# MEASURING TEXT-GRAPHIC ACTIVITY

Fred H. Lakin

Rehabilitation R&D Center, Palo Alto VA Medical Center
3801 Miranda Ave, Palo Alto, California, USA

## ABSTRACT

Interface designers and psychologists need to be able to formulate hypotheses about user activity in graphics interfaces and then experimentally test them. The results of such experiments are both interesting in their own right AND can be used to change the design of interfaces so as to better support the graphics activity taking place. In order to do this, what is needed is a high level representation for the activity of graphics interface use. In terms of such a representation: 1] hypotheses can be described; 2] interface activity can be measured to test the hypotheses; and perhaps even 3] changes to the interface design can be specified. This paper presents a structural model of writing and drawing which provides a metric for measuring such activity and is designed to serve the three purposes above.

In the SAM model, the product of writing and drawing is simplified to 'text-graphic objects', and the activity of writing and drawing becomes 'text-graphic manipulation'. The text-graphic objects have structure. This structure arises directly from an attempt to account for the manual manipulations observed in non-computer image production such as occurs on blackboards. According to the model, the needs of manual manipulation determine the text-graphic pattern as the simplest organizing structure for images. SAM stands for Structure-Arises-out-of-Manipulation. Included in the SAM model is a notation for the structure of text-graphic objects. This notation allows high level description of blackboard type image activity.

A graphic structure editor based on the SAM model has been defined. The editor was implemented in PAM, a language which generalizes LISP to handle text-graphic objects (PAM stands for PAttern Manipulation).

The model-based editor has been used to provide measurement of and interactive assistance for text-graphic manipulation. The simplest measurement is simply a chronological record of each successive manipulation and image state. The lowest level assistance is structure based agility aids. Next, direct user manipulation of structure is facilitated. And at higher levels, the editor is a tool for exploring the rules used by humans to collect elementary visual objects into conceptual groups. Examples of such groupings discovered by analysis of user editor activity are presented in the paper.

Implementations have been done in MACLISP, Smalltalk and Franz LISP.

KEYWORDS: graphics interface, measuring text-graphic activity.



## SECTIONS



## I. PROBLEM DEFINITION: MEASURE WRITING AND DRAWING ACTIVITY

The goal here is to measure 'writing and drawing activity'. In order to define this goal more precisely, an instance of such activity will be selected and examined.

### Blackboard Activty Is Prototypical Of Writing And Drawing

Human beings often write and draw to facilitate cognitive tasks. If they do not use a specific visual language (VLSI symbology, flowcharts, drafting, etc) then they are doing general purpose writing and drawing. Blackboards are often used in this way. The phrase 'blackboard activity' will be used to describe the live, spontaneous imaging that takes place on blackboards in meeting rooms and classrooms. Such images are 'colloquial' or 'natural'; they are general purpose as opposed to the visual jargon of formal visual languages like those handled by CAD/CAM or VLSI systems.

Large wall displays created by groups provide a rich area for measurement of writing and drawing:
  1] The desire to communicate about the task encourages the group to fully display task-related cognition on the external display.
  2] The agility requirements are a good test of representation.
  3] Blackboard activity uses specific visual languages as components.

In order to keep track of complex topics, working groups find that a display can function as an 'external group memory' [Ball71]. The operator must make sure that all significant aspects of the group's task-related cognition are represented on the external display. This completeness serves the group's members who are working together (and also, of course, researchers using records of the display as corpus or protocol).

Any of the popular visual symbol systems CAN be called into play during blackboard activity. Consider figure 1, which has text paragraphs, vertical lists, arrows, cartoons, Venn diagrams, freehand machine drawings, etc. A measuring system must have the generality to take into account these various dialects of imaging.

## An Example Of Blackboard Activity

To clarify this problem domain (ie to provisionally indicate that which is to be measured), we need a specific example of writing and drawing. Figure 1 shows a group having a discussion, and being aided in their thinking by a wall display which reflects their thoughts. In this case the display is not a blackboard but a large piece of paper because it supports crisper images and photographs more clearly. The blackboard activity which generated the final frame in figure 1 will be taken as typical for this application domain. The image in figure 1 took approximately 23 minutes to create using dry markers and newsprint.

