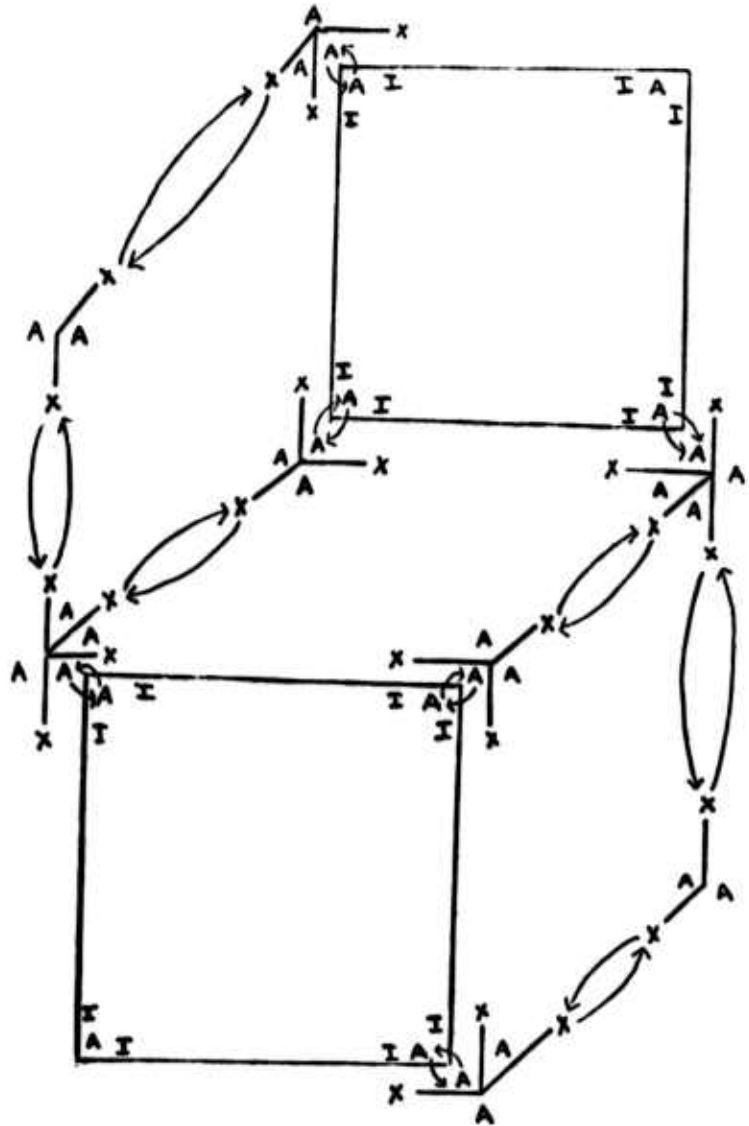
Considering protocol A.53 next (see Figure V.12), the initial goal here is to search the outside. The "trigger" vertices don't overcome this goal, which is active until V11. At that point, the subject moves inside, possibly in an attempt to link up with vertex 7, which he has seen. When this fails he moves to the outside, linking this inside vertex with an outside chunk he has kept. From view V15 to V61, an exact goal structure is

Figure V.12



106b

Figure V.13



to induce. The subject generally stays to the outside contour, constantly noting various three-dimensional interpretations of these vertices. Though the word cube is mentioned, there does not appear to be any specific cubic images generated for comparison. The apparent inconsistencies of these possible cubic corners is noted in comment [S72] at V33. Thus, it appears the inconsistencies of depth directions generated by the trigger vertices are noted without constructing complete inconsistent images, or at least without attempting to first confirm these cubic images.

At V61, the subject finally decides to consider the picture in "local" terms (elements [S111] to [S113]). This is equivalent to a decision to perceive the picture in two-dimensional terms, instead of the three-dimensional concepts which appear to be inconsistent. This is an indication within this protocol, of the fundamental nature of the decision. From V61 on the subject operates under the goals to recognize two-dimensional objects and to incorporate them into ITM with their known pictorial environments. This interpretation is indicated by the active goal episode chart of the appendix. Indeed, with only slight modification, the current implementation could be used to generate behavior representative of that following V61. This activity results in the creation of the final ITM representation as displayed by Figure V.13, which shows correspondence with the verbal and drawn descriptions obtained from the subject.

One final protocol is considered, AB3 (see Figure V.14). Again the initial goal here is to search the outside (SEO). Upon completion of this, the subject recognizes the image as that of a cubic projection, with extra internal lines. This image is incorporated into ITM. The subject then directs attention to begin linking up the inside directed segments. The trigger vertices encountered now do result in the three-dimensional inside

Figure V.14

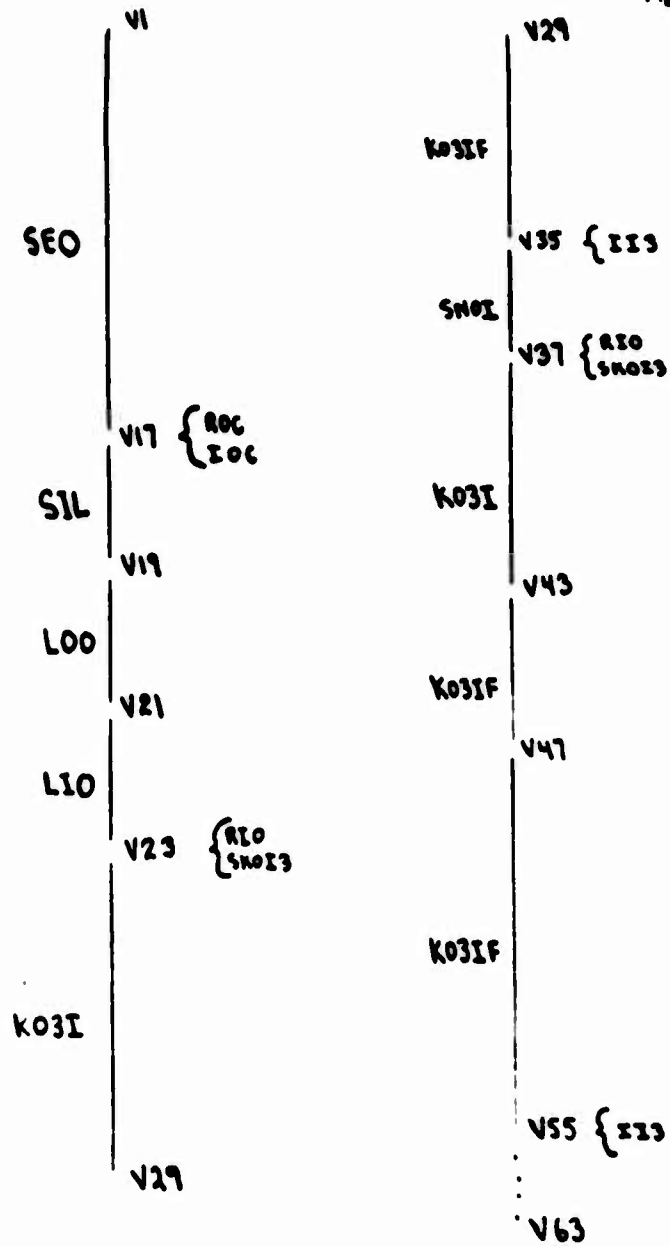
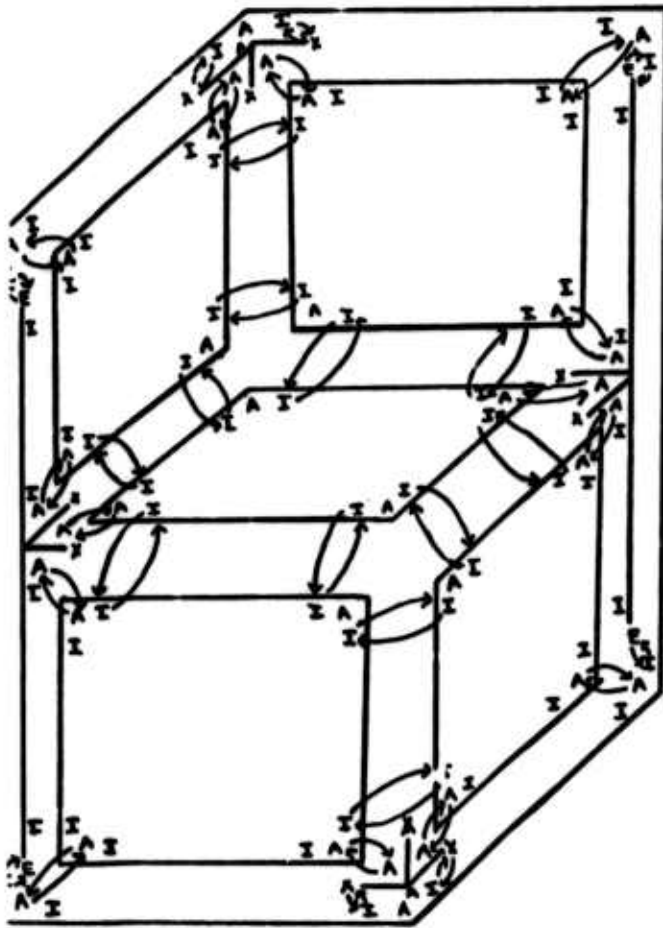


Figure V.15



object proposals. A series of perceptual strategic goals becomes active, as shown in the episode chart, which results in the scanning and perceiving of the two internal cubes. Again the final description is not free of ambiguity or inconsistency (see Figure V.15). The fact that the ITM representation contains the outside contour image allows the subject to anchor the drawing within this consistent two-dimensional framework and achieve a correct drawing. The verbalizations during drawing and the sequence of drawing (outside contour first) favor the inferred representation.

Considering these three protocols in full, one is afforded a greater appreciation of the ability of the proposed implementation (perceptual theory) to assimilate the extensions necessary to generally account for the perception of line drawings in three-dimensional terms. The flexibility of the proposed image representation is well illustrated by the examples, also.

3. Generalized Visual Image and Perception System

What are the extensions and modifications necessary to translate VIPS into a system representing, and thus explaining, the complete human visual information processing system? First of concern is the image representation itself, in terms of what alterations and additions must be made. Given these basic imagery (representation) changes, the necessary modifications to the proposed visual processes to accommodate and use these alterations can be considered.

IMAGE GENERALIZATION

The image representation must be altered in two fundamental ways. First, it must be sufficiently generalized to be capable of embodying all possible (probable) visual form perceptions, not just the restricted environment of straight line drawings. Secondly, the image representation requires an alteration or addition to adequately represent the incomplete informational input provided by peripheral vision. The hole-movement task which has been modelled by the implemented system excludes all peripheral information by masking.

The course taken to realize the generalization of representation is to propose generalized image chunk and element classes. Figure V.16 presents the chunk and image body element types of VIPS and their associated generalized classes. The general image chunk classes are not described further below, as their names indicate the concepts represented. The image body element generalizations determine the form of the image structure and are discussed.

