Toward Empirically-Based Software Visualization Languages

Sarah Douglas, Christopher Hundhausen and Donna McKeown

Department of Computer and Information Science
University of Oregon

Toward Empirically-Based Software Visualization Languages

Sarah Douglas[†], Christopher Hundhausen[†], and Donna McKeown*
[†]Computer & Information Science Dept.
*Psychology Dept.
University of Oregon
Eugene, OR 97403
{douglas, chundhau}@cs.uoregon.edu, donna@dynamic.uoregon.edu

Abstract

Single-user software visualization (SV) systems purport to empower people without expertise in graphics programming to develop their own visualizations interactively, and within minutes. Underlying any single-user SV system is a visualization language onto which its users must map the computations they would like to visualize with the system. We hypothesize that the usability of such systems turns on their ability to provide an underlying visualization language that accords with the ways in which their users conceptualize the computations to be visualized. To explore the question of how to design visualization languages grounded in human conceptualization, we present an empirical study that made use of a research method called visualization storyboarding to investigate the human conceptualization of the bubblesort algorithm. Using an analytical framework based on entities, attributes, and transformations, we derive a semantic-level visualization language for bubblesort, in terms of which all visualizations observed in our study can be expressed. Our empirically-based visualization language provides a framework for predicting the usability of the visualization language defined by Lens [11,12], a prototypical single-user SV system. We draw from a follow-up usability study of Lens to substantiate our predictions.

1 Introduction

Predicated on the intuitive idea that a mapping between an executing program and computer graphics can give one insight into the program's dynamic behavior, computer-based software visualization (SV) systems provide techniques for facilitating four central activities involving that mapping: its (1) design, (2) specification, (3) observation, and (4) interactive exploration. Early SV systems such as Balsa [2] and Tango [17] defined system user models in which different actors performed those activities. In the Balsa model, for example, *client programmers*—algorithmaticians and animators—were responsible for designing and specifying the algorithm-to-graphics mapping, whereas the *end users*—script authors and script viewers—actually observed the mappings (see Figure 1). Although, in theory, the same person could have assumed all four roles, in practice they were played by different people, and at different times.
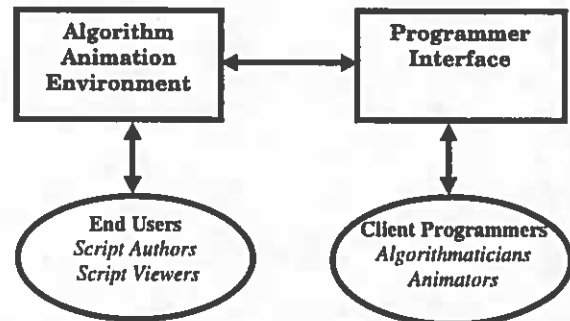

Figure 1. Traditional SV system user model, as defined by the Balsa system

That division of labor can be largely attributed to the sharp difference in the levels of expertise that the Balsa system demanded of client programmers and end users. Because they needed to possess a detailed knowledge of a low-level graphics system, client programmers required a much higher level of expertise than end users, who required no knowledge whatsoever of either the program or the techniques used to animate it, in order to work with the Balsa system. Such a conceptual model, which encourages SV's central activities to be conducted by different people, and at different points in time, has been the dominant SV system paradigm for nearly a decade.

Motivated by a desire to bridge the gap between client programmer and end user—that is, to empower those without low-level graphics programming skills to rapidly prototype their own visualizations—SV researchers have begun to explore ways of realizing an alternative SV model, in which SV's central activities can be undertaken by the same person (see Figure 2). Recently, Mukherjea and Stasko [11,12] developed the first system to adopt such an alternative conceptual model. Their *Lens* system defines an interactive environment into which an algorithm written in C may be loaded. Using a combination

of dialog box fill-in and direct manipulation, Lens users can interactively specify their own mapping between the algorithm (which runs as a separate process under the control of the dbx debugger) and graphics. At any point in the process, users can choose to run the algorithm and observe the visualization they have defined.
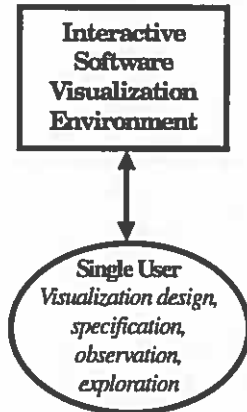

Figure 2. The single user SV system user model

Notice that, from the perspective of the system user, at least three activities are involved in the process of visualization programming under the single-user SV paradigm:

(1) the conceptualization of the computation and its representation, both in terms of pseudocode or a specific programming language, and in terms of a mental visualization;

(2) the mapping of that conceptualization to the *visualization language* defined by the system; and

(3) the manipulation of the SV system's user interface to program the visualization.