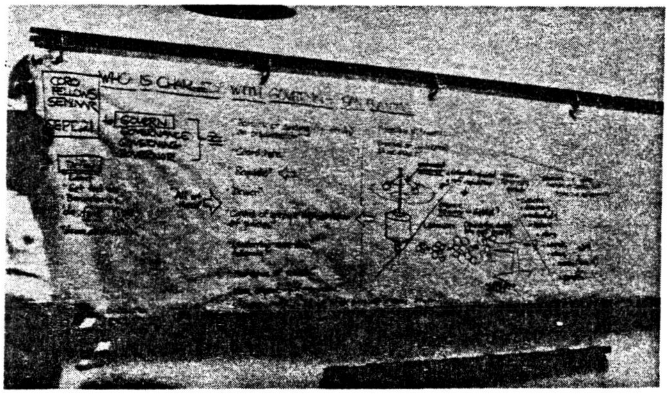


Figure 1. 'Blackboard activity' on a paper display

## Chunking The Objects In Blackboard Activity

A method was needed to study the pheonemon of nonspecific writing and drawing as it was practiced 'in the field' on existing equipment. The approach used was to take time-lapse photographs of the development of the display in figure 1. The chronology revealed the manipulable units for the operator, and may reflect conceptual grouping as well. This study was useful for putting both agility and the nature of objects in more concrete terms. Figure 2 shows a few diagramming times and objects for figure 1.

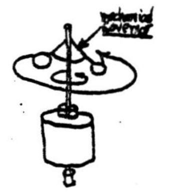
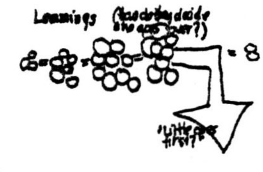
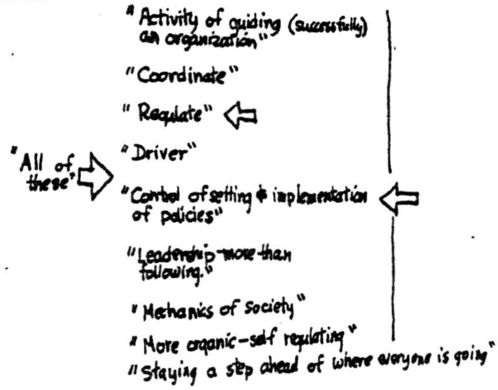
| Time to create: | Object: |
|---|---|
| approx 23 min | entire image |
| 5Ø seconds |  |
| 7Ø seconds |  |
| 2 min 3Ø seconds |  |



Figure 2. Objects and times from blackboard activty

This study shows that measurement of writing and drawing must deal with the the structures that the human uses in seeing and understanding the display. Those structures are much more complex than rectangular bit patches; they are reflected in the kind of visual objects listed in figure 2. At its simplest, blackboard activity is concerned with groups and sub-groups of objects which maintain their identity when overlapping one another. "Little ones first" (the text label at the tip of the big hollow arrow in the 7Ø second object) is an object which goes with the arrow. It can overlap it or not; it's identity remains constant. Intelligent assistance of blackboard activity will only come with an interface whose 'knowledge' of objects on the display corresponds with the user's knowledge.

## The Purpose of Measuring Writing and Drawing

To further clarify the purpose of the measuring proposed here, a comparison with text is appropriate. The 'rules' and 'units' for generalized text editing are well enough understood that the same editor can be used for many different kinds of text manipulation. This is a victory for the representation used, which provides an easy and unappreciated way to measure 'text activity'. We don't have character string editors -- we have 'text' editors, where 'text' implies words, sentences and paragraphs as organizing super-structures for characters. Due to the power of this underlying representation, text editors have some general features which are good for virtually all domains -- letters, poems, programs, forms, etc. The point of the next section is to attempt a model of writing and drawing which can supply the same representational power as the idea of 'text' does for generalized text editors.

## II. PROBLEM SOLUTION: A MODEL OF TEXT-GRAPHIC MANIPULATION

A model is proposed which provides a frame of reference for defining and measuring writing and drawing activity.