ELEMENT GENERALIZATIONS

The image body element classes are defined by their members' function within an image body structure and are realized through a generalization of those element types used in VIPS. Inversely, each VIP element type is a specific instantiation of one of the generalized functional classes to be described here. This set of generalized element classes and the interrelating interconnections which are allowed are sufficient to produce all visual form image representations.

The XIT type element of the implemented system is an example of the functional class of elements which interconnect image body elements of chunks representing spatially disjoint information. An element A of this class normally references another element B of this class through an interconnecting link of space-traversing meaning. Element B references element A through a link of accordingly inverse space-traversing meaning. Element A and element B are in separate image chunks. Each element references a Position type element which specifies its location within the currently bound locational frame.

Each element of this class will have properties which adequately specify this inter-location (Space) traversal. Using the XIT element as an example, each XIT element specifies the direction (DVX) and range (RVX) of the space within the visual field which it representationally traverses. These two properties are sufficient to define a straight line segment, given the perceived location of the initial XIT element as a starting location for the line segment. The referenced XIT element (by LVX) embodies complementary traversal information to allow the link to be meaningfully traversed (accessed) in either direction.

This traversal of space or change of location may be across figure of ground. The XIT elements as proposed are consistently utilized to traverse the figure, embodying an existing line segment. Elements of the general class may span "empty" or ground space to interrelate differing image chunks of differing perceived spatial locations.

An element of this functional class also provides the necessary non space-traversing links to other image elements within the same chunk. An XIT element provides the links named by the properties DIV1 and DIV2 to interrelate it with other (ANGLE) elements of the vertex or line chunk in which it belongs.

The functional element class representing a generalization of the implemented INTERNAL type image element is similar to that just described. The elements of this functional class are primarily used to again span locations within the visual field (image), except that this link is now between elements of the same chunk. The INTERNAL image elements within OBJECT image chunks represent and allow traversal of object line-segment sides. Again, the links are between two elements of the same class, with complementary spatial information allowing meaningful traversal in either direction. These elements allow the image representation of visual concepts which span locations by tying together spatially disjoint groups of visual and spatial features.

These intra-chunk space-traversing elements provide local (Same location) links to other image elements of the chunk which are embodying other feature information at that perceived location. These local links necessarily maintain attention within the chunked concept when traversed ("figure - ground" distinction). These elements allow the image representation of visual concepts which span locations by tying together spatially disjoint groups of visual and spatial features. The INTERNAL type elements of

the implemented system provide one such link, named by DIV1, within any line or object chunk.

The functional class which the ANGLE element of VIPS exemplifies is that class which embodies the set of interesting visual features found at some location within the visual field (image). What features are of interest, and thus embodied within such a feature element, is determined by the active visual perception or image process.

These feature elements also provide the necessary means for their traversal by directed attention through the image representation. The ANGLE type elements of the proposed implemented system provide the links named by DIV1 and DIV2, which allow their consideration and through-traversal to an interrelated INTERNAL or XIT type element. Traversal of a feature element within an image representation does never involve any change of perceived location. Traversal of the interrelating links which they provide allows access to elements or elements of the image existing at the same location within the same chunk. Feature elements of differing chunks may be linked by equivalency.

One last generalized functional class is discussed now. This is a generalization of the END type image element of the implemented system. Elements of this functional class are used in image chunks which represent space-traversing image concepts which cannot be satisfactorily specified solely by properties of space-traversing elements. Each such chunk has a pair of these End-class elements, which are externally (equivalently) interrelated to either a pair of intra-chunk or a pair of intra-chunk space-traversing elements. The END elements of the implemented representation are linked to XIT elements when in LINE chunks, and linked to INTERNAL elements when utilized in SIDE chunks. Their function is to link the complex space

traversing concept which its chunk embodies (SIDE or LINE) to the appropriate elements of other chunks at its ends.

The Position type element as it has been defined, discussed, and implemented is already sufficiently generalized. The application of the fuzzy planar location reference frame in conjunction with a similarly bound fuzzy depth reference provides the ability to assign perceived or imagined locations as observed. The Chunk-Header element is provided as a means of gaining reference to an image chunk and to embody general chunk information, such as the chunk type. As such, with an appropriate extension of usable symbols, this element is sufficiently generalized also.

Thus, all of the image elements (except the QUICKSEE, to soon be discussed) have been generalized to realize distinct functional classes. It is proposed that these general element classes may be straight-forwardly instantiated to realize a visual image representation of any probable visual form environment. The image representation of VIPS is an instantiation for the environment of straight-line line drawings.

PERCEPTION SYSTEM GENERALIZATION

The memories, the processes, and the interrelationships existing between them which are discussed in Chapter I.4 form the basic system structure of the human visual processor (Figures IV.1, IV.2, IV.3). TCLTM also plays an important role, as discussed earlier in this chapter (Figure V.1).

The perceptual theory stated generally is that successive partial representations of the external visual environment are meaningfully integrated by the assimilation-accommodation process, with appropriate aid from the recognition and incorporation processes, to realize a complete, visual image representation. The final image representation is capable of generating verbal and drawn descriptive outputs and of generally guiding behavior within that given environmental context. It resides in ITM as a structure of interrelated image chunks. It is necessary to recall into or maintain in STM the image chunks of interest to actively guide or produce such related activity.

THE PERCEPTUAL GOAL

Strategic control is maintained by an ever-present internal state, the currently active goal. This currently active goal is fundamental in determining what information of that which is currently accessible and available is relevant and worthy of attention. This goal also influences the class of image or other representational structure into which the currently interesting information is to be integrated. The active goal, in conjunction with the currently attended image content, is basic to the determination of eyemovement and fixation point choices. Thus, the goal influences the structuring and attending of the perceptual image.

The question arises: What is the source of these goal types? The initial goal type to become active is primarily a derivative of the current overall cognitive goal type active prior to the initiation of the visual perception. Any sensory stimulation is interpreted in light of the pre-existent and reactivated state of the system. The subjects of the hole-movement task had as their pre-set cognitive goal one reflecting the desire to achieve a perception of the presented line drawing sufficient to enable subsequent verbal and drawing reproduction. From this goal were derived either the goal SEO or the goal RNO, being the first active goals of the two implemented perceptual strategies.

Given an initial goal derived from a pre-perception condition, the sequence of currently active goals is generated through an interaction of the active goal and the available visual (icon) and image (STM) information. This is the essence of any assimilation-accommodation scheme; the strict interdependence between currently active goal and the currently available information of the image representation. At any point, the interaction between active goal, active image contents, and the newly fixated visual information may result in the initiation of a new active goal. A network of uni-levelled goal transitions was inferred from the observed behavior of the hole-moving task. With the availability of peripheral information, some stacking of goals, due to the multiplicity of hypotheses possible with more input information, may occur. It is proposed that the stacking of hypotheses and goals within the visual process is still very shallow, probably of a depth of at most two or three. This is a direct result of the strong interdependence between the icon and accessed image and the active goal during assimilation - accommodation. This interdependence reduces drastically the value of any long-range plan of attack embodied within an elaborate stacking of future goals.

Support for the role of an active goal (goal type) in the human processing of visual information comes from several sources external to this thesis. Eye-movement recordings of reading activity indicate the goal-oriented direction of attention across lines of presented text (Yarbus, 1967). Yarbus reports that varying the instructions concerning the eventual questions to be asked about a presented complex picture greatly alters the subject's eye fixation activity. In more rudimentary visual search tasks involving target location (Williams, 1967), search patterns were shown to display random characteristics only in the absence of information about the search target.

These studies indicate the affect of a pre-set cognitive goal upon the subject's direction of attention within a presented visual environment. Their results indicate the close interdependence between the active perceptual goal and the presented visual environment throughout the subsequent visual processing activity. Areas of the visual environment are fixated in accordance with their believed relevance to the perceptual task.

A pre-existing goal likewise has been shown to have an effect upon the representation of the visual perception. Chase and Clark (1972) conducted an experiment involving the comparison of pictures and sentences for meaning equivalency. When the sentence was provided first, the subsequent encoding of the picture appears to be contingent upon (influenced by) the form of the sentence. The sentence provides a pre-existent goal (internal state) which guides the development of a picture representation suited to the required comparison.

PERIPHERAL VISION

Peripheral vision input, the primary informational difference between the

implemented and generalized processing system, appears to play two significant roles. First, in conjunction with any currently proposed perceptual hypothesis or active goal, it is instrumental in determining the necessary pre-set range and direction for performing saccadic eye-movements. Hebb (1949) noted this role of peripheral vision input. The use of extra-foveal information in the selection of a new fixation position has been confirmed from data obtained in several visual search tasks (Williams, 1966;1967). Either the direction and range of a saccadic eye-movement is pre-set by AA-SYS or a peripheral element of the current icon (see below) is marked as the area of the visual field to next fixate. VI-SYS, with access to the icon (VI), could then itself generate the direction and range in terms of a fixed retinal location relation. VI-SYS would then deactivate only after properly fixating the prior peripheral area, this possibly involving one or more corrective saccades.

Secondly, peripheral vision appears to play a significant role in both generating and confirming proposed hypotheses concerning the external visual environment, supplementing the more accurate information obtained in the fixation point area. In a visual search task involving pattern discrimination (Gould and Dill, 1969), peripheral vision appeared to supply sufficient information to allow superficial judgements of pattern "similarity" and to effect accordingly the probability of foveal fixation of a pattern in a matching task. External research indicates support for both processing roles of peripheral vision that are proposed above. Directly acquired data of Experiment II will be seen to provide data favorable to this interpretation of peripheral vision's roles in perception.

PERIPHERAL INFORMATION REPRESENTATION

The proposal for the representation of the peripheral vision input to the visual perception system is to generalize upon the QUICKSEE type image body element. This type of element has been introduced into the proposed representation to realize the inferred incomplete embodiment of vertex information developed in the case of a subject moving continuously through an encountered vertex. It is this characteristic of incomplete representation which is proposed to typify the available information of peripheral visual input. As such, an appropriate generalization of the QUICKSEE element is proposed as the extension sufficient to embody peripheral information.

This element, in its generalized form, will be called the Peripheral element. Elements of this class are proposed to embody a density measure of possible interesting visual environment characteristics at locations of high discontinuity in the periphery. Which of the available peripheral elements are considered at any point is again a function of the active perceptual goal. An element of this class references a position type element embodying its fuzzy location in relation to the fixation location. The imaginal space of perceived locations is not known at this (fictive) stage of processing. These elements fail to provide any interrelating links to other image elements of the visual input. These two stated characteristics are the results of a consideration of experiment II data to be discussed below.

The adequate representation of peripherally available visual information by one such symbolic element type is admittedly doubtful. This is due to the indistinct dichotomy existing between foveal and peripheral visions. There are not two separate and distinct levels of visual acuity existing, as such vision dichotomy appears to imply.

