A Framework for Constructing Visualization, Animation, and Direct Manipulation Interfaces

視覚化、アニメーション、直接操作インターフェース作成のための枠組

by
Shin Takahashi
高橋 伸

A Dissertation

Abstract

This thesis describes a framework for developing kinds of non-WIMP-based user interface software, designated as a bi-directional translation model[1]. Here, '*non-WIMP-based*' interface means that it cannot built only by combining WIMP widgets, such as buttons and menus. In particular, the targets of this thesis are two types of GUI software. One is direct manipulation interfaces for figures and diagrams that visually represent various abstract objects and relations in an application data, such as direct manipulation interfaces for network and hierarchical data structures. They enable the user to draw and modify diagrams to input abstract data into an application. Another is animations of changing abstract objects and relations in an application data, such as animations of sorting algorithms. They show their transitions visually, and are useful for understanding the behavior of an application or an algorithm.

The bi-directional translation model models the general process of visualization, picture recognition, and animation generation. It is originated in Kamada's general framework for visualizing abstract objects and relations. We have extended it by integrating recognition of figures/diagrams and animation of changing abstract objects and relations, which is one of the contributions in this thesis. The key idea of the framework is that these functions are *translations* between different data representations. The model introduced four data representations: AR (Application Data Representation), ASR (Abstract Structure Representation), VSR (Visual Structure Representation), and PR (Pictorial Representation). In the model, visualization is a translation from AR to PR via ASR and VSR. Recognition of figures is an inverse translation from PR to AR via VSR and ASR. The essential translation is that between ASR and VSR, because there is no a priori mapping between these representations. ASR is the representation in our model that represents the structure of application data, and it does not have explicit visual appearance information. On the other hand, the purpose of VSR is to represent the high-level structures of pictures. The mapping between ASR and VSR is the essential conversion between abstract data for an application and pictorial data for presentation. Nevertheless, the structure of ASR and that of VSR are usually similar. As users understand the meaning of abstract application data via pictures, visualized pictures should *represent* abstract application data, and thus they should have a similar structure. The programmer specifies these translations by visual mapping rule sets, which are declarative rules that define mappings between ASR data and VSR data. As both structures are similar, mapping rules are usually simple.

Animations are also handled in the framework extended naturally to the time dimension. In the framework, animations are regarded as *operations* on PR that are translated from operations on AR via operations on ASR and VSR. The translation among operations is executed maintaining consistency with the mapping relations among the AR, ASR, VSR, and PR data. We have chosen an interpolation-based method for implementation of the extended framework for animations. That is, animations are generated by interpolating a sequence of pictures translated from the running application's internal data. Rather than specifying animations (motions or transformations) directly, the

---

[1]WIMP stands for Windows, Icons, Menus and Pointing device.

programmer specifies transitional operations, i.e., the methods used for interpolating two pictures before and after the invocation of an operation, with transition mapping rules. Therefore, animations are determined by two types of mapping rules: visual mapping rules and transition mapping rules.

In addition, we describe the systems implemented based on the framework. TRIP2 is a system that achieved the bi-directional translation between AR and PR. We applied TRIP2 to produce several examples such as the Othello application. It is implemented on NextStep using Objective-C and Prolog. To build direct manipulation interfaces using TRIP2, the programmer writes visual and inverse visual mapping rules in Prolog. TRIP2a is a tool for constructing abstract animations that depict the behavior of program executions. It is implemented by extending TRIP2 so that we can use two functions together: bi-directional translation between abstract data and pictures, and translations from abstract data into animations. The programmer can use the same visual mapping rules of TRIP2 for TRIP2a. TRIP2a/3D is the successor of the TRIP2a system that specializes in generating animations and can also handle three-dimensional representations and event-driven animations, which is useful for animating parallel program executions. In order to be able to view animations on various platforms, the animation viewers are separated from the translation module in TRIP2a/3D. The programmer writes visual mapping rules for each animation in KLIC which are compiled with translation modules to generate translators that output animation data from the log data of the application's execution. Generated animation data can be viewed with viewers on various platforms such as X-Window, MS-Windows, and Java. We describe various algorithm animations and visualizations of program executions generated with these systems.