### The SAM Model: Structure Arises Out Of Manipulation

In the SAM model, the product of writing and drawing is simplified to 'text-graphic objects', and the activity of general purpose writing and drawing becomes 'nonspecific text-graphic manipulation'. The text-graphic objects have structure. This structure arises directly from an attempt to account for the manual manipulations observed in figure 1 (described above). According to the model, the needs of manual manipulation determine the text-graphic pattern as the simplest organizing structure for images. SAM stands for Structure-Arises-out-of-Manipulation.

Given an object of attention in a context, and selective manipulation, structure follows by necessity. To be able to point at one object and command it to move in relation to the other objects on the screen implies a part/rest distinction understood by both human and computer. Structure is the visual rule used in SAM for making part/rest distinctions. Nothing like selective manipulation can even take place without structure: the capability to have visual atoms and groups (patterns), and to selectively manipulate them is essentially a structural phenomenon.

The SAM model: the human uses the hand controls to manipulate his object of attention on the text-graphic display.

This model is schematically illustrated in figure 3. The figure shows a top level representation of a human manipulating text-graphic objects. The 'object of attention' is indicated by being darker and thicker in the illustration.

### Consequences Of The Sam Model

To briefly summarize the consequences of the model in figure 3:

* Focus of attention = object of attention
* Context of attention = text-graphic display
* Structure = tree structure for all objects
* History of user's manipulations =
  a] chronological record of 'moves' (control movements)
  b] chronological record of 'snapshots' of successive states of the object of attention
  c] chronological record of 'snapshots' of successive states of the text-graphic display
  d] chronological record of the grouping structures for successive states of the text-graphic display

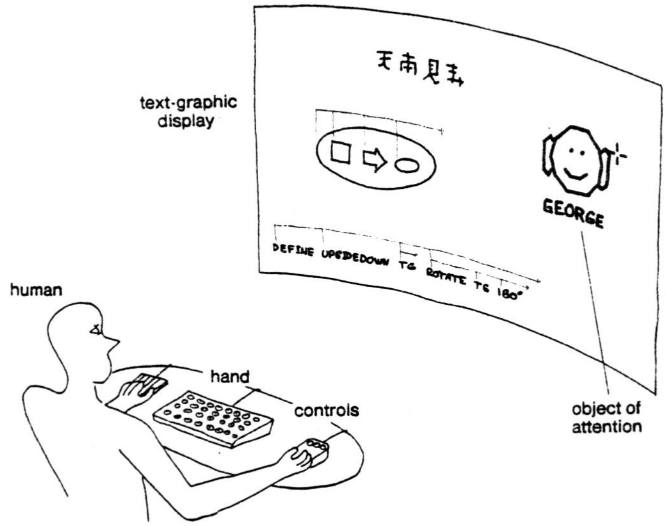



Figure 3. SAM model of text-graphic manipulation

Everything on the display in figure 3 is a text-graphic object. And exactly what is an object? Figure 4 shows the logic of manipulable objects in the SAM model. Text-graphic integration is important to the model; a single structure underlies both text and graphics. Note the three levels of atoms (atomic objects are all called lines in SAM): drawlines, characters, and textlines. Although this will seem illogical to someone programming a vector graphics display, it makes sense from a user's ordinary orientation in which a line of text and a drawn line are both indivisible