Instead there is a continuous, though rapidly decreasing level of visual acuity with input visual angle eccentricity (Mandelbaum and Sloan, 1947). There is not agreement as to the actual area of the visual field considered to be foveal vision, though generally the estimate is somewhere between a 1 to 3 degree central angle. As such, there This discussion assumes that visual acuity is related to the amount of features available for representation and thus to the type of symbolic element which may be utilized for such encoding. As such, there may not be one specific element type which is used to represent peripheral vision information.

In the discussion to follow concerning the proposed visual processor, the generalized QUICKSEE (peripheral) element is used. The controversy concerning such a proposal is simply being duly noted here. The basic hypothesis is that peripheral vision content can be sufficiently represented, by discrete symbolic means. The discussion of data to follow indicates support for this hypothesis.

THE SINGLE (INITIAL) FIXATION

Part I of Experiment II was designed to give some results relevant to inferring what information becomes available from an initial fixation upon a presented line drawing. As such, and as noted previously (Chapter 12), each line drawing was presented for one-quarter to one-third of a second duration with the fixation point being pre-set. Immediately upon removal of the stimulus the subject was to simultaneously draw and verbalize his partial perception of understanding of the drawing. Thus, the results not so much reflect the immediately available "iconic" memory as they do the results of the processing of this input information by the perceptual system. There is no attempt to interrupt the processing to get more direct

indications of the iconic representation (the contents of VI in the proposed system) as done by Sperling (1960). The primary interest here is to determine what are the possible results of the processing of the initial glance. This can yield indications as to its influence over the subsequent visual processing which would follow.

The results indicate that the initial glance is indeed a valuable source of perceptual hypotheses, in the form of constructed images with high correspondence to the features of the input information. These images serve as the basis for the operation of the assimilation-accommodation procedure, which accordingly directs attention and develops a final consistent representation of the presented stimuli.

The discussion to follow references Appendix A. Fixation. The appendix contains several groupings of data and associated interpretations. Each grouping begins with the presented stimuli, with a star (*) serving to denote the pre-set fixation location. The following page displays a subject's verbal and drawn response to the presented stimuli. Following this is a page of intermediate interpretation of this response in the form of pictorial representations of the images constructed and their relationship to the proposed icon image.

To begin, consider group I of the appendix. As can be seen the fixation point was in the lower left. Figure (a) of the interpretations page (page 3) is the inferred contents of the visual information cell (VI) at this fixation point. There is a full representation of the immediate, fixated corner and the circles indicate peripheral information elements which embody visual field discontinuities as discussed previously in the chapter. As noted above, no attempt was made to get more direct indications of this input representation. Through the discussions to follow, I hope to show that this representation (involving the peripheral elements in conjunction with a full fixated

vertex representation) bears some functional relationship to the obtained results. This is still an uncertain area, needing further experimentation.

As can be seen from the verbal and drawn response (page 2), the subject reacts to the "trigger" vertex that has been fixated and accordingly proposes a cubic image for verification. This is seen to fit fairly well with the inferred input information, as shown in Figure (b) of page 3 of the appendix. The subject apparently further notes the existence of the peripheral element not accounted for by the regular cubic projection shown in the middle of the middle edge of Figure (b). This leads to uncertainty in the initial proposed image, and so to the transparent image development as noted in Figure (c). This response indicates the influence of peripheral vision in the development of the perceptual image proposals. With no peripheral vision the proposal would have been only that of Figure (b), as occurs repeatedly in the protocols of the hole-movement task. With the peripheral vision a problem with that is noted, possibly resulting in two competing images for possible confirmation. This could serve to direct attention immediately to the trouble spot to determine if that can easily be assimilated into the initial solid cubic projection.

Group II is now considered. As can be seen the fixation is in the lower left, and the inferred input is shown in Figure (a) of the interpretations on page 6 of the appendix. The subject notes that the external outline is that of a square, again noting the importance of that strategy of visual perception, as previously embodied by RAAS for the hole-moving task. Figure (b) indicates how that image proposal could be related to (derived from) the inferred input representation.

The "trigger" vertex that has been fixated leads to an attempt at three dimensional interpretation again. Figure (c) indicates a possible attempt at fitting this perception

proposal to the input representation. The subject does pause appreciably after noting the three-dimensional possibility. There appears to have been too many inconsistencies existing between this image proposal and the input information. Thus, the subject returns to the image of Figure (b) to guide the assimilation-accommodation process. The need to determine the intersection of the inward directed lines appears to be the goal which next would have been utilized, as in RAAS for the hole-moving. Again, peripheral vision appears to have a valuable position in the processing. It has provided sufficient information to lead to an apparently rather certain "square" proposal for the external edge without further scanning required. The image hypothesis without peripheral vision input would have been the cubic projection image.

Group III further gives indications of both peripheral vision's power and its limitations. Two subject responses are included in this group. As seen the fixation is in the lower corner of a cubic projection. Figure (a) of the interpretations (page 10) indicates the inferred input information.

Both subjects report seeing two solid block type objects. Figures (b) and (c) on page 10 of the appendix indicate possible interrelationships between the two proposed object perceptions and the input information. There one object has been completely bound only to peripheral elements. It is not clear that solely derived from those elements, as the "trigger" vertex of the fixation area has led to an attempt at a three-dimensional interpretation in the foreground. This course of interpretation surely has some effect upon the development of the proposed cubic image interrelating the other peripheral elements of the visual information. This indeed is an indication of the ability of peripheral vision to aid in the proposal of perceptual images, as one object proposal interrelates solely peripheral information elements.

Yet, both subjects were unclear concerning the spatial interrelationship existing between the two object proposals (See drawn responses). It appears that the peripheral elements do not adequately embody all of the information necessary to propose or realize the equivalency and other spatial interaction relations which may occur between objects. Further eye fixations are required to fully develop these object interaction representations.

In conclusion, the initial fixation is a source of perceptual hypotheses (image chunks). These image chunks provide a basis for subsequent assimilation - accommodation. Peripheral information plays a significant role in forming these proposals. The representation of peripheral input that is proposed appears to be functionally adequate as a basis for the explanation of the behavior observed.

EYEMOVEMENT SEQUENCES

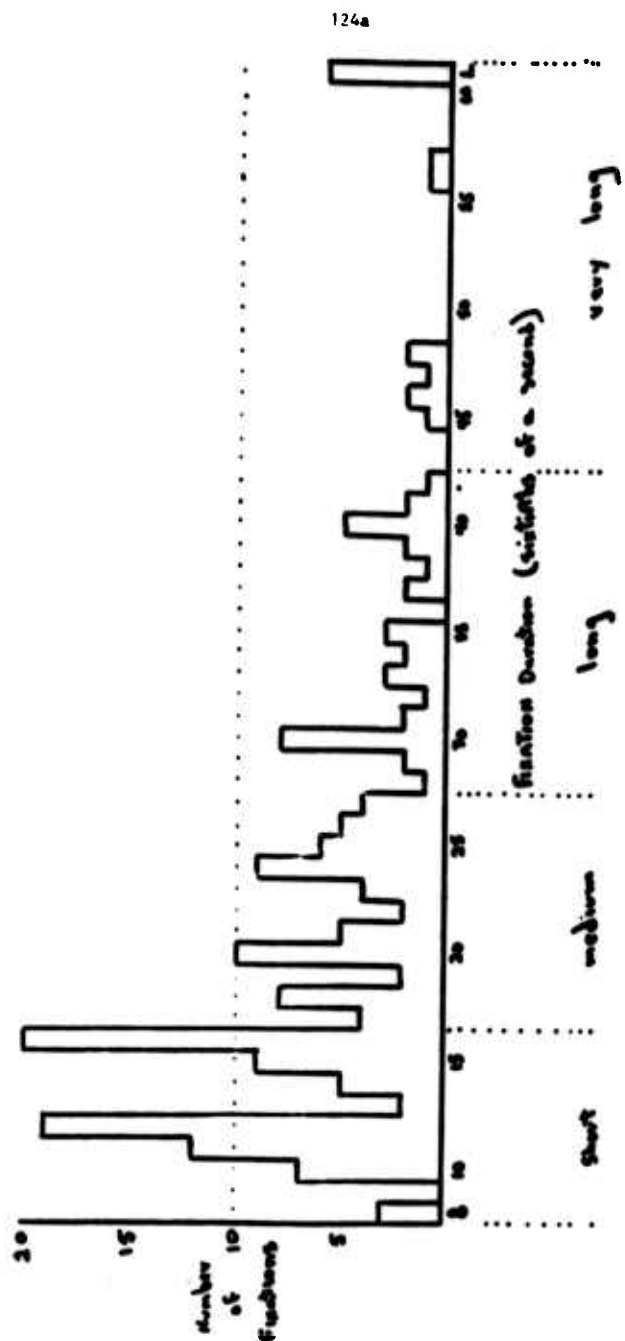
Attention now turns to a discussion of several of the eye movement protocols obtained through Part II of Experiment II. Appendix A:EYEMOVE contains eight sets of data. Each set consists of a page illustrating the eyemovement sequence observed for a subject upon the line drawing shown. An initial section of the eyemovement sequence has been appropriately interconnected by dotted lines. Two numbers are associated with each dotted fixation point. The fixation number in the sequence is given, and then in parenthesis the duration of the fixation in sixtieths of a second is indicated. Following each eyemovement sequence is a page presenting the resulting verbal descriptive response obtained and the subject's drawing from memory. As mentioned, there are eight sets, two subjects on four line drawing stimuli. These line drawings were used previously in the hole-moving task. No subject recall of the complete pictures occurred. This indicates that TCLTM images did not play a significant role.

One general data dimension to consider is the duration of eye fixation. The graph of Figure V.17 indicates that fixation duration varies. The activity of assimilation - accommodation requires varying amounts of processing at each fixation point, as exemplified by VIPS. Assuming that fixation duration is directly related to the amount of processing done at a fixation, this general level of data consideration does not discourage an assimilation-accommodation explanation.

More specifically, Figure V.18 presents the fixation durations for the first four fixations of the eight sequences. In all but one case, the first fixation is consistently longer than the second (actually the longest of the first four) and the difference between their averages is significant. The initial fixation has been described as a source of perceptual hypotheses. The duration comparison indicates that such extensive processing may be occurring at that point. The exceptional case involves a line drawing which includes two overlain objects (as perceived). As the equivalent examples of VIPS activity indicate, this may cause re-recognition (image accommodation) which results in appropriately longer fixation times during the object scan for confirmation.

In general, a relatively lengthy fixation is followed by several shorter fixations. This is favorable to the behavior interpretation proposed here. Recognition activity is followed by several fixations for hypothesis confirmation (assimilation) or rejection requiring image accommodation. The variability of hypothesis success and necessary environment assimilations during confirmation yield the irregularities in the cycles of fixation durations.