Supporting the debugging of visual mapping rules with our systems is difficult. As a step toward solving this problem, we describe techniques for visualizing constraint systems in visual mapping rule sets. One way is to show constraint systems as three-dimensional graphs. By looking at constraint graphs, the user can see the structure of constraint systems more directly than in their textual form. Another way is to animate constrained graphical objects to show their degrees of freedom in constraint systems by using cartoon techniques. These techniques are useful for understanding and debugging constraint systems in visual mapping rules. Using these techniques, we developed a browser for constraints in VSR of TRIP systems. It is implemented using Java and integrated with the animation viewer of TRIP2a/3D.

Acknowledgments

5

Contents

List of Figures

11

Chapter 1

INTRODUCTION

1.1 Background

There is a common saying that "a picture is worth a thousand words." Human beings have displayed various types of information visually since the Paleolithic age. Cave paintings at Lascaux and Altamira indicate that humans have always drawn pictures, which serve as messages for communication. During the medieval period, pictures were used to represent divine concepts, and were not merely copies of real scenes. People often draw a rough sketch to represent their thinking, and these sketches represent visual communications to themselves.

The importance of visual display is increasing in the computer age. Powerful computers and high-speed networks are producing vast amounts of diverse information that is difficult and time-consuming to retrieve and understand. Visualization is one technique that can be used to alleviate this problem. With well-drawn graphs, such as scatter plots and line charts, users can more easily understand statistical information. The results of large scientific computations cannot be fully understood without accompanying sophisticated visualization. UML[1] diagrams are widely employed to visualize complex structure of large software. Visualization is, therefore, an indispensable technique for modern computer systems.

Pictures are helpful not only for presenting information to the user, but are also useful for the user to provide input to computers. Graphical user interfaces (GUIs) enable users to input information visually. Computers show visual representations of various data to the user, who edits the visual representations to change the corresponding data. Users actually feel as if visual representations are the data, and think that they are handling the data directly. This is an instance of technique called *direct manipulation*(DM)[109]. Shneiderman pointed out four principles of DM interfaces in [109]:

- Continuous representation of the object of interest.

- Physical actions or labeled button presses instead of complex syntax.

- Rapid, incremental, reversible operations whose impact on the object of interest is immediately visible.

- Layered or spiral approach to learning that permits usage with minimal knowledge.

The concept of direct manipulation is widely incorporated into current GUI software. In particular, it is essential for the applications that handle visual information such as figures, diagrams, and pictures.

---

[1] Unified Modeling Language. For details, see `http://www.uml.org/`

One of the goals of user interface research is to allow users to communicate with computers in diverse ways, including DM-style interfaces. Which method is better is determined by its purpose. The user interface should be chosen so that it is appropriate for the information that it handles. Most traditional communication between users and computers has been performed through character-based methods. The user inputs commands with a keyboard, and the computer presents its output in the form of text. This style of user interface is still dominant in most current UNIX systems. Expert users love such systems, and they are indeed useful for textual tasks, such as writing documents and developing programs. However, such command-line interfaces are not suitable for the communication of visual information. GUIs enhance visual interactions between users and computers. Computers show information visually, and users see and react to these visual representations with a keyboard and a pointing device. However, current GUIs are insufficient in some important aspects. Although representation of data is visual with a GUI, users cannot freely input visual information. Most current GUI systems are WIMP-based; that is, they consist of **W**indow, **I**con, **M**enu, and **P**ointing device. In a WIMP-based GUI, basic user inputs are selection and clicking of menus, icons, and buttons using a pointing device and input from a keyboard. They are limited in the sense that users do not easily input visual data such as pictures, figures, and diagrams.