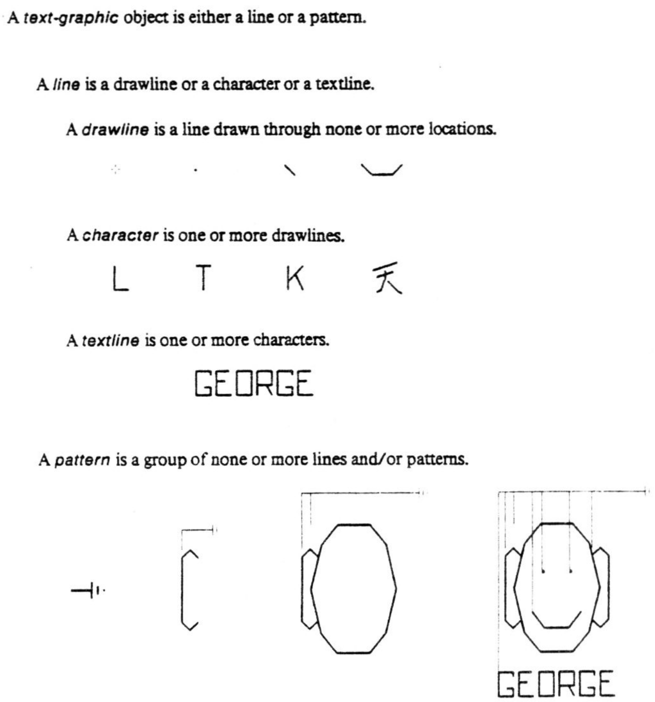
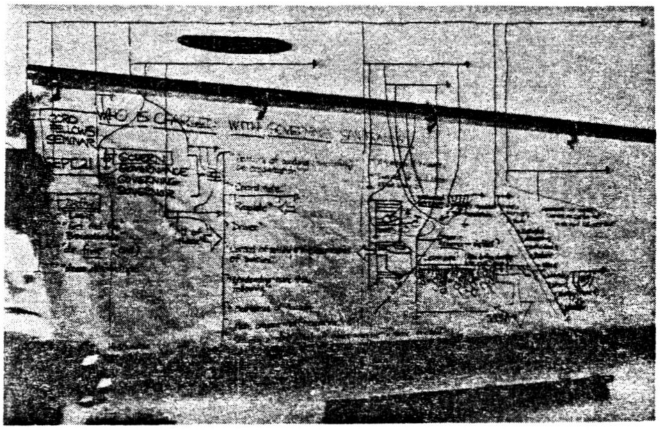



Figure 4. Structure from manipulation for text-graphic objects

atomic objects for manipulation. Central to the model is the tree notation for the structure of objects. This is important, as the units for text-graphic activity will be the basis for measuring it (even as characters, words, sentences and paragraphs provide manipulative structures which go across many different text manipulation activities).

### Grouping Structures for the Blackboard Image

The tree notation from figure 4 can now be used to diagram for measurement structures in the text-graphic objects produced by writing and drawing. For example, figure 5 shows a logical structure for the natural blackboard image from figure 1 . Here we can readily see the kind of objects that humans think about and want to manipulate. The 7 top-level members are all themselves patterns; and 2 of those member patterns are quite complex (each nested to a level of 5).



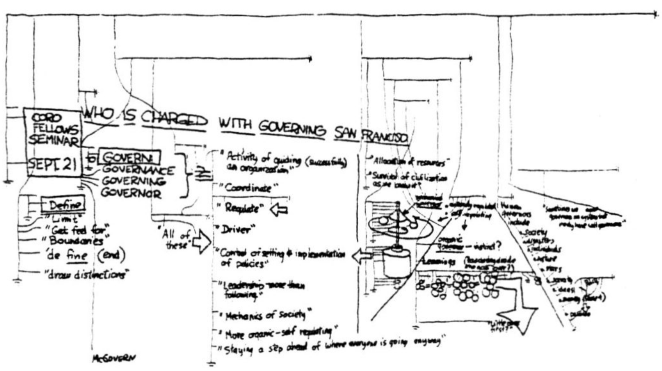



Figure 5. Tree notation showing structure for blackboard activity in figure 1

### III. APPLICATION: MEASURING TEXT-GRAPHIC MANIPULATION

In order to facilitate measurement of text-graphic manipulation, a graphic structure editor has been built which records the dynamics of user image activity as defined by the SAM model.

### Embodiment Of The Model: handPAM, A Graphics Editor

Now let us consider handPAM (for hand PAttern Manipulation) as an embodiment of the SAM model [Lakin80a, Lakin80c, Lakin81]. It is a manual instrument based on the simple visual logic from the model. In handPAM the user's object of attention has become the OBJATN, a global variable. 'Driving attention around the display' is simply interactively changing the binding of that variable using the handPAM controls. And the context of his attention has become BIGPAT, the largest pattern containing every object on the display.