Thus, any single-user SV system must define a visualization language onto which computations to be visualized with the system must be mapped.

The central hypothesis of our work is that the usability of computer-based SV systems—and, in particular, those that assume the single user model—turns on their ability to offer a visualization language that accords with the ways in which humans conceptualize the computations to be visualized with the systems. The goal of our research is two-fold: (1) to explore methods for designing SV languages grounded in empirical studies; and (2) to substantiate and refine our hypothesis by integrating empirically-based SV languages into single-user SV systems and conducting further empirical studies.

In this paper, we present our foray into (1). We begin, in Section 2, by describing an empirical study that makes use of a research technique called *visualization storyboarding* to explore human conceptualization of the bubblesort algorithm. In Section 3, we present an analytical framework that allows us to formulate a *semantic-level* visualization language for bubblesort based on our study. In Section 4, we demonstrate the utility of the approach by showing how the language can be used to evaluate the usability of SV systems, using a follow-up study we conducted with the Lens software for supporting evidence. Section 5 surveys related work. We conclude, in Section 6, by identifying several directions for future research.

2 An empirical study of the human conceptualization of bubblesort

In this section, we introduce a research method called visualization storyboarding, and we describe an exploratory study that made use of the method to investigate the ways in which humans visualized the bubblesort algorithm. See [5] for a more thorough treatment of the study.

2.1 Method: Visualization storyboarding

How does one empirically study human conceptualizations of computer programs independently of computer-based technology? We see at least two requirements for the medium employed by the empirical method. First, the medium must be *dynamic*; that is, it must be flexible enough to express the constantly-changing form that human visualizations are likely to take on. Second, since we are interested in allowing humans to express their conceptualizations in the most natural way possible, the medium should be one with which humans are familiar and comfortable; it should support, and not be an impediment to, the expression of their ideas.

Based on the storyboarding technique introduced by the animation industry in the 1930s, *visualization storyboarding* takes advantage of the familiarity and dynamism of art supplies and human conversation to create a situation in which humans can describe computer programs naturally. Two person teams are provided with a full spectrum of colored construction paper, scissors, and different colored pens; they are then asked to work together to describe a computer program to someone who has never seen it. The resulting *storyboard* of the algorithm's execution—a dynamic presentation consisting of animated construction paper cut-outs, drawings, conversation, and gestures—is videotaped, becoming the empirical data to be analyzed at a later time.

Visualization storyboarding applies two qualitative research techniques—*constructive interaction* [10] and *conversational analysis* [18]—to the situation of software

visualization design. By using two participants, constructive interaction creates a situation of collaborative problem solving in which each participant must inform the other, in an explicit verbal and visual record, of problems, causes, and solutions each has encountered. To interpret participants' interaction, we review the videotapes of participant sessions using conversational analysis techniques developed by Douglas [4].

2.2 Procedure

We videotaped three same-sex pairs of participants (two consisting of women, one consisting of men) during 45 to 70 minute sessions. All participants were graduate students in computer science at the University of Oregon, and all reported that they understood the bubblesort algorithm. Nonetheless, we provided participants with C code for their reference (see Figure 3).

```
#include <stdio.h>
main()
{
  int n,i,j;
  int temp;
  int a[50];

  int count;

  printf("Input number of elts in array\n");
  scanf("%d",&n);

  printf("Enter the elements\n");
  for (count=0; count<n; ++count)
    scanf("%d",&a[count]);

  for (j=n-2; j>=0; --j)
    for (i=0; i<=j; ++i)
    if (a[i] > a[i+1])
      { temp = a[i];
        a[i] = a[i+1];
        a[i+1] = temp;
      }
}
```

Figure 3. The C *bubblesort* algorithm given to participants

After providing our pairs of participants with art supplies, we presented them the following task: "If you were to explain the bubble sort algorithm to someone who had never seen it, how would you do it? Feel free to use any of these resources (construction paper, scissors, etc.) to assist you in your explanation." If they decided to design a visualization for a sample data set, we recommended that they only step through a few of the iterations for simplicity's sake.