To facilitate discussion below, the fixation durations have been classified. Durations of 8 to 16 sixtieths of a second are short (See the bottom of Figure V.17).



Figure

Figure V.18

Fixation Durations (sixtieths of a second)

Fixation Number

	1	2	3	4
S1	24	10	15	24
P1	40	10	12	34
S2	35	13	11	17
P2	34	15	11	15
S3	33	10	32	11
P3	24	26	21	15
S4	42	10	41	26
P4	43	11	20	12

Stimulus
Sequence

* *

* *

MEAN	34.4	13.1
Average Duration	5.6	3.8

Fixations from 17 to 27 sixtieths of a second are intermediate. Fixations of 28 to 42 sixtieths of a second are long and those that are longer still are very long. This is not a classification based upon statistical clustering.

Now considering the first eye-movement illustration S1. In terms of fixation lengths the first is intermediate, followed by two short fixations. The first fixation has been noted above as being a provider of proposed perceptual (visual) images. The two following quick fixations can be then interpreted as functionally confirming feature groupings as proposed. The fourth fixation is again intermediate in length, followed again by two short fixations. Thus, as inferred with the hole-moving task, the subject appears to be engaged in a cyclical application of the recognition process in cooperation with subsequent eye-movements guided by the assimilation - accommodation process on the basis of those images proposed (constructed) in STM. Both fixations 7 and 8 are intermediate in length, followed by a long fixation 9. This could indicate an incorporation phase for those objects thus far satisfactorily confirmed. The subject then commences to rescan the picture, viewing further parts of the picture previously seen only through peripheral vision. This could be to fully develop the representation to be sufficient for drawing purposes.

Turning to sequence P1, it is noted that the starting fixation point differs. This was not intended, but did yield an interesting result. Again the subject engages in a long first fixation, followed by two short fixations. The higher location of the initial fixation has apparently allowed a better perceptual image proposal for the top figure of the line drawing than in sequence S1. This has drawn the attention of the perceptual process to that figure earlier than in S1. This difference is not only indicated in the eye-movement sequence, but also is carried through into the sequence of the verbal description.

This subject (P) utilizes long fixations more frequently than subject S. This could be indicative of a more piece-meal approach to incorporation. The top object may be incorporated at fixation 4 and the slim rectangle to the right may be recognized and incorporated at fixation 5. This would imply an initially heavy reliance upon peripheral information for object recognition. From a general analysis of the sequences recorded and those illustrated in the appendix, it appears that such initial reliance is utilized. This is followed by a rescan often visiting feature locations previously not fixated. Often later attention is redirected to areas of object interaction several times, apparently to confirm the proposed image interactions. The lower right corner of the right-hand square is never fixated in P1, while other interaction areas near the center of the picture are re-fixated frequently. The subject has difficulty when drawing this picture area. This would correspond to the inability of peripheral information to embody object image interactions accurately as noted from the single fixation data discussed above.

S2 and P2 involve the perception of an impossible drawing when considered three-dimensionally. Hole-movement protocols involving this stimuli were previously discussed with regards to the three-dimensional extension of the programmed system. One thing of interest in the S2 sequence is the subject's reliance upon peripheral vision for the perception of the bottom cube. The front face of that cube is never fixated. Attention is primarily directed upon the ambiguous face and edges of that cube in an apparent attempt at disambiguation. Disambiguation is not realized. As with the hole-movement protocol AP3, the ambiguity does not interfere with the verbalized description, but does cause the drawing procedure to have problems.

Eye movement sequence P2 has been displayed on two pages, P2(a) and P2(b)

P2(a) shows the first twelve eye-movements connected by dotted lines, the fixation 12 being a very long fixation. This initial eyemovement sequence may be interpreted as related to the behavior which produces an image which is capable of giving rise to the initial segment of the verbalized description, bracketed by "a". Again the front face of the bottom cube has not received any significant fixation attention. This initial sequence is followed by ten more fixations, these connected by dotted lines in P2(b). In this sub-sequence the bottom cube is completely scanned, and then attention is focussed for two long fixations along the middle line of the upper part of the drawing. These eyemovements appear to bear a correspondence to the second part of the verbalized description bracketed by (b). It appears that the subject develops two distinct picture representations (perceptual images) here, realizing the inadequacy of the ambiguous description for drawing purposes. Again evidence of the pre-set cognitive goal (the drawing task to be accomplished) upon the perceptual process is demonstrated.

Both S3 and P3 show the alternation of fixation durations and the frequent re-fixating of object interaction areas. An initial scan of the complete picture, which appears to rely heavily upon peripheral indications is followed by a rescan to insure the proposed image's correspondence to the presented drawing.

Protocols S4 and P4 are important for their similarity in the use of peripheral vision to realize three-dimensional perceptions. In both instances only areas of the front faces are fixated. The three-dimensional "trigger" vertices are fixated and the interaction between the cube and sloped plane receives the expected consideration. Peripheral vision here is deemed sufficient by the assimilation-accommodation process as a source of facial confirmation information.

Only general consideration has been given above to the eye-movement protocols given in the appendix. No move-by-move description or explanatory generation is attempted. That some such descriptions appear possible in terms of the activities of the processes and the contents of the memories proposed by the thesis is the degree of success desired here. The interested reader is urged to consider the appendices further, as the data presented has many valuable points.

Thus, through the discussion of the two appendices of Experiment II data, the proposed theory of visual form perception gathers further anecdotal or inferential support. As indicated, much research is needed to gather further data applicable to the refinement of such a theory of visual image representation and perceptual processing.

Chapter 1.6 Conclusion

This chapter considers several other recent studies involving the memory of primarily visual concepts and discusses these proposals in light of this thesis. An in-depth discussion of general characteristics of the human cognitive representational system follows. Finally, several suggested directions for further research will be noted as a conclusion to the paper presented here.

1. Study Comparison

Any choice of studies for comparison must be highly exclusive. As such, three recent studies have been selected to be contrasted with VIPS. All three involve the use of a symbolic representation of visual information in a computer-implemented simulation of cognitive behavior.

BAYLOR'S BLOCKS

George Baylor (1971) proposes that an image representation of a cube is an integral part of the processing machinery involved in the solution of a block visualization task. The subject is given a painted block description and instructions as to how this block is then cut into pieces. The subject is required to answer various questions as to the number of resultant pieces with given painted conditions (i.e. how

many have one blue face?). As a result of the analysis of a verbal protocol, Baylor proposes the existence of two interacting representational problem spaces which become involved in the problem solving activity.

"S-Space" (symbolic) holds true generalizations (facts) concerning blocks and also is used to hold an encoded (parsed) representation of the question to be answered concerning the segmented block. This encoded question representation is applied to the block representation of "I-Space" (image space) by available matching operators which locate image components for counting. Finally, an answer is deposited in S-Space. Thus, S-Space acts as a translational buffer memory between the verbalizations and image representation (STM) and as a semantic component of long term memory. Baylor's proposal is admittedly not a complete cognitive model.

The block representation residing in I-space is hierarchical in nature. This structure reflects the hierarchy of part containment (i.e. block, to faces, to edges, to vertices). As the initial block is segmented according to the prescribed cutting, other specialized intermediate hierarchical groupings become applicable and are used (i.e. corner, border, center cubes). At the bottom level, vertices are interrelated by spatially meaningful links. Although these spatially meaningful links are incorporated into Baylor's image (I-Space) representation, this information appears to be irrelevant to the task's solution and thus also to the model's operation. The input question is parsed into a hierarchical semantic structure which then serves as a goal stack. This hierarchy of goals directs the altering and accessing of the hierarchical cubic representation found in I-Space. Through a matching of structure schemes, the relevant information is obtained and placed in S-Space as the answer. Thus, though spatial information is a part of Baylor's image representation, it is not put to significant

use. Furthermore, this image is not accessed by means related to that of accessing external visual information. Thus, Baylor's image is not a visual image as defined in Chapter 1.1. Baylor's image representation's structure closely resembles the usual (his) encoding of verbal information. I-Space as defined does not merit the claimed separation.

NEWELL'S STIMULUS ENCODING

Newell (1972) develops an encoding representation for colored bottles. He then defines several associated processes which prove sufficient to accomplish a visual series extrapolation task. The salient features for determining the generating characteristics of the series, and thus for proposing the next series element, are the bottles' orientations and colors. Not surprisingly, the representation proposed for the visual input (and memory) of each bottle is merely a list of just these features. For example, "(OBJ BTL VT RED)" represents a vertically oriented, red bottle.

Newell does not propose a visual image representation of the visual environment, even though the information source is spatially structured (the series is presented by a left-to-right row of bottles). The spatial structure is representative of the bottle sequence in this task. A visual image representation which must contain spatial and form characteristics, would be superfluous to the subject's accomplishment of the given task, and thus is so for the inferred system also. The task can be interpreted as being that of determining the bottle sequence's generating syntax. A representation which is processed syntactically has been appropriately inferred. This representation does not satisfy the visual image definition either.

Two special situations which are encountered by Newell's system during solution

of the task indicate the applicability of an assimilation-accommodation process. Newell proposes a constant hierarchy of feature relevancy from an analysis of subject verbalizations on twenty-three series extrapolations. This hierarchy determines the order of features as acquired on input and as subsequently considered in determining the generating syntax of the series. This constant hierarchy leads to several erroneous series proposals by the implemented system.

The implemented system also encounters trouble in achieving the overall series organization. Newell says, "There must be a delay in actually organizing the perceived sequence, since subsequent objects have not yet been observed and they may affect the organization" (p 21). When encountering two type bottles for input (X and Y) and the first four are as XYXY, the system prematurely organizes them as $((X Y) (X Y))$. If another XY pair is encountered the final series structure will be $(((X Y) (X Y)) (X Y))$, which is inappropriate. Newell discusses and implements further processes to perform the necessary reorganization as $((X Y) (X Y) (X Y))$.

These problems relating to feature relevance and overall sequence organization are symptoms of the proposed system's failure to use existing STM contents as the source of organization and expectations. The most relevant stimulus feature (or features) is not chosen on the basis of a static hierarchy of relevance. It is (should be) chosen on the basis of the current representation of the series in STM, the current context. The partial series image, as represented in an intermediate STM state, is also used to propose expected overall series organizational characteristics. Any current regularity of the incomplete input serves to suggest a structural representation of the complete input. This proposal provides expectations and determines feature relevance with regards to the newly fixated input. An assimilation-accommodation process can

then operate to note consistency of new inputs with expectations and accordingly effect alterations, generate new proposals, and determine new feature relevancies as necessary.

For example, given XYYX the process constructs an overall series representation of (X Y) (X Y) (X Y) prior to the input of the last two elements. No "delay" is necessary, as Newell proposes. Instead, stimulus encoding (perception) entails early (hypothetical) structuring which facilitates subsequent assimilation - accommodation.