In this context handPAM offers DRAWING to create graphic objects, TEXTING to create text objects, spatial GRABBING of objects into attention, tree guided attention shifters like FIRST, REST, NEXT, and UP, and spatial & tree manipulations of any object in attention.

handPAM implementation: The text-graphic editor monitors the movements of the keyset, keyset and mouse, and uses then as commands to manipulate the OBJATN in BIGPAT.

handPAM is not unlike an piano (with its notes, octaves, keys and scales) in that there is a structure to the visual objects handPAM manipulates, a structure which both enables that manipulation and is inherent to the objects manipulated (figure 4). This system is a combination of two popular approaches for representing diagrams: it uses a linked list structure, and yet the objects thus linked are so literally visual that geometric calculations can be easily done on them as they stand (in fact, the display processor uses the representation to paint the screen). handPAM is actually a generalization of Warren Teitelman's InterLISP editor to text-graphic patterns on a static display (Teitelman credits Peter Deutsch for the original idea of a structure editor, [Teitelman78]). The implementation used to produce the images in figures 6 and 8 was done in Smalltalk-76 at Xerox PARC [Lakin80b].

### Image Example: The Project Diagram

Figure 6 shows an image created using handPAM. The agility of handPAM is adequate for blackboard activity; figure 6 only took 17 and one half minutes to create. The figure is a project diagram " ... for what was at the time a proposed project" [Henderson82].

What emerges from study of such examples of handPAM usage is that structure offers a shared framework of orientation with respect to purposeful manipulation of the image. This framework is of enormous help to handPAM in knowing what the user means when she signals "DRAG THAT HERE!". Structure supports agilty because it offers humans a way of getting their hands on the image.

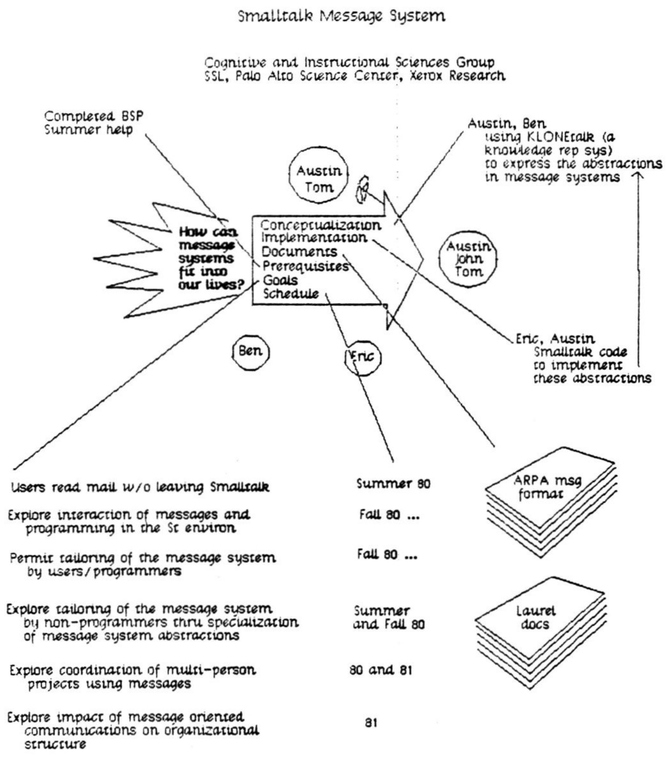



Figure 6. The 'Project Diagram' created on handPAM

## Generalizing LISP to Visual Objects

handPAM is one aspect of the complete PAM system; the other aspect is writtenPAM, the programming language used to implement handPAM. writtenPAM is as a generalization of LISP from textual symbolic expressions to text-graphic forms (objects). The whole PAM system is designed for the manipulation of text-graphic patterns -- first manually, and then, later, programmatically. In writtenPAM, as in LISP, programmatic processing power is based upon providing atomic objects, ways of structuring them into complex objects, and equality tests for both atomic and complex objects. This leads eventually to 'computing with text-graphic forms' [Lakin80c]. It also means that processing power is available for doing analysis of diagrams as text-graphic objects (and diagramming as text-graphic manipulation). Such processing will be necessary when writing programs which automatically measure text-graphic activity.