2.3 Observations

All pairs of participants made extensive and intriguing use of the art supplies with which we supplied them. In fact, they all used the art supplies materials to create homemade animations of bubblesort operating on a sam-

ple data set; they augmented their construction paper animations extensively with their own gestures and comments. Figures 4, 5, and 6 depict our participants' storyboards. Note that, because they are static in nature, these figures cannot do full justice to our participants' dynamic presentations, which must be seen on videotape to be appreciated.
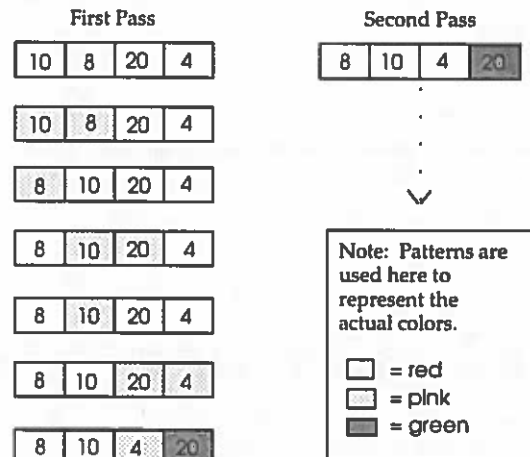

Figure 4. The visualization trace created by the Number Pair to demonstrate bubble sort

One pair, whom we shall call the Number Pair, used numbers (e.g. 5, 4, 3) to indicate the magnitude of each array element, and used color to illustrate significant state changes (see Figure 4). They initially shaded all elements in the array red to indicate that the elements were unsorted. Each pass of the algorithm's outer loop was depicted as a column of rows. Successive horizontal rows of array elements within a column corresponded to successive passes of the algorithm's inner loop. Within each horizontal row, the two array elements compared during that pass were shaded pink to indicate that they had been compared. At the end of a pass of the algorithm's outer loop, the pair colored the element that had reached its rightful place in the array green, and began a new column for the next pass of the array's outer loop.

Another pair, whom we shall call the Color Pair, used color to indicate magnitude. The legend they constructed (see the triangles at the top of Figure 5) defined a canonical (spectral) order for a five-element array. By illustrating the array after each successive pass of the array's outer loop, they demonstrated how the bubble sort gradually placed an unsorted array into that order. Like the first pair, this pair had built, by the end of the sort, a history "array" of the sort; however, they depicted each pass of the algorithm's outer loop with row instead of a column.

The third pair, whom we shall call the Football Pair, also used color to denote magnitude (see Figure 6).
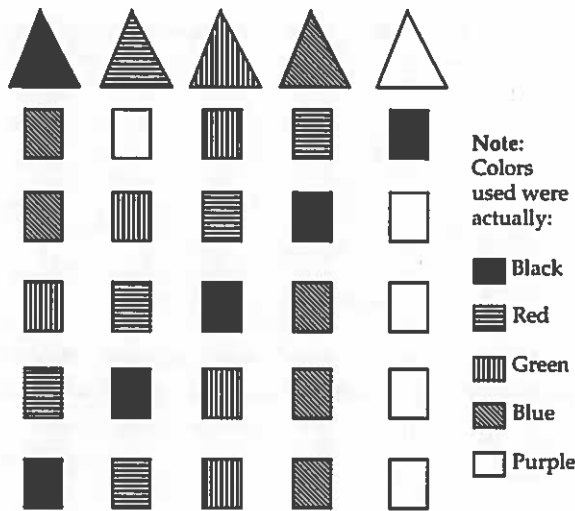
3

Figure 5. The visualization trace created by the Color Pair to demonstrate bubble sort

However, color played only a secondary role in their visualization, which drew extensively from concepts of American football to depict various aspects of the algorithm. Sorting elements were represented by football players whose varying weights were represented by color, as indicated by the legend. At the beginning of each pass of the sort, the football was given to the first player in line.
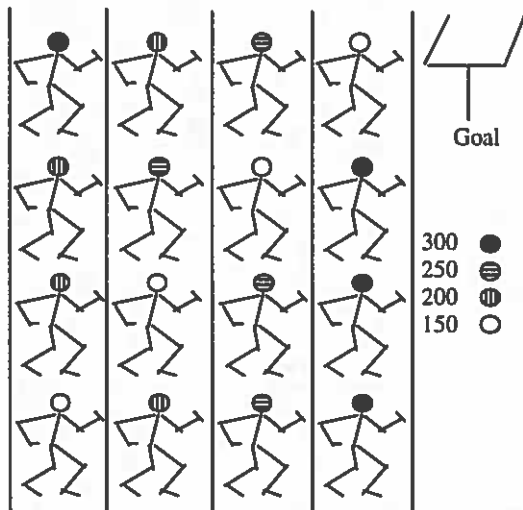

Figure 6. The visualization trace created by the Football Pair to demonstrate bubblesort

The ball carrier and the player next in line symbolized the two elements being compared at any given time. If the current ball carrier weighed more than the next in line, then the ball carrier knocked that player over, thereby exchanging places with that player in the array. If, however, the ball carrier weighed less than the next player in line, the ball carrier fumbled the ball to the next

player, who then became the ball carrier. This process of ball advancement continued until the ball carrier reached the end of the line, whereupon a new pass of the sort commenced.