MORAN'S PATH IMAGE

In another recent proposal, Moran (1973) develops a representation to account for observed subject verbalizations during a path imaging task. The representation developed there is again hierarchical in nature (as Baylor). The sequence of line segment inputs exists at the bottom level of the structure, with an organizing structure of visual concepts developed over this. The structure is used as the basis (means) for the "chunking" of information. A single symbol of the hierarchy in STM is capable of representing and activating (recalling into STM, if necessary) all of the elements which are below it in the structure.

Several task requirements are again particularly relevant to the design of that proposal. The sequence of line segments to be imaged retains importance throughout the task. Indeed, the task is stated in terms of "imaging a path", not a picture (line-drawing). The line-segment elements of the path are determined randomly. Thus, the subject has no expectation that a coherent, easily recognizable path would be presented. Actually, the opposite characteristic (of path randomness) is probably naturally assumed. As such, this appears to have inhibited dependence upon object

expectation proposals and forces the subject to develop a conservative image representation at all times. Partial objects are actually more to be expected than completed ones. This yields a hierarchic representation in terms of realized parts similar to Baylor. Randomness does not prevent expectations (Feldman, 1961), but in this case it introduces a readiness for partial concept representation as a useful available contingency.

Some difficulty exists as to the representation's ability to realize inferences based primarily upon spatial (locational) interrelationships. This results from not including locational symbols in the image representation. The scarcity of protocol evidence of such inferences allows the inefficient (not straight - forward) implementation of these inference realizations. In the protocols of this thesis, inferences based upon locational relationships occur frequently. As such, the visual image is bound to a framework of locational (positional) symbols which create an imaginal space.

REPRESENTATION COMPARISON

Behavior on three tasks involving the representation of what is primarily considered to be visual (or visual image) information yields three fundamentally differing proposals. In two tasks (Newell, Moran), the sequence of inputs retains importance and is carefully imbedded into the memory structures. In two tasks (Moran, Baylor), partial concept representations (due to cutting of a cube or expected, frequent incompleteness) are of importance and a hierarchical structure reflecting component part subordination is proposed. Two proposals attempt to fully represent the cognitive processor (Newell, Moran), and their visual representations reflect a concern for STM research. None of the three visual representations satisfy the visual image definition of Chapter 1.1.

In each study, the task determines what features and relations of the visual information are necessary for the successful accomplishment of its requirements. The visual modality factor appears to have had much less effect upon the form of these three proposals than do other task-related considerations. The task of this thesis, being the realization of a perception of a given line drawing has likewise resulted in another differing proposal for the symbolic (cognitive) representation of visual information. In this case the modality factor is the primary task feature. This has resulted in a representation satisfying the visual image definition.

2. Flexibility of Representation

Behaviorists have long attempted to realize a parsimonious theory of human internal information representation. They propose that all internal representation is in the form of mental words. Associationists have also been in this fight. Recent proposals state that all memory is realized in the form of concept - relation based hierarchical structures (Kintsch, 1972; Bower and Anderson, 1973).

In contrast, it is proposed here that the representational aspect of human cognition is a very flexible aspect. The apparent need for concise representation within our limited operational or active memory precludes the development of a parsimonious theory of human information representation consisting of a single or several satisfying universals of form and content. The ability to fit the active information representation to the goals at hand is proposed as being the primary characteristic of the human representational system. As George Miller has said :

"it seems to me that the very fact of our limited capacity for processing information has made it necessary for us to discover clever ways to abstract the essential features of our universe" (G. A. Miller, 1967, p49).

LANGUAGES OF ART

The discussion to follow relates this proposal to a most interesting treatise on the representations and sources of meaning. It is a book by Nelson Goodman (1968) called *Languages of Art*. In the book he exactly defines a set of terms and proceeds to extensively compare and interrelate these into a coherent system of meaning. Several such terms and their interrelationships are relevant to this attempt at describing the selectivity of representation which is proposed here. Though Goodman writes primarily to explain the representative power of art, the concepts which he develops are useful to the understanding of human cognitive (internal) representation.

An object of the environment "possesses" all of those properties for which the corresponding predicate applies and is true. Thus, in Figure VI.1, object O has properties A,B,C,D. These true predicates (ie. "HAS A") apply to the object from the property as a meaningful reference. An object necessarily exemplifies all true properties which apply as opposite directed references. An object symbol (or representation) possesses all properties of the object which it represents. (See Figure VI.2.) An object symbol may furthermore exemplify any property which it possesses. As Goodman puts it, "Exemplification is possession plus reference" (p 53). In Figure VI.3, object symbol O exemplifies only properties A and B of those which it possesses.

Goodman says that "just which properties of a symbol are exemplified depends upon what particular system of symbolization is in effect" (p53); and further,

Figure VI.1

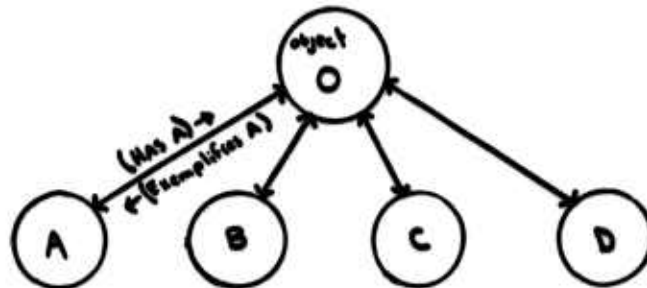
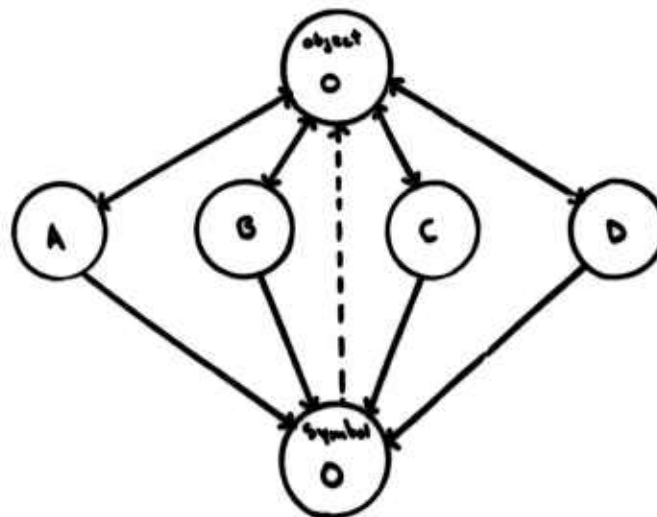


Figure VI.2



136b

Figure VI.3

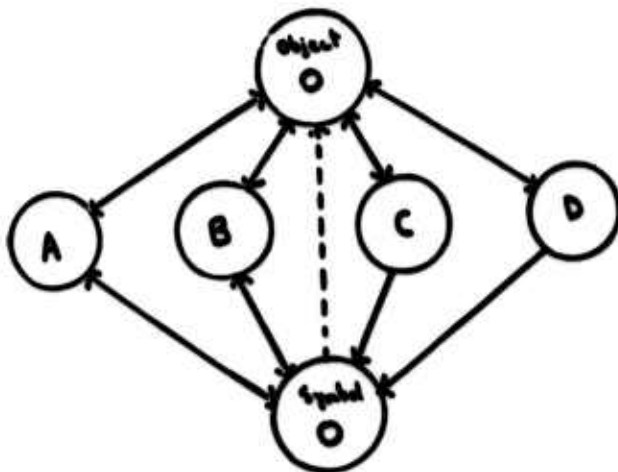
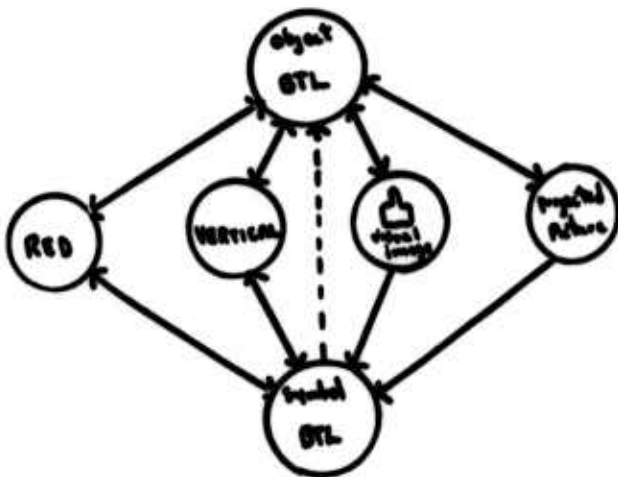


Figure VI.4



"Establishment of the referential relationship is a matter of singling out certain properties for attention, of selecting associations with certain other objects" (p88). Putting these two statements together, certain properties that are possessed by an object will be of interest, which properties are exemplified by the chosen object symbol (representation). An associated, particular system of symbolization will be in effect as a result.

In the three comparative studies, and in this thesis itself, the task leads subjects to determine which properties of the involved visual information are necessary for its successful completion. The task determines which properties receive attention and are exemplified by the memory representations. A system of symbolization results. For example, only the properties of color and orientation are exemplified by the bottle object representations of the symbolization system proposed by Newell (see Figure VI.4).

Goodman defines the useful concept of "representation-as". This term differs from just plain "representation". While "representation" involves solely an operation of denotation, "representation-as" involves both denotation and classification. Classification is selectivity of exemplification. Thus, classification naturally alters the form of representation as shown by the study examples. The human ability to "represent-as", which is well demonstrated by artists' accomplishments as noted by Goodman, is exactly the flexibility of representation which is being proposed.

In the visual image representation proposed by this thesis, a block is "represented-as" a set (structure) of interrelated, non-hierarchic, facial-projection representations (image chunks). This form of symbolic representation facilitates visual recognition and processing as has been discussed. A complex hierarchy of object

(block) parts (as Baylor) contains unnecessary information and is wrongly structured for efficient use by the visual processor.

As Goodman predicts, the systems of symbolization which have been inferred to account for the observed subject behavior on the tasks of each study differ significantly from each other. Yet each inferred representation adequately embodies the information which is sufficient for the accomplishment of the given tasks. Instead of pursuing an argument of conflict here as to which proposal is closer to the content and form of a "universal" representation (image) of visual information, I claim that all of these studies serve, in concert, as initial results supporting a general theory of representational flexibility.

A THEORY OF COGNITION

While Goodman's ideas have merit, he does not say enough. He does not attempt to describe the cognitive system which makes possible the use of the representational abilities which he proposes. He does not do this as he has other goals. The goal of this thesis, on the other hand, is exactly to describe such an operative cognitive system. Therefore, the following theory is presented. The theory embodied by VIPS is a specific instantiation of this general theory of cognition. Indeed, the specification of a theory of any class of cognition must be grounded in an overall theoretical basis of all cognition.