## Measurement In The handPAM Environment

First it seemed that structure merely offered humans a way of getting their hands on the image. But now it turns out that in thus getting their hands on the image, they leave 'tracks'. Given sufficient structural agility, users will create groupings wherever they need them. The structure left behind reflects their grappling with the image to manipulate it for their cognitive purposes. Indeed, the structure offers a shared framework of orientation (for user and handPAM) precisely because it has been formed through purposeful manipulation of the image.

To take full advantage of these structures, we must measure them. Measuring involves: first, recording the structural dynamics of a user session; and second, attempting to determine which of those dynamics reflect the cognitive groupings of the user.

The handPAM environment facilitates this task in a semi-Procrustrean fashion (figure 7). Since the human needs an artifact to manipulate text and graphics anyway, why not supply one which makes for easier measuring of the use of the artifact? The paramount task in measurement is to define the units of text-graphic activty -- the temporal and spatio-visual chunks which are meaningful to the user. A recording of the user session in these units is a history. A history of a user session can be helpful to both the user and researchers. For the user, it provides an InterLISP style history of events; for researchers, a corpus of visual linguistic activity for analysis.

This 'chunkwise' history list provides cognitive psychologists with a tool for exploring the role of graphics and external memory in problem solving. As small insights are garnered toward this end, then they can be applied to improving the system's representation of what the user is up to, and supplying assistance based on this knowledge.

```
   structuring and manipulating text-graphic objects

                    Text-Graphic Instrument
                          /          \
                         /            \
                        /              \
user                   /                \
entree => Performing                Recording

     immediate hand tool        measurement & analysis
```

Figure 7. A medium for graphic activity that measures

## Measuring Structures In The 'Project Diagram' Activity

Figure 8 shows the tree structure which resulted from the spontaneous structural manipulations of the user. This is an example of structural agility; the groupings were created 'on the fly', part and parcel with the rest of the image. The 17 and one half minutes creation time includes all structural manipulations. The structure in figure 8 consists of 27 top level members of BIGPAT, some of which are themselves patterns.

Let's take another look at how figure 8 shows the manipulative structures for the 'Project Diagram' in figure 6. Note that the tree notation gives a clear display of the development chronology of the display: the first object created is pointed to by the first branch at the upper left, and the last object is pointed to by the last branch at the upper right (the ground-sign indicates the tail of the pattern). Likewise the sub-groupings in the image are denoted by sub-branches. The image has 27 top level members, 8 of which are themselves patterns. 1 of the these patterns is complex, having 2 members which are simple patterns. Each pattern was a manipulative unit for the user during the screen lifetime of the image. Most of the groupings make manipulative AND conceptual sense, with perhaps the exception of the objects surounding the hollow arrow. It is not clear why all those objects -- the 4 circles with names in them, the arrow itself, the list inside it, the steering wheel

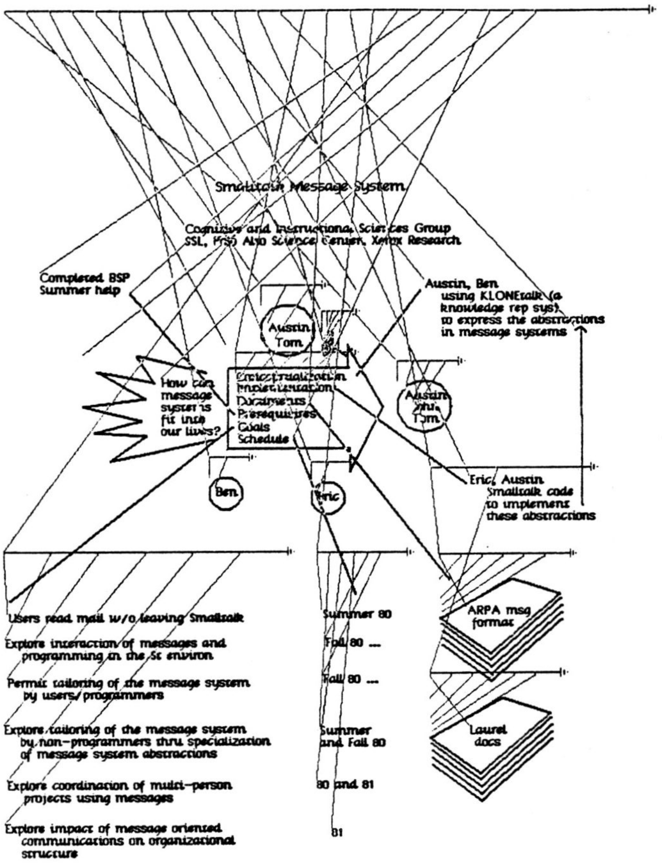



Figure 8. Manipulative structures for the 'Project Diagram'