3 Formulation of a semantic-level SV language for bubblesort

In this section, we motivate an analytical framework that provides a means for characterizing visualization semantics. We then use the framework to derive a semantic-level SV language, in terms of which all of the bubblesort visualizations observed in our study can be expressed.

3.1 Analytical framework

In deciphering the languages employed by our participants in their visualization storyboards, we find it useful to distinguish between two linguistic levels: *lexical* and *semantic*. We use the *lexical level* to refer to the graphical vocabularies of their visualizations. In contrast, we use the *semantic level* to refer to the meanings of elements in those graphical vocabularies.

At a lexical level, our participants' visualization languages varied on three main fronts. First, they used varying *graphical entities*. For example, the Number Pair used numbers, the Color Pair used triangles and squares, and the Football Pair used football players. Second, the visualizations conveyed various kinds of information by using different *attributes* of those entities. For example, the Color Pair used *color* to represent sorting element magnitude; the Football Pair used the *position* of a football to denote the two elements currently being compared. Finally, the visualizations employed different *transformations* among entities and attributes to convey various kinds of information. For example the Number Pair altered number square color to indicate comparison, whereas the Football pair changed the location of the football to indicate comparison.

Despite the lexical differences among our participants' visualization languages, notice that they all encoded the bubblesort at a similar level of abstraction—one that conveys bubblesort's overall functionality in terms of a similar kernel of fundamental abstractions, including *sorting element, array of sorting elements, comparison* and *exchange*.

Given the three dimensions along which our participants' visualization languages differed lexically, and given that all of their visualizations portrayed bubblesort's abstract functionality using a similar semantics, we propose a framework for analyzing their visualizations that characterizes them in terms of the (lexical level)

4

Semantics	Number Pair Lexicon	Color Pair Lexicon	Football Pair Lexicon
Sorting element	Square	Square	Stick Figure
Magnitude of element	Number symbol	Color	Color (as weight)
Array of elements	Contiguous row of squares	Non-contiguous row of squares	Contiguous row of figures
Inner loop pass history	Rows of sorting elements	—	—
Outer loop pass history	Columns of rows	Rows of sorting elements	Rows of sorting elements
Legend explicating ordering on sort elements	—	Triangles with color spectrum	Column of color/player weight pairs

Table 1. Mappings between lexical entities and attributes of the human visualizations and their semantics

Semantics	Number Pair Lexicon	Color Pair Lexicon	Football Pair Lexicon
DO outer loop	Start new column of rows	Create new row of squares	Create new row of football players
DO inner loop	Create new row of squares	—	—
a) Reference elements to be compared	Color elements pink	—	Location of football
b) *Compare* elements (same, <, >)	—	—	Intuitions about how player size relates to running, tackling, and fumbling
c) *Exchange* elements	Exchange numbers	Exchange colors	Ball carrier advances by tackling next football player in line (thereby exchanging positions with that player)
d) Don't *exchange* elements	—	—	Fumble football to next player in line
Terminate outer loop	Color square in correct order green	—	—
Terminate Sorting	Ordering of natural numbers, all squares green	Color squares match legend	Players ordered by weight

Table 2. Mappings between lexical transformations of the human visualizations and their semantics

entities, attributes, and transformations by which they represented bubblesort's (semantic level) abstract functionality. As we shall see below, by using such a framework, we uncover a strikingly similar semantics for the visualization languages, allowing us to unify the languages at a semantic level.

3.2 Deriving a semantic-level SV language for bubblesort

Table 1 summarizes the mappings between the lexical entities and attributes created by each pair, and their underlying semantics. For example, the Number Pair used numbers in a row of contiguous squares to designate an array. Notice that some semantic-level entities present in a typical semantic-level pseudocode description of bubblesort, such as temporary storage and subscripts, were not represented in our participants' visualizations. Participants perhaps believed that such entities would serve to confound the abstract functionality that they were

striving to capture. Conversely, the color legend—used by both the Color Pair and the Football Pair—cannot be derived from the algorithm's semantics. Instead, it is an artifact of the visualization itself, most likely arising out of a perceived need to explicate the ordering relationship among graphical entities.

Table 2 summarizes the mapping between transformations created by each pair, and their underlying semantics. For instance, the Number Pair used color to indicate that two elements were being compared (the elements turned pink), and to indicate a difference between sorted elements (which turned green) and unsorted elements (red).