There are some *visual* applications. For example, drawing and painting software enable users to create figures and pictures with a pointing device. Designers use CAD systems to design architecture, industrial products, electronic circuits, etc. In visual programming environments (VPEs), programmers can construct programs visually and interactively. However, these systems have not yet reached our goal of allowing users to communicate with computers in diverse ways. Drawing editors cannot understand the pictures that users draw; that is, drawn pictures are treated as images from which systems cannot extract meaning. CAD systems and VPEs recognize drawn figures, which are used for input from users. In this sense, these systems are communicating with users via visual representations of data. However, the syntax of the pictures that they can understand is rigid and built in; it is difficult to provide software that can recognize and understand figures with user-defined syntax. More flexible visual communication between computers and humans, and better communication among humans with the assistance of computers, are desirable goals for computer-human interaction. In addition, computers would be able to assist drawing work intelligently if they could understand what the user is drawing.

Various tools have been developed to ease the burden of visualizing information. Scientific visualization tools such as AVS[30] and Khoros[4] are widely used by physicists for visualization of large numeric data, such as the results of fluid mechanics' simulations. Spreadsheet applications that can produce business graphics easily from a table are another widespread example. Mathematica[129] and MATLAB[84, 35] are powerful environment that helps scientific calculation and visualization of the result. However, they basically provide ready-made visualization libraries for numeric data. Although it may be possible to customize them by changing various parameters, it is difficult to implement user's own visualization methods. In addition, they are not designed to handle relational and structural data.

The visualization of dynamic information is referred to as "animation." In particular, *algorithm animation*, animations that depict the behavior of algorithms, is one highly active area of user interface research. *Sorting Out Sorting*[6] is a well-known example that shows animations of various sorting algorithms. A number of systems to support the development of algorithm animations have been proposed, such as BALSA[15], Zeus[16], Tango[116], and Pavane[103]. However, these are tools for making devised animations. It is still a difficult and tedious task to make animations for personal purposes, such as verifying or debugging the behavior of programs.

1.2 Goals and Approach

To approach this ultimate goal, support software for constructing such user interfaces must be improved. Libraries for graphical user interface (GUI) software allow their more rapid development with less effort. Although current GUI libraries and tools are useful for the development of WIMP-based interfaces, they are of no use for making more interactive parts of user interfaces. Most programmers still write these from scratch using primitive library functions.

The objective of this study is to reduce the cost of developing GUI software. In particular, we aim at two types of interfaces: One is direct manipulation of pictures, figures, and diagrams that visually represent the structure of abstract data in an application. Another is animation of changing abstract data in an application which shows transitions of abstract objects and relations visually. They are not merely an aggregation of buttons and menus, and they cannot be developed only by combining WIMP widgets provided by user interface toolkits such as Java Swing set[125] and GTk+[98]. Many researchers in this area are addressing this problem, and a number of techniques have been developed for this purpose, including Programming by Example (PBE) and Programming by Demonstration (PBD). Garnet[91] and its successor, Amulet[14], are well-designed toolkits for GUI construction. These systems incorporate a number of innovative techniques and ideas. However, they are insufficient, in that they do not support the development of user interface software that can facilitate flexible and intelligent visual communication.

Our approach is, first, to create an integrated framework for visualization, recognition, and animation. This framework decomposes the processes of visualization, recognition, and animation into several sub-processes. We can implement visualization, recognition, and animation programs by developing modules that achieve each sub-process. The merit of creating a framework is that it avoids ad hoc construction of user interface software. We develop the integrated framework by extending Kamada's general framework for visualizing abstract objects and relations[65, 67], which is the model of the TRIP system. TRIP realized general visualization of abstract objects and relations based on its framework for general visualization. It regards the visualization process as the *translation* from application's data representations into pictorial representations. However, it handles only visualizations of abstract application data. This thesis extends their framework to accommodate recognition of drawn figures and animations that depict change of abstract application's data.