on top of the arrow, and the 'flame' with its message -- were not collected into a higher pattern. The explanation is that the user never had to move those objects and thus never collected them for manipulative purposes.

By careful measurement of assisted displays, we may get a handle on notions like what a 'visual sentence' is. The user contributed groupings in figure 8 are the first step. For instance, perhaps each top-level object can be considered a complete 'visual paragraph'. Some of the 'paragraphs' have only one member -- a 'visual sentence' -- while others have more than one.

## Dynamics Of Image And Structure

Thus far the discussion of measurement has been confined to image and grouping structure analysis of single, frozen frames (figures 5, 8). But using the visual computational power of the handPAM environment, dynamic records can be kept of the user's activity, showing the evolution of images over time.

Figures 9 and 1Ø diagram the general schema for a 'dynamic visual corpus'. Figure 9 shows successive frames recording entire image appearance, OBJATN, and cursor position ( ⊕ ) over time. Figure 1Ø shows the concurrent grouping structures over time for the same image sequence.

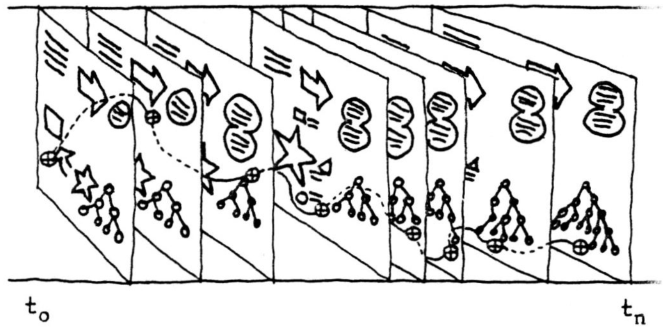



Figure 9. Image dynamics over time facilitate measurement ...


Figure 1Ø. ... taken together with structural dynamics

Taken together, figures 9 and 1Ø provide a high level, 'chunkwise' history list of the user's image transactions. In general, this record becomes a tool for cognitive psychologists to use in exploring the role of graphics and external memory in problem solving. To define that role explicitly in terms of structured image manipulation is to begin to build a system's representation of manipulation for the purpose of graphic communication. Such a representation could later provide a base for intelligent assistance of that activity [Lakin83a].

## CONCLUSION

Text-graphic manipulation provides a rich linguistic phenomenon for research, with measurement of this activity the first step toward understanding it. In particular, when writing and drawing is used by groups during problem solving, the representation of the group's task related cognition is both synoptic and explicit. Group display offers a good opportunity to study intelligence as text-graphic manipulation in constrast to text-only symbol manipulation.

Unlike speech, text-graphic manipulation must be mediated by an artifact, so it might as well be one which records and measures. The handPAM environment provides grouping operators which allow users to structure images for manipulation, and the structures left behind reflect their grappling with the images to manipulate them for cognitive purposes.

The current application of the PAM system is in exploring the role of graphics in the problem solving and communication of the cognitively disabled. As insights are garnered toward this end, they will then be applied to improving the system's representation of the user's objectives, and to supplying assistance based on this knowledge. Assistance will be both of a direct nature (help the user manipulate images) and of a remedial nature (monitor user performance and construct training regimes based on progress). The long term goal of this work is the creation of practical cognitive prosthetics [Lakin83a,83b].

## NOTES ON THE ILLUSTRATIONS

Figure 4 was constructed in the experimental handPAM environment at SAIL. Figures 6 and 8 are screen images from the implementation at Xerox PARC. Figures 3 and 4 are copyrighted by the Association for Computing Machinery, Inc., 198Ø and reprinted by permission (appearing in [Lakin8Øa]). Figures 9 and 1Ø are copyrighted by the IEEE, 1983 and reprinted by permission (appearing in [Lakin83a]).