Derived directly from the analysis presented in Tables 1 and 2, the SV language presented in Table 3 unifies our participants' SV languages at a *semantic* level. While we might have chosen to express that language using any number of different formalisms, our analytical framework lends itself naturally to expression in terms of abstract data types (ADTs), which capture the semantics

ADT Name	Operations	Semantics
Sorting_element	*instantiate(magnitude)*	Create graphical appearance for element based on magnitude
	compare_highlight()	Graphically indicate a comparison involving this element
	inorder_highlight()	Graphically indicate that this element is now in its rightful place in the sort array
	move(new_index)	Move this element to location defined by *new_index*
Sort_array	*instantiate (array of integers)*	Create graphical appearance for sort array (facilitates data initialization)
	start_outer_loop()	Graphically indicate the beginning of a pass of the algorithm's outer loop
	start_inner_loop()	Graphically indicate the beginning of a pass of the algorithm's inner loop
	compare(elt1, elt2)	Graphically indicate a comparison between *elt1* and *elt2* by calling on *compare_highlight()* operations of *elt1* and *elt2*
	exchange(elt1, elt2)	Graphically indicate an exchange involving *elt1* and *elt2* by calling on the *move(new_index)* operations of *elt1* and *elt2*
	no_exchange(elt1, elt2)	Graphically indicate that there was no exchange involving *elt1* and *elt2*
	terminate_outer_loop()	Graphically indicate the end of a pass of the outer loop, possibly calling upon *inorder_highlight*
Legend	*instantiate(array of integers)*	Graphically explicate ordering on sort elements

Table 3. Semantic-level SV language for bubblesort in terms of ADTs

underlying our participants' visualization languages, but which intentionally hide the lexical details of their graphical representation. Indeed, as we have seen, such lexical details vary considerably depending on the people doing the visualization.

In Figure 7, we use our ADT language in conjunction with a procedural Pascal-like language to express all of our participants' visualizations as a common program. Having provided, within our study, the lexical details of how each operation manifests itself graphically, our participants can thus be viewed as the *implementors* of the ADTs used in that program.

4 Using semantic-level analysis to evaluate existing single-user SV software

In this section, we demonstrate how the semantic-level analysis introduced in the previous section can assist in the evaluation of computer-based SV systems. Section 4.1 uses our framework to formulate hypotheses regarding the usability of the bubblesort SV language provided by Lens, a single-user SV system developed by Mukherjea and Stasko [11,12]. In Section 4.2, we scrutinize our hypotheses by drawing from an actual usability study we conducted on Lens.

4.1 Analyzing the Lens SV language for bubblesort

Figure 8 presents the final frame of the bubblesort visualization that one would typically define in Lens. Array elements are depicted as sticks whose height corresponds to magnitude. Two elements are flashed to indicate a comparison, and smoothly change places to indicate an exchange. An ascending row of sticks indicates that the array has been sorted.

In Tables 4 and 5, we illustrate the manner in which the semantic-level entities, attributes, and transformations identified in our empirical study map to the Lens lexicon. As those tables indicate, the visualization language defined by Lens accords quite well, at a semantic level, with the visualization languages defined by our participants; indeed, it provides support for depicting the bubblesort in terms of a subset of the semantic-level abstractions present in our semantic-level SV language.

Notice, however, that the Lens semantics does not fully cover the semantics of the visualizations we encountered in our study. In particular, we can identify two shortcomings. First, the Lens language supports no means for maintaining a history of bubblesort's inner and outer loops; instead, the array of sticks is initialized at the beginning of the sort, and all passes of the sort are

6

```
VAR
  a   : ARRAY OF INTEGER;
  le  : legend;
  sa  : sort_array;
i, j, temp: INTEGER;

BEGIN (* Bubblesort *)
  read data into a;
  sa.instantiate(a);
  le.instantiate(a);
  FOR j := n - 2 TO 0 DO BEGIN
      sa.start_outer_loop();
    FOR i := 0 TO J DO BEGIN
      sa.start_inner_loop();
      sa.compare(i, i+1);
      IF a[i] > a[i+1] THEN BEGIN
        sa.exchange(i, i+1);
        temp := a[i];
        a[i] := a[i+1];
        a[i+1] := temp
      END (* IF *)
      ELSE
        sa.no_exchange(i, i+1);
    END (*inner FOR *)
  END (* outer FOR *)
END; (* Bubblesort *)
```

Figure 7. Expressing participants' visualizations by annotating a procedural language with calls to the ADTs defined in Table 3

performed on that same set of sticks. Here, it is important to point out that the Lens interface does provide the option of starting and stopping the visualization at any point; thus, intermediate snapshots of the array can be momentarily viewed, although they cannot be preserved for later inspection. Second, the Lens language provides no means for defining a legend for explicating the ordering on sorting elements.