The basis of all cognition is representational knowledge. Representational knowledge is embodied by long term memory. Long term memory is a complex semantic structure of symbols and relations. A symbol is a node in the network-like long term memory structure. A relation is a directed connection between two two

symbols. Being a semantic structure, long term memory is a representation of meaning. The meaning of a symbol or relation is defined by (to be) its location within the long term memory structure. Relations are always paired. A pair of relations is said to be a property of the symbol from which both relations originate. One relation of the property pair references a symbol which labels the link between symbols which the other relation constitutes. A symbol is simply a group of relational properties. A symbol may be connected by relations to communicable components of a language-based substructure of long term memory which allow its external expression. A symbol may also be connected to components of the various modality-specific perceptual systems which allow it to be used in the perceptual activity and to be an element of the resultant perception.

The symbols and relations of cognition are physically embodied in the structure and the patterns of activation of the neurons of the brain. Brain neuron activity can be (is) effected by the activity of the complete central nervous system. Human behavior not involving the use of symbolic representation is not cognition. For example, simple reflex actions (e. knee jerk reflex) appear to be based solely upon lower-order nerve cell states alone (sensory, motor, and spinal).

The symbols, relations, and structure of long term memory develop as the product of the organism's (person's) continuous attempt to adapt to (represent) its environment. The process of representational adaptation is an evolutionary process. As such, long term memory is a function of an innate initial state, of innate potentials, and of the history of its involvement in all cognition. Due to the evolutionary character of long term memory, the existence and meaning of any symbol is necessarily semi-permanent.

To adapt to an environment is to be capable of effecting successful (adequate) behavior within that environment. As the basis of cognition, long term memory must represent not only an environment but also the means of responding to and acting upon that environment. Therefore, cognitive processes are themselves structures of symbols and relations that are part of long term memory. To be operative, cognition has one requirement external to long term memory. Cognition requires a systematic cycle of long term memory activation and interpretation. It is this that life and time provide (or are).

As processes are embodied within long term memory, it is both the means and product of its own evolution. Long term memory evolves by the action of three process types. The processes of remembering incorporate selected environmental factors, created by other task-related processes, into the long term memory structure. Developmental processes determine the state of availability of innate potentials. Learning processes alter long term memory in three ways. One is the alteration of existing symbol and relation meanings by the addition of new relations which restructure of long term memory. Two is by the creation of new symbols and the incorporation of these into long term memory. The third is by developing activatable structures as rules and environmental factors suited to the expected tasks.

At any point in time (during any cycle of activation and interpretation) only a small fraction of long term memory is active or selectively attended. The active memory is a representation of current context. Active memory consists of a set of environmental factors and an active rule of a cognitive process (the active process). The environmental factor is the unit of contextual environment representation. The rule is the unit of cognitive process. As a part of long term memory, the environmental factor

and the process rule are both structures of symbols and relations. To define an active representational system (eg. visual imagery) is to specify the symbols, the relations, and the structure types which constitute its units of activation (environmental factors). To define a cognitive process is to do the same for its rules.

The set of environmental factors which are active at any point in time constitutes the environmental context for the active rule during that cycle of activation and interpretation. Each active environmental factor plays a process-related function and, as such, is said to be (be in) a certain component of active memory. The possible functions are defined by and exist with regards to the active process rule. General characteristics of the structure and content of an active environmental factor are determined by the functional role it plays. Thus, a currently active rule determines the general nature of its representational environment. Cognitive processes exist on the basis of a task to be accomplished (adequate behavior). As such, general characteristics of the current environmental representation are indirectly related to the current task.

The flexibility of the environmental representation to suit the current context to the task at hand is necessarily restricted and imperfect. At any time, flexibility is bounded at the extreme by the currently available symbols, relations, and structure of long term memory. The flexibility is also limited by the evolutionary state of sophistication of the active rules (processes) which become involved in the task solution. The state of process sophistication and specialization is itself inversely affected by the symbols and relations which are available for activation as rule or environmental factor contents. The state of organism development determines the availability of innate potentials which also affect the range of flexibility. Selectivity of activation is the operational equivalent of Goodman's selectivity of exemplification.

Upon the commencement of each cycle of activation and interpretation, a rule and a set of environmental factors are activated by selective attention. There are several general criteria which determine this selection. The rule is selected on the basis of the currently active process (a partition of the cognitive rules) and on the basis of the contents of environmental factors which have been active during the immediately prior cycle (primarily) or during one of the past several cycles (recently active). The environmental factors which are activated initially are determined in conjunction with the rule selection. They are a selection from the recently active environmental factors which determine the rule selection and which are therefore activated to play functional roles with regard to the selected rule. During the cycle of activation and interpretation, further symbols, relations, and even environmental factors may be activated. This selection is determined by the active rule and by the meanings of the already active environmental factors.

An active rule or environmental factor remains effectively imbedded in the overall long term memory structure, thus facilitating activation of any closely related knowledge. The recently active environmental factors and the long term memory closely related to the currently active memory constitute a readily activatable "fringe of consciousness". As the phenomenologist Husserl states, "The focal is girt about with a 'zone' of the marginal;" (Husserl, 1962, p107).

Perception is the class of cognition which is the active interpretation by the organism of the state of a subset of its cells. The cells which are interpreted are either a part of or have a direct, effective link with the organism's surface. In an organism of sufficient complexity, these cells are modality-specific sensory receptors. The status of the sensory receptor cells is the one further determinant of the

selective activation of environmental factors which occurs during perception. An active rule of a perceptual process can activate new environmental factors on the basis of currently active environmental factor meanings and the states of selected receptor cells or receptor cell combinations (receptive fields). The result of this interpretive activity is a symbolic representation which is constructed as an interrelated group of environmental factors.

Conscious experience at any point in time is the internal phenomenon which accompanies (results from) the activation and interpretation of the active rule and environmental factors. Conscious experience is not the contents of active memory. It is not the active and causal "stuff" of cognitive behavior but is merely the epiphenomenal creation of that behavior. This is not to den that conscious experience may itself form a domain of representational knowledge in long term memory. The activation of any part of that representation in the context of an active rule and other environmental factors will necessarily result in a conscious experience which differs from the original (that represented).

Conscious experience is related to the meanings of the symbol structures of active memory. Differences between conscious experience have two sources. One is an actual difference in the contents of active memory (in terms of activated symbols, relations, and structures). The other is that the meanings of the activated long term memory elements have been altered by the intervening evolution of long term memory. Due to the changes of symbol meanings which accompany this evolution, it may not be possible to recreate (reexperience) a conscious event even though the contents of active memory are identical with that of the past experience.

PSYCHOLOGICAL INDICATIONS

A study by Trabasso and Bower (1968) illustrates the flexibility of active memory or the selectivity of "representation-as". They presented subjects with cards showing four letters in a row. The task given was to sort the cards into two piles on the basis of the second letter being L or B, for example. In learning trials, letters were randomly paired with L and B in the other positions. In further trials following learning, the letter in the fourth position was consistently paired with the second position (ie. D with L, E with B). During these trials the subjects were required to recite all letters while sorting the cards on the basis of the second position as before. Following this test, trials were given in which the second position was left blank (no L or B). No subject could correctly sort the cards even though the letters in the fourth position (D or E) were shown and could serve as a basis for sorting. Thus, seeing, recognizing, and verbally repeating the letters did not ensure (cause) selection of the letters in the fourth position as a part of the active environmental factors upon which the action of the active rules, and thus behavior, was based.

Recent research in cognitive psychology has begun to provide evidence (and proposals) for the multiplicity of internal information representation. Tversky (1969) has produced results which indicate the ability of subjects to represent the same stimulus in visually and verbally determined structures in STM. The mode used reflects task requirements and thus indicates flexibility. Paivio (1971) also presents results indicating the use of both visual and verbal memory representations as mediators in association and recall tasks.

The human ability to produce an internal representation which represents things

as is necessary or sufficient for the task at hand is most valuable. As with any organismic characteristic, this ability is only an innate potential. It is predicted that a person without a flexible representational system may not perform as well on some tasks. Such tasks which do not easily conform to their regular mode of representation may prove difficult or insurmountable. Lieberman (1972) carried out a study involving a task which requires mental imagery (of a cube) in coordination with sequential goal-directed behavior to solve a three dimensional maze. Pre-experimental tasks were conducted to classify people as being good spatial imagers, good verbal representers, or both. Preliminary results indicate that having both abilities, thus representational flexibility, shows positive correlation to the degree of success realized on the complex maze task. These findings are not over a statistically significant base and thus are only preliminary.

Two other recent studies appear to indicate representational flexibility. Frost (1971) showed subjects line drawings which varied both in spatial orientation and in semantic category. The instructions led the subjects to expect either a recall or recognition test following the initial learning period. Tests were conducted to determine if the instructions had affected memory organization (representation). The results indicated that they had. Loftus (1971) conducted a similar comparison and had similar results for a task involving continuous paired associates.

RELATED PROPOSALS

Horowitz (1970) has proposed three basic modes of memory representation. The "enactive" mode is used to represent actions, exemplifying features of direction and force. The "image" mode is used to represent spatial information. The "lexical" mode

consists of grammars which can represent sentences (language) and sequential information in general. These modes are derivatives of the "enactive", "iconic", and "symbolic" modes which had been described by Bruner (1964).

Based upon a survey of relevant physiological and neurological research, John (1967) proposes the existence of two representational system classes. The "specific" system is tied to sensory (stimulus - bound) activity as are PIC and VI (iconic memory) of VIPS. The general system reflects repetitive, successful activity not directly tied to sensory input (eg. RLTM and LTM of VIPS). A specific system may interact with a general system during certain behavior (ie. perception by VIPS).

With the return of complex internal representation to credibility from its long period of losing battles with behaviorism, research has produced several books dealing with the applicability of various information representations to tasks of the environment. Visual Thinking by Arnheim (1971) attempts to describe the abilities of visual imagery and illustrates its apparent use by interesting examples. Experiences in Visual Thinking by McKim (1972) presents the reader with a series of tasks to be performed using imagery representations and discusses the apparent applicability of this mode of representation to these. Synectics by Gordon (1961) discusses the applicability of "representation by analogy" to the solution of rather complex creative (engineering) tasks. In The Image of the City, Lynch (1960) investigates the nature of people's memory representations of cities and discusses its relevance to future city design. Claude Levi-Strauss discusses the nature of culturally based totemic symbol structures and the forces which cause their alteration in his book The Savage Mind.

THE ROLE OF REPRESENTATION

This thesis proposes that the symbolic representation or image of any body of information is used to guide behavior with regards to that information. Behavior at a global or strategic level is based upon the structural characteristics of the information representation being employed. The local (actual) "values" of the information structure's elements determine the behavior at the tactical (specific situational) level.

The structural characteristics of an active representation, the types of relations and symbols employed, are fundamental in determining the classes of inferences the cognitive processing system can realize. These readily available inferences, proposals, or expectations are in turn fundamental determinants of behavior. The inferred representational structures of this thesis (and the others) do not only embody efficiently the information relevant to the task given. They also determine what inferences may be easily made concerning that information and serve to guide (shape) the behavior at the strategic level.