Second, we examine whether each functional module in the framework can be shared among different instances of GUI software. Sharable modules can be built in advance and provided as a library/tool for building GUI software. The essential part of the GUI software cannot be shared, but we have made it possible to build with simple specifications. Third, based on this framework, we have built tools that support the construction of GUI software. We have built several systems (TRIP2, TRIP2a, TRIP2a/3D), and applied them to the construction of various applications. Using these tools, users can construct GUI software by providing only declarative specifications. We have also developed a tool for visualizing the structure of constraints in the specifications, which is helpful to understanding and debugging the specifications.

1.2.1 TRIP

As described above, our integrated framework is originated in Kamada's general framework for visualizing abstract objects and relations[65, 67]. We have extended it by integrating recognition of figures/diagrams and animation of changing abstract objects and relations. This section briefly describes Kamada's framework, and the prototype system TRIP developed based on the framework.

The objective of their work is to fill the gap between application and graphics systems that is usually only a set of drawing instructions. There were few graphics tools that are helpful to

visualize complicated high-level abstract relations in an application into pictures that represent their structures. TRIP and its framework intended to explore a general approach to the visualization of abstract objects and relations, and help visualizing abstract application data.

The framework models the visualization process as the *translation* from textual languages into pictures. A set of pictures can be viewed as a visual language, in the sense that picture elements are arranged under certain layout rules. Thus, visualization is the translation from textual languages into visual languages. When translating textual data into pictorial data, the programmer should determine the layout of pictures so that viewers can easily understand the underlying structures. The programmer needs to map a relational structure in textual data into a layout rules of picture elements. We call this mapping as *visual mapping*. In the framework, application data are represented by abstract objects and relations, which are mapped to graphical objects (such as boxes and circles) and graphical relations (such as horizontal/vertical listings and line connections). The mapping is specified not with procedures, but with declarative mapping rules. Thus, the programmer can easily change the layout of pictures by changing declarative mapping rules.

TRIP (TRanslation Into Pictures) is a prototype system based on the visualization framework. It is independent of both input textual languages and output visual languages and can handle various visualization problems. The process of visualization proceeds as follows (Figure 1.1):

- First, an analyzer parses original textual representation data, and generates the relational structure representation data. Analyzers depend on the textual representation, because it needs syntax data of original textual representation for parsing them. Parser generators such as YACC[77] can be used to help developing analyzers.

- Second, relational structure data are mapped to visual structure data. Abstract objects are mapped to graphical objects, and abstract relations among them are mapped to graphical relations among corresponding graphical objects. In TRIP, relational structure and visual structure are represented in Prolog. Abstract relations and objects are represented by Prolog predicates. Graphical relations and objects are special predicates that are defined in COOL — a picture layout module in TRIP.

- Finally, actual layout of graphical objects are computed by solving graphical constraints and a picture is generated. They basically handles diagrams structured by horizontal/vertical listings and explicit line connections. They developed COOL(COnstraint Object Layout system) to layout picture elements. It solves constraints and outputs pictures. It can handle over-constrained equation systems by using least squares method. It also has general undirected graph drawing module, which is useful for visualize network diagrams. The graph layout module uses the spring-model algorithm[66].

Using TRIP, they generated various diagrams from abstract textual data. For example, they generated kinship diagrams from English sentences that states kinship relations, list diagrams that show cons-cells from S-expressions, and E-R diagrams from ER schema. In addition, they have shown that the programmer can generate various types of tree diagrams from the same tree data only by changing the mapping rules.