Semantics	Lens Lexicon
Sorting element	Stick
Magnitude of element	Stick height or width
Array of elements	Non-contiguous row of sticks
Inner loop pass history	—
Outer loop pass history	—
Legend explicating ordering on sort elements	—

Table 4. Mappings between semantic-level entities and attributes identified in our empirical study and the Lens lexicon

Semantics	Lens Lexicon
Initialize array magnitudes (Done once per loop in participant visualizations)	Create row of black sticks
DO outer loop	—
DO inner loop	—
a) Reference elements to be compared	Flash elements, or change fill color
b) *Compare* elements (same, <, >)	—
c) *Exchange* elements	Exchange sticks
d) Don't *exchange* elements	—
Terminate outer loop	—
Terminate Sorting	Sticks ordered by increasing height or width

Table 5. Mappings from semantic-level pseudocode transformations to Lens visualization

Based our semantic-level analysis of the Lens SV language, we can make two hypotheses regarding its usability:

Hypothesis 1: Lens users will not have problems with the semantics behind the lexical elements in terms of which the abstract functionality of bubblesort can be depicted in Lens. Indeed, all of the semantic entities available in the Lens language are also in our semantic-level SV language.

Hypothesis 2: Lens users will encounter two kinds of usability problems with the Lens language stemming from its failure to support two kinds of semantics present in our semantic-level SV language: *array history* and *legends*. Since the Lens language does not support the former, we might expect Lens users to complain about an inability to distinguish among intermediate passes of the algorithm's loops. Since the Lens language does not support the latter, we might expect Lens users to have problems interpreting the target ordering of the array.

In the following section, we scrutinize these hypotheses by considering a usability study we conducted on Lens.

4.2 Using a follow-up usability study on Lens to substantiate the hypotheses

Overview of study. Two pairs of students who participated in the study presented in Section 2 went on to program a visualization for bubblesort using Lens. Participants were given a three page description of proce-

dural instructions for implementing a predefined visualization of bubblesort corresponding to the one described in Section 4.1 (see Figure 8); we then asked them to use the Lens software to program the visualization. As was the case with our visualization storyboarding experiments, both sessions were videotaped and later analyzed using conversational analysis techniques developed by Douglas [4]. See [8] for a more thorough treatment of the study.
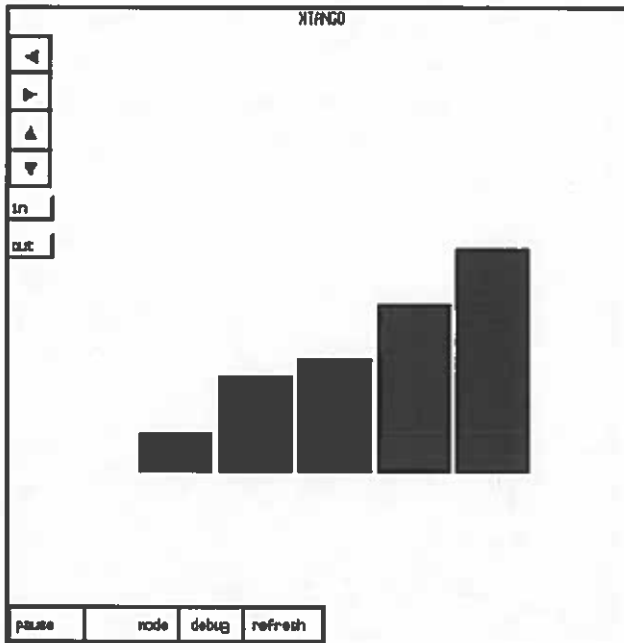

Figure 8. The final frame of the standard bubblesort visualization in Lens

Observations relevant to Hypothesis 1. As hypothesized, we observed that the overall usability of the Lens language was quite good. By the time they had watched the entire Lens visualization, both pairs of participants clearly indicated that they grasped the mappings between elements of the Lens lexicon and their underlying semantics. Since there existed a good match between the underlying semantics of the Lens visualization, and the semantics of the visualizations defined by our two pairs of participants, the fact that the Lens visualization used different lexical elements did not hinder our participants' comprehension.

Observations relevant to Hypothesis 2. Participants did not explicitly complain about the fact that the Lens visualization did not represent sort history. Thus, we have no reason to believe that the failure of the Lens language to provide a direct means for representing sort history was the cause of a usability problem.

We did, however, document convincing evidence that the lack of a legend in the Lens visualization caused a usabilty problem. Indeed, both pairs of study participants

were initially confused about the mapping between element magnitude and stick height in the Lens visualization. As the visualization progressed, they eventually came to understand that mapping. Nonetheless, our semantic-level analysis suggests that, if the Lens language had supported the definition of an order legend, participants might have grasped the visualization more quickly.

5 Related work

The research described in this paper purports to use empirical studies as a basis for SV language design and evaluation. Two lines of related work have employed empirical studies with humans within the context of software visualization, while a third line of related work has specifically addressed the problem of SV language design. In this section, we briefly survey each of these.