Computer science research has investigated and demonstrated the important role of information representation in the accomplishment of a given task. Indeed, anyone who has done any computer programming at all soon becomes aware of the important function of information representation. Various representations have been formally studied and shown to have important effects upon the efficiency (time) of program operation, upon the amount of memory space required, and even upon the ability to successfully complete a given task. These results have appeared in the area of artificial intelligence (Amarel, 1968) and in the computer science discipline in general (Knuth, 1968).

3. Future Research, Conclusion

All aspects of visual perception, imagery, and information representation require further investigation. If any progress can be said to be made by any study now, it is that the study suggests or begins to delimit areas for future experimentation. Experimentation is most fruitfully carried out in light of an existing theory. This serves to generate questions and hypotheses (guides behavior at the strategic level). Though subsequent experimental results may eventually lead to the theory's final discredit, the theory will already have proven its value as the source of investigation, of new knowledge, and of the associated "scientific revolution" (Kuhn, 1962).

This outlook is at odds with the recent statement by R. E. Shaw:

"To a theorist B, machine A₀ is a model of a psychological phenomenon A to the extent that B can use A₀ to answer questions that interest him about A." (Shaw, 1971, p.)

This is not my understanding of science. Rather, a machine is a model to the extent that it generates interesting questions and provides hypothetical answers against which empirical results can be compared and in light of which empirical results may be interpreted. A machine is also a model to the extent that it can be altered (can evolve) to account for subsequent conflicting indications which do not merit revolutionary measures.

By providing a needed and well specified theory of visual form imagery and perception, this thesis can serve as a significant source of research ideas. The theory

has the backing of the local data from which it has been inferred. Also, as indicated in earlier chapters, the attempt has been made to make the inferred processing system consistent with external, relevant psychological research and to explain the visual perception and imagery phenomena observed therein.

In terms of general cognitive system characteristics, further memory research is necessary. Further STM (active memory) research with visually presented stimuli that are not just letters or words which transfer easily to equivalent verbal codes is needed to help determine the amount of information retainable and to gain further ideas concerning its possible representation. Further research is needed to better determine the structure and operating characteristics of recognition long-term memory. To answer such questions as, "What partial information is sufficient to suggest an object's existence (recognition)?" "How easily can new objects for recognition be entered and how is this accomplished?" "Is the operation of RLTM serial (sequential traversal of a decision net, as proposed here) or parallel?" further results are needed.

Active experimentation is also suggested with regards to the determination of the existence and characteristics of an intermediate term memory. This memory has been proposed here as it has proven useful in accounting for the observed behavior. There is still insufficient evidence of its existence and secondarily of its characteristics. The existence of an intermediate term memory with the capability of maintaining recent context factors in a readily accessible state appears likely. The existence of TCLTM as defined in Chapter 8 also appears likely to me. The context capability (characteristic) of cognition is not satisfactorily explained by a memory proposal solely involving the classic short-term and long-term memories.

As for the proposals relevant to the explanation of human visual information processing (imagery, perception), several areas for research are indicated. Further investigation of the role played by an active cognitive goal in visual processing is necessary. The effect of goals upon the direction of attention and the representation of the realized perception requires further investigation and specification.

Research with regards to the functional role of peripheral information in visual perception is necessary. One experiment could involve the masking of a pre-set fixation area to help determine the type of information contained in peripheral inputs and to what use this information can be put by the visual system. Indeed, can visual form perception be realized when only peripheral input is allowed? If so, how is the process altered? More neurological studies to determine nerve and nerve-complex detection abilities of the eye can also yield some needed input to theory development.

Finally, further experimentation is needed to investigate the information representation capabilities of the human. Response time experiments, operating under the assumption that the length of time to derive an answer is directly related to the directness of the answer's representation, can yield valuable results. Experiments which induce one informational goal and then probe the ability to answer unexpected questions concerning the same body of information, may yield further interesting results as to the flexibility and specificity of human memory representations.

Thus, by developing a completely operational visual system based upon inferences made from observed subject behavior, a more fully developed perceptual theory is here proposed than is possible from more restricted experiments. Though much of the theory is highly questionable, its value lies in the areas of research which it outlines for further inquiry toward understanding.

REFERENCES

- [Adelman & Smith, 1971] Expectancy as a determinant of functional units in perceptual recognition. *Cognitive Psychology*, 1971, 2, 117-129.
- [Amarel, 1968] S. Amarel. On representations of problems of reasoning about actions. *Machine Intelligence 3*, New York: American Elsevier, 1968.
- [Anderson & Bower, 1973] John R. Anderson & Gordon H. Bower. *Human associative memory*. Washington, D.C.: Winston, 1973.
- [Arnheim, 1956] Rudolph Arnheim. *Art and visual perception*. Berkeley, Cal.: University of California Press, 1956.
- [Arnheim, 1969] Rudolf Arnheim. *Visual thinking*. Berkeley, Cal.: University of California Press, 1969.
- [Attneave, 1954] F. Attneave. Some information aspects of visual perception. *Psychological Review*, 1954, 61, 183-193.
- [Attneave, 1957] F. Attneave. Physical determinants of the judged complexity of shapes. *Journal of Experimental Psychology*, 1957, 53, 221-227.
- [Attneave & Frost, 1969] The determination of perceived tridimensional orientation by minimum criteria. *Perception and Psychophysics*, 1969, 6, 391-400.
- [Bartlett, 1917] Frederic C. Bartlett. An experimental study of some problems of perceiving and imaging. *British Journal of Psychology*, 1917, 8, 222-266.
- [Bartlett, 1932] Frederic C. Bartlett. *Remembering: a study in experimental and social psychology*. Cambridge: Cambridge University Press, 1932.
- [Baylor, 1971] George W. Baylor. A treatise on the mind's eye: an empirical investigation of visual mental imagery. Ph.D. Thesis, Carnegie-Mellon University, Department of Psychology, 1971. Also issued as three reports from the Université de Montréal, Institut de Psychologie, 1969-1971.
- [Bower, 1972] Gordon H. Bower. Mental imagery and associative learning. In [Gregg, 1972], 51-88.
- [Brooks, 1968] Lee R. Brooks. Spatial and verbal components of the act of recall. *Canadian Journal of Psychology*, 1968, 22, 349-368.

- [Bruner, 1964] Jerome S. Bruner. The course of cognitive growth. American Psychologist, 1964, 19, 1-15.
- [Bugelski, 1971] B R Bugelski. The definition of the image. In [Segal, 1971], 49-68.
- [Buswell, 1935] How people look at pictures. Chicago: University of Chicago Press, 1935.
- [Chase, 1973] William G. Chase, editor. Visual information processing. New York: Academic Press, 1973.
- [Chase & Clark, 1972] William G. Chase & Herbert H. Clark. Mental operations in the comparison of sentences and pictures. In [Gregg, 1972], 205-232.
- [Clark & Chase, 1972] Herbert H. Clark & William G. Chase. On the process of comparing sentences against pictures. Cognitive Psychology, 1972, 3, 472-517.
- [Craig, 1943] K.J.W. Craig. The nature of explanation. Cambridge: University Press, 19743.
- [Ebbinghaus, 1885] Memory: a contribution to experimental psychology. Translation: New York: Dover, 1964.
- [Fagan, 1972] J. F. Fagan. Infant's recognition memory for faces. Journal of Experimental Child Psychology, 1972, 14, 453-472.
- [Farley, 1971] Arthur Farley. Toward an understanding of visual depth perception: a survey, Unpublished report.
- [Feldman, 1961] Simulation of behavior in the binary choice experiment. Proc. Western Joint Comp. Conf., 1961, 19, 133-144.
- [Fernald, 1912] Mabel F. Fernald. The diagnosis of mental imagery. Psychological Monograph, 1912, 14 (Whole No. 58).
- [Festinger, et al., 1967] Efferece and the conscious experience of perception. Journal of Experimental Psychology Monograph, 1967, 637.
- [Frost, 1971] Interaction of visual and semantic codes in long term memory. Unpublished Ph.D. thesis, University of Oregon, 1971.
- [Gibson, 1966] James J. Gibson. The senses considered as perceptual systems. Boston, Mass.: Houghton Mifflin Company, 1966.
- [Gombrich, 1960] Ernest H. Gombrich. Art and illusion. New York: Pantheon Books, 1960.
- [Goldiamond & Hawkins, 1958] Goldiamond, I. and Hawkins, W.F. Versuchsversuch: The

log relationship between word-frequency and recognition obtained in the absence of stimulus words. *Journal of Experimental Psychology*, 1958, 56, 457-483.

- [Goodman, 1968] Nelson Goodman. *Languages of art: an approach to a theory of symbols*. Indianapolis, Indiana: Bobbs-Merrill, 1968.
- [Gordon, 1961] *Synecdoche, the development of creative capacity*. New York: Harper, 1961.
- [Gould & Dill, 1969] Eye movement parameters and pattern discrimination. *Perception and Psychophysics*, 1969, 6, 311-320.
- [Gregg, 1972] Lee W. Gregg, editor. *Cognition in learning and memory*. New York: Wiley, 1972.
- [Gregory, 1966] Richard L. Gregory. *Eye and brain: the psychology of seeing*. New York: McGraw-Hill, 1966.
- [Gregory, 1970] Richard L. Gregory. *The intelligent eye*. New York: McGraw-Hill, 1970.
- [Haber, 1968] Ralph Norman Haber, editor. *Contemporary theory and research in visual perception*. New York: Holt, Rinehart & Winston, 1968.
- [Haber, 1969] Ralph Norman Haber, editor. *Information processing approaches to visual perception*. New York: Holt, Rinehart & Winston, 1969.
- [Haber & Hershenson, 1973] Ralph Norman Haber & Maurice Hershenson. *The psychology of visual perception*. New York: Holt, Rinehart & Winston, 1973.
- [Hayek, 1952] Frederic Hayek. *The sensory order*. Chicago: University of Chicago Press, 1952.
- [Hebb, 1949] Donald O. Hebb. *The organization of behavior*. New York: Wiley, 1949.
- [Hebb, 1968] Donald O. Hebb. Concerning imagery. *Psychological Review*, 1968, 75, 466-477.
- [Hochberg, 1968] Julian E. Hochberg. In the mind's eye. in [Haber, 1968], 309-331.
- [Hochberg, 1970] Attention, organization, and consciousness. In Mostofsky (ed.), *Attention: contemporary theory and analysis*. New York: Appleton, 1970.
- [Hochberg, 1973] Elementary units in the visual perception of objects. In Mehler (ed.), *Handbook of Cognitive Psychology*. Englewood Cliffs, New Jersey: Prentice Hall, 1973.
- [Holt, 1964] Robert R. Holt. Imagery: the return on the ostracized. *American Psychology*, 1964, 19, 254-264.