1.3 Objectives

The objectives of this thesis are as follows:

original textual
representation

Analyzer ◀— syntax data

relational structure
representation

visual mapping ◀— mapping data

visual structure
representation

COOL ◀— layout library

target pictorial
representation
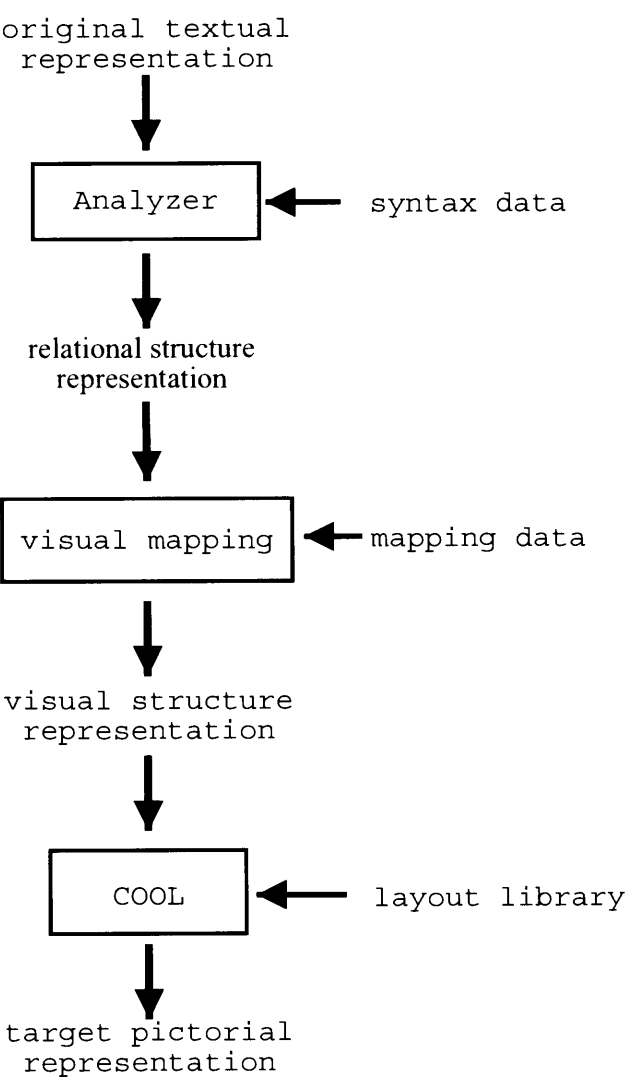
Figure 1.1: The visualization model of TRIP (from [67]).

An Integrated Framework

The first objective of this study was the development of a general framework for constructing visualization, direct manipulation, and animation interfaces. This framework describes the structure of these user interfaces, and supports the development of toolkits for building direct manipulation and animation interfaces. According to this framework, these interfaces consist of modules that translate four types of data representations, i.e., application data representations, two intermediate data representations, and pictorial representations, into the other types of representation. Thus, the development of these user interfaces allows implementation of translators among four data representations. Essential translators are those between two intermediate representations. These must be created by the user for each target application, but they can be built by writing only declarative mapping rules that specify the translations. Other translators are common among different targets, so they can be built as a library and provided for the user.

The basis of this framework is the model of the TRIP system[67], which handles the visualization of abstract data, and is regarded as a translation from abstract data into pictures. Our framework extends this to handle the recognition of pictures and animations, which are also regarded as translations between data representations in the framework.

Tools implemented based on the framework

The second objective of this study was the implementation of three tools based on the above framework. One is the TRIP2 system, a tool that can visualize abstract data as a figure, and can also recognize modifications performed on the visualized figure. This function can be regarded as a kind of direct manipulation of abstract data, which is achieved by bi-directional translation between abstract data and figures specified by a set of visual mapping rules and a set of inverse visual mapping rules. The other two tools are the TRIP2a and TRIP2a/3D systems for making animations that represent the execution of the programs. In our framework, animations are also achieved by translations generated by translating a series of changing abstract data into a corresponding series of changing pictures. Therefore, animations can also be specified with declarative rules, i.e., a set of visual mapping rules and a set of transition mapping rules. TRIP2a was developed by extending TRIP2, and it is integrated with the TRIP2 system. Therefore, both direct