5.1 Using empirical studies to evaluate SV systems

Pioneered at the Georgia Institute of Technology, the first line of related work has focused on the problem of empirically verifying the benefits of computer-based SV systems in pedagogical settings [1,9,15]. Our work differs fundamentally from this work both in the research method (quantitative factors analysis versus qualitative visualization storyboarding), and in the research goal (evaluation of SV efficacy versus SV language design).

5.2 Using empirical studies to document human visualizations

Ford [6] employs a research method similar to ours to document human visualizations of computer programs. In Ford's study, 46 beginning computer science students, assembled into two-, three-, and four-person teams, were videotaped as they sketched visualizations for C++ code fragments. Ford's analysis of the 180 visualizations resulting from the study revealed a wide range of alternative visualizations for the same code fragments; the visualizations were classified using Cox and Roman's [3] abstraction classification system.

In contrast to our study, Ford's study considered the visualization of general imperative and object-oriented programming concepts—for example, variables, pointers, loops, conditionals, arrays, and classes; visualization conceptualization was not considered at the level of algorithm semantics. Further, Ford made no attempt to use the study results as a basis for an SV language. Instead, Ford's interest has been in devising a set of empirically-based visual program abstractions for aiding

programmer comprehension of conventional text-based program views [7].

5.3 Language design

Drawing from an analysis of human pen-and-paper descriptions, Radiya and Radiya [13] introduce a model called HAL (Human approach to describing ALgorithms) for graphically describing algorithms However, whereas they were interested in using such descriptions as basis for a visual programming language, we are interested in applying our research to SV languages.

To design the Lens SV language, Mukherjea and Stasko [11,12] studied visualizations of 42 algorithms drawn from several problem domains, including sorting, searching, graph theory, and graphics. Programmed by over 25 people from 4 different institutions, all of the visualizations they studied had been implemented in XTango [16], an animation toolkit based on Stasko's path-transition paradigm [17]. It should be no surprise, then, that the language they designed is a subset of the XTango language. Mukherjea's and Stasko's approach differs from ours in that it used visualizations already programmed in a particular SV language as a basis for design; we are interested in grounding SV languages in visualizations described independently of computer-based technology and languages.

6 Conclusions and future work

In this paper, we have shown how visualization storyboarding studies, together with a semantic-level analytical framework, can be used to derive an empirically-based, semantic-level SV language for bubblesort. As we have demonstrated, such a language can prove useful in diagnosing language-based usability problems in existing single-user SV software.

However, we suspect that our approach holds even greater promise as a tool for computer-based SV system design. We envision that semantic-level languages like the one presented here could serve as a good starting point for defining both the functionality and user interface of highly-usable single-user SV systems. Future research will need to confirm the approach's potential as a design tool by actually using SV languages derived using the approach as the basis for single-user, computer-based SV software. To that end, we plan to complete usability studies on a computer-based system that implements the bubblesort SV language presented in this paper by the end of the year.

Aside from confirming the approach's potential as a design tool, we are interested in exploring several related research questions that have been inspired by the research presented in this paper. We conclude by identifying five of these.

Studying a wider range of computations. How universal are the semantic primitives that we have identified for the bubblesort visualization language? Do they exist in the visualization languages of other sorting algorithms? We suspect that visualization languages are highly computation-dependent. We would like to conduct further empirical studies that focus on a wider range of algorithms in order to explore both the range of visualization languages that are possible, and the extent to which visualization languages can be generalized to specific classes of computations.

Studying more complex computations. What affect does the sophistication of a computation have on the semantic-level visualization language in which it is visualized? We initially chose to study the bubblesort algorithm precisely because of its simplicity. Perhaps because of that simplicity, we found that people visualize it at a similar level of abstraction. We suspect that the more sophisticated the algorithm, the wider the range of semantic levels at which humans will choose to visualize it. What analysis and design techniques could be used to deal with visualization languages with varying semantic levels?

Studying the relationship between SV language expressibility and SV language usability. The work presented here considers the impact of SV language design on usability only in the small case of bubblesort. In practice, SV system designers are interested in supporting the definition of visualizations for a broad range of algorithm classes. For example, the Lens system has been used to define visualizations for sorting, searching, graph, computational geometry, and even scientific simulation algorithms [12]. How does the expressibility of an SV language—that is, the range of computations whose visualizations it can express—affect its usability? Will a library of different SV languages—each suited to specifying visualizations for a particular computation or class of computations—prove more usable than a monolithic SV language capable of expressing visualizations for any possible computation?

Studying alternative SV language paradigms. All of the participants in our visualization storyboarding experiment chose to visualize the bubblesort algorithm by identifying interesting events in the algorithm, and mapping those interesting events to corresponding graphical transformations. Called *visualization-by-annotation* in the SV literature [14], this specification paradigm is the one most commonly supported by computer-based SV systems.