## REFERENCES

Ball71 Ball, Geoffry H. and Gilkey, James Y., "Facilitation and Explicit Group Memory -- Their Application in Education", SRI International IR&D No. 183531-4Ø9, Menlo Park, CA, Dec 1971.

Brunon71 Brunon, Joseph, "Group Dynamics and Visual Thinking", Journal of Architectural Education, No. 3, 1971.

Engelbart62 Engelbart, D.C., "Augmenting Human Intellect: A Conceptual Framework," SRI International, Menlo Park, California, Oct 1962.

English67 English, W.K., Engelbart, D.C., and Berman, M.L., "Display-Selection Techniques for Text Manipulation", IEEE Trans. on Human Factors in Electronics, Vol. HFE-8, No. 1, March 1967.

Futrelle78 Futrelle, R.P. and Barta, G., "Towards the Design of an Intrinsically Graphical Language", SIGGRAPH '78 Proceedings, pages 28-32, Aug 1978.

Henderson82 Henderson, Austin and Malone, Tom. Personal communication concerning the diagrammatic interview of Austion in June 198Ø which generated figure 6. For more details of this project, contact Austin or Tom.

Ingalls78 Ingalls, Dan, "The Smalltalk-76 Programming System Design and Implementation", Fifth Annual ACM Symposium on Principles of Programming Languages. Tucson, Arizona, Jan 23-25, 1978.

Kay77 Kay, Alan and Goldberg, Adele, "Personal Dynamic Media", IEEE Computer, March 1977.

Lakin75 Lakin, Fred, FACILITATING COGNITIVE GROUP DISPLAY, book manuscript, 1975.

Lakin78 Lakin, Fred, STRUCTURE AND MANIPULATION IN TEXT-GRAPHIC IMAGES, book manuscript, 1978.

Lakin8Øa Lakin, Fred, "A Structure from Manipulation for Text-Graphic Objects", published in the proceedings of SIGGRAPH '8Ø, Seattle, Washington, July, 198Ø.

Lakin8Øb Lakin, Fred, "Diagramming a Project on the Electric Blackboard," video tape for SIGGRAPH '8Ø, July 198Ø.

Lakin8Øc Lakin, Fred, "Computing with Text-Graphic Forms", published in the proceedings of the LISP Conference at Stanford University, August 198Ø.

Lakin81 Lakin, Fred, "Computing for the Right Brain", published in the proceedings of the 1981 West Coast Computer Faire, San Francisco, April 1981.

Lakin83a Lakin, Fred, "A Graphic Communication Environment for the Cognitively Disabled", published in the proceedings of IEEE Spring Compcon '83, San Francisco, March 1983.

Lakin83b Lakin, Fred, "A Graphic Communication Environment For The Brain Injured Based On A Cognitive Model of Image Manipulation", invited address at HEAD TRAUMA REHABILITATION CONFERENCE: COMA TO COMMUNITY, Stanford University Medical Center, March 1983.

Martin81 Martin, Paul A., INTEGRATING LOCAL INFORMATION TO UNDERSTAND DIALOG, PhD thesis, Stanford University, March 1981.

McCarthy6Ø McCarthy, John, "Recursive functions of symbolic expressions and their computation by machine," Comm. ACM, 196Ø, vol. 3, no.4.

McCarthy79 McCarthy, John and Talcott, Carolyn, LISP PROGRAMMING AND PROVING, Class notes CS 2Ø6, Computer Science Dept., Stanford University, 1979.

Sibbet76 Sibbet, David, "Introduction to Group Graphics", The Corospondent, CORO Foundation North. Cal. Public Affairs Quarterly, Summer 1976.

Sibbet8Ø Sibbet, David, I SEE WHAT YOU MEAN -- A WORKSHOP GUIDE TO THE GRAPHIC FACILITATION OF GROUPS, The Offices of David Sibbet, San Francisco, CA.

Teitelman78 Teitelman, Warren, InterLISP REFERENCE MANUAL, Xerox Palo Alto Research Center, 1978.