- [Hendricks, 1970] Image formation and cognition. New York: Appleton-Century-Crofts, 1970.
- [Hubel & Wiesel, 1962] David H. Hubel & T. N. Wiesel. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. Journal of Physiology, 1962, 160, 106-123. Reprinted in [Hubel, 1964], 260-277.
- [Hunt, 1971] Earl B. Hunt. What kind of computer is man? Cognitive Psychology, 1971, 2, 57-58.
- [Musser, 1931] Images. New York: Dover Press, 1962 ed.
- [James, 1890] William James. The principles of psychology. New York: Dover Press, 1890.
- [John, 1967] E. Roy John. Mechanisms of memory. New York: Academic Press, 1967.
- [John, 1971] E. Roy John. Brain mechanisms of memory. In Psychobiology. New York: Academic Press, 1971.
- [Kintsch, 1972] Walter Kintsch. Notes on the semantic structure of memory. in [Tulving and Donaldson], 1972.
- [Knuth, 1968] Donald E. Knuth. The art of computer programming. Volume 1: Fundamental Algorithms. Reading, Mass.: Addison-Wesley, 1968.
- [Lashley, 1942] Kurt S. Lashley. The problem of cerebral organization in vision. Biological Symposium, 7, pp 301-322, 1942.
- [Leibovic, 1972] K. N. Leibovic. Nervous system theory: an introductory study. New York: Academic Press, 1972.
- [Lettvin, et al., 1959] Lettvin, Maturana, McCulloch, and Pitts. What the frog's eye tells the frog's brain. Proceedings IRE, 47, pp 1940-1951, 1959.
- [Lieberman, 1972] Robert Lieberman. Behavior on a complex three-dimensional maze task. Unpub. report, Computer Science Dept., Carnegie Mellon University, 1972.
- [Loflus, 1971] Geoffrey R. Loftus. Comparison of recognition and recall in a continuous memory task. Journal of Experimental Psychology, 91, pp 355-365, 1971.
- [Lynch, 1960] Kevin Lynch. The image of the city. Cambridge, Mass.: MIT Press, 1972.
- [Michels and Zusne, 1965] Metrics of visual form. Psychological Bulletin, 63, pp 74-86, 1965.

- [Miller, 1956] George A. Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-96. Reprinted in several books.
- [Miller, 1967] George A. Miller. The psychology of communication. New York: Basic Books, 1967.
- [Miller, 1973] Robert J. Miller. Cross-cultural research in the perception of pictorial materials. Psychological Bulletin 1973, 80(2), 135-150.
- [Miller, Galanter, & Pribram, 1960] George A. Miller, Eugene Galanter, & Karl H. Pribram. Plans and the structure of behavior. New York: Holt, Rinehart & Winston, 1960.
- [Minsky, 1968] Marvin L. Minsky, editor. Semantic information processing. Cambridge, Mass.: M.I.T. Press, 1968.
- [Minsky & Papert, 1969] Marvin L. Minsky & Seymour Papert. Perceptrons. Cambridge, Mass.: M.I.T. Press, 1969.
- [Moran, 1973] Thomas P. Moran. The symbolic nature of visual imagery. Stanford, Cal.: Proceedings of the Third International Joint Conference on Artificial Intelligence, 1973.
- [Muntz, 1971] W.R.A. Muntz. Sensory processes and behavior. In Psychobiology, New York: Academic Press, 1971.
- [Neisser, 1967] Ulric Neisser. Cognitive psychology. New York: Appleton-Century-Crofts, 1967.
- [Neisser, 1972] Ulric Neisser. Changing conceptions of imagery. In [Sheehan, 1972], 233-251.
- [Newell, 1972] Allen Newell. A theoretical exploration of mechanisms for coding the stimulus. Department of Computer Science, Carnegie-Mellon University, 1972. Reprinted in [Melton & Martin, 1972], 373-434.
- [Newell, 1973] Allen Newell. Production systems: models of control structures. In [Chase, 1973], 463-526.
- [Newell & Simon, 1972] Allen Newell & Herbert A. Simon. Human problem solving. Englewood Cliffs, N.J.: Prentice Hall, 1972.
- [Noton & Stark, 1971] David Noton & Lawrence Stark. Saccades in eye movements during pattern perception. Science, 1971, 171, 308-311.
- [Paivio, 1971] Allan Paivio. Imagery and verbal processes. New York: Holt, Rinehart & Winston, 1971.

- [Perky, 1910] C.W. Perky. An experimental study of imagination. American Journal of Psychology, 1910, 21, 422-452
- [Piaget, 1967] Jean Piaget. Six psychological studies. New York: Random House, 1967
- [Piaget, 1970] Jean Piaget. Structuralism. New York: Basic Books, 1970
- [Posner, 1969] Michael I. Posner. Abstraction and the process of recognition. In Bower and Spence (eds.), The psychology of learning and motivation, Volume 3. New York: Academic Press, 1969
- [Posner, 1973] Michael I. Posner. Coordination of internal codes. In [Chase, 1973], 35-73.
- [Posner and Keele, 1968] On the genesis of abstract ideas. Journal of Experimental Psychology, 1968, 77, 353-363.
- [Posner and Mitchell, 1967] Chronometric analysis of classification. Psychological Review, 1967, 74, 392-409
- [Pribram, 1971] Karl Pribram. Languages of the brain. Englewood Cliffs, N.J. Prentice Hall, 1971.
- [Pylyshyn, 1973] Zenon W. Pylyshyn. What the mind's eye tells the mind's brain: a critique of mental imagery. Psychological Bulletin, 1973, 80, 1-24
- [Quillian, 1968] M. Ross Quillian. Semantic memory. In [Minsky, 1968]
- [Rabinow, 1957] J. Rabinow. Optical coincidence devices. U.S. Patent 2, 795, 705. June 11, 1957.
- [Reed, 1973] Stephen K. Reed. Psychological Processes in Pattern Recognition. New York: Academic Press, 1973
- [Reitman, 1971] Judith S. Reitman. Mechanisms of forgetting in short-term memory. Cognitive Psychology, 1971, 2, 185-195
- [Richardson, 1969] Alan Richardson. Mental imagery. New York: Springer-Verlag, 1969
- [Roberts, 1960] L. G. Roberts. Pattern recognition with an adaptive network. I.R.E. Conf. Rec., 1960, in [Uhr, 1966]
- [Segal, 1971] Sydney Joelson Segal, editor. Imagery: current cognitive approaches. New York: Academic Press, 1971
- [Seltridge, 1959] O.G. Seltridge. Pandemonium: A paradigm for learning. Reprint in [Uhr, 1966]

- [Shaw, 1971] Robert E. Shaw. Cognition, simulation, and the problem of complexity. *Journal of Structural Learning*, 1971, 2, 31-44.
- [Sheehan, 1972] Peter W. Sheehan, editor. *The function and nature of imagery*. New York: Academic Press, 1972.
- [Shiffrin, 1970] R. M. Shiffrin. Memory search. In D. A. Norman *Models of human memory*, New York: Academic Press, 1970.
- [Shiffrin and Atkinson, 1969] R. M. Shiffrin and R. C. Atkinson. Storage and retrieval processes in long-term memory. *Psychological Review*, 1969, 56, 179-193.
- [Shiffrin & Geisler, 1973] Richard M. Shiffrin & Wilson S. Geisler. Visual recognition in a theory of information processing. In [Solso, 1973], 53-101.
- [Simon, 1967] Herbert A. Simon. An information-processing explanation of some perceptual phenomena. *British Journal of Psychology*, 1967, 58, 1-12.
- [Simon, 1972] Herbert A. Simon. What is visual imagery? an information processing interpretation. In [Gregg, 1972], 183-204.
- [Solso, 1973] Robert L. Solso, editor. *Contemporary issues in Cognitive Psychology: the Loyola symposium*. Washington, D.C.: Winston, 1973.
- [Sperling, 1960] George Sperling. The information available in brief visual presentations. *Psychological Monographs*, 1960, 74 (11, Whole No. 498).
- [Sternberg, 1967] Saul Sternberg. Two operators in character recognition: some evidence from reaction time measurements. *Perception and Psychophysics*, 1967, 2, 45-53.
- [Sperling, 1969] S. Sternberg. The discovery of processing stages: Extensions of Donder's method. *Acta Psychologica*, 1969, 276-315.
- [Stromeyer & Portka, 1970] The detailed texture of eidetic images. *Nature*, 1970, 225, 346.
- [Titchener, 1909] E.B. Titchener. *Lectures on the Experimental Psychology of the thought processes*. New York: MacMillan, 1909.
- [Trabasso & Bower, 1968] T. Trabasso and G.M. Bower. *Attention in learning, theory and research*. New York: Wiley, 1968.
- [Tulving, 1972] Ender Tulving. Episodic and semantic memory. In [Tulving & Donaldson, 1972], 351-403.
- [Tulving & Donaldson, 1972] Ender Tulving & Wayne Donaldson, editors. *Organization of memory*. New York: Academic Press, 1972.

- [Tversky, 1969] Barbara Tversky. Pictorial and word encoding in a short term memory task. Perception and Psychophysics, 1969, 6, 225-233.
- [Uhr, 1966] Leonard Uhr, editor. Pattern recognition. New York: Wiley, 1966.
- [Uhr, 1973] Leonard Uhr. Pattern recognition, learning and thought. Englewood Cliffs, N.J.: Prentice Hall, 1973.
- [Weisstein, 1973] Naomi Weisstein. Beyond the yellow-Volkswagen detector and the grandmother cell: a general strategy for the exploration of operations in human pattern recognition. In [Solso, 1973], 17-51.
- [Waltz, 1972] David L. Waltz. Generating semantic descriptions from drawings of scenes with shadows. Ph.D. Thesis, Department of Electrical Engineering, M.I.T., 1972. Project MAC, Technical Report TR-271.
- [Watson, 1913] J.B. Watson. Psychology as the behaviorist views it. Psychological Review, 1913, 20, 158-177.
- [Waugh & Norman, 1965] Nancy C. Waugh & Donald A. Norman. Primary memory. Psychological Review, 1965, 72, 89-104.
- [Wickelgren, 1972] W.A. Wickelgren. Coding, retrieval and dynamics of multitrace associative memory. In [Gregg, 1972].
- [Williams, 1967] L.G. Williams. The effects of target specification on objects fixated during visual search. Acta Psychologica, 1967, 27, 355-360.
- [Winston, 1970] Patrick H. Winston. Learning structural descriptions from examples. Ph.D. Thesis, Department of Electrical Engineering, M.I.T., 1970. Project MAC, Technical Report TR-76.
- [Yerbus, 1967] A.L. Yerbus. Eye movements and vision. New York: Plenum Press, 1967.
- [Zusne, 1970] Leonard Zusne. Visual perception of form. New York: Academic Press, 1970.