It is widely known, however, that visualization-by-annotation is not the most appropriate method for ex-

pressing all possible computation-to-visualization mappings. Can the empirical study and analysis technique introduced here inform design in cases in which the computation-to-visualization mapping employs an alternative SV mapping—for example, the declarative method by which the traditional visualization for the Towers of Hanoi algorithm is most elegantly expressed? If not, what can we learn from visualization storyboarding studies in such cases, and are there alternative analysis techniques that can help us?

Incorporating gestures and speech into visualization languages. Our analysis of the empirical study presented in Section 2 only considered our participants' storyboards—the graphical traces of the bubblesort that they created. Yet, as we stated in that section, participants augmented their storyboards extensively with gestures and speech, both of which contributed greatly to their overall presentations.

Unfortunately, the videotapes of our participants' sessions, which were filmed using a single camera positioned above the participants, proved inadequate as a means for coordinating participant gestures and speech with transformations in their storyboards. By using a second camera focused on the participants themselves, however, we could capture their gestures and speech, and consider their semantics in our analysis. We look forward to exploring that possibility in future studies.

Acknowledgment

References

[1] A. Badre, M. Beranek, J.M. Morris, and J.T. Stasko. Assessing program visualization systems as instructional aids. Technical report GIT-GVU-91-23, College of Computing, Georgia Institute of Technology, Atlanta, GA, 1991.

[2] Brown, M. *Algorithm Animation*. Cambridge, MA: The MIT Press, 1987.

[3] K.C. Cox and G-C. Roman. Abstraction in algorithm animation. In *Proc. 1992 IEEE Workshop on Visual Languages* (Seattle, WA), pp. 18–23, 1992.

[4] S.A. Douglas. Conversational analysis and human-computer interaction design. In P. Thomas (ed.) *The Social and Interactional Dimensions of Human-Computer Interfaces*. Cambridge: Cambridge University Press, (in press).

[5] S. Douglas, D. McKeown, and C. Hundhausen. Exploring human visualization of algorithms. Technical report CIS-TR-94-27, Department of Computer and Information Science, University of Oregon, Eugene, 1993.

[6] L. Ford. How programmers visualize programs. Research report 271, Department of Computer Science, University of Exeter, Exeter, U.K., 1993.

[7] L. Ford and D. Tallis. Interacting visual abstractions of programs. In *Proc. 1993 IEEE Workshop on Visual Languages* (Bergen, Norway), pp. 93–97, 1993.

[8] C.D. Hundhausen. Exploring the potential for conversational analysis in the evaluation of interactive algorithm visualization systems. Unpublished technical report available at http://www.cs.uoregon.edu/~chundhau, 1993.

[9] A.W. Lawrence, A.N. Badre, and J.T. Stasko. Empirically evaluating the use of animations to teach algorithms. Technical report GIT-GVU-94-07, College of Computing, Georgia Institute of Technology, Atlanta, GA, 1994.

[10] N. Miyake. Constructive interaction and the iterative process of understanding. *Cognitive Science 10*, pp. 151–177, 1986.

[11] S. Mukherjea and J.T. Stasko Applying algorithm animation techniques for program tracing, debugging, and understanding. In *Proc. 15th IEEE International Conference on Software Engineering* (Baltimore, MD), pp. 456–465, 1993.

[12] S. Mukherjea and J.T. Stasko. Toward visual debugging: Integrating algorithm animation capabilities within a a source-level debugger. *ACM Trans. on Computer-Human Interaction 1(3)*, pp. 215–244, 1994.

[13] A. Radiya and V. Radiya. A model of human approach to describing algorithms using diagrams. In *Proc. 1992 IEEE Workshop on Visual Languages* (Seattle, WA), pp. 261–263, 1992.

[14] G-C. Roman and K.C. Cox. A taxonomy of program visualization systems. *IEEE Computer 26(12)* (December), pp. 11–24, 1993.

[15] J.T. Stasko, A. Badre, and C. Lewis. Do algorithm animations assist learning? An empirical study and analysis. In *Proc. INTERCHI '93 Conference on Human Factors in Computing Systems* (Amsterdam, The Netherlands), pp. 61–66, 1993.

[16] J.T. Stasko. Animating algorithms with XTANGO. *SICACT News 23(2)*, pp. 67–71, 1992.

[17] J.T. Stasko. Tango: A framework and system for algorithm animation. *IEEE Computer 23* (September) pp. 27–39, 1990.

[18] L. Suchman. *Plans and situated actions: The problem of human–machine communication*. Cambridge: Cambridge University Press, 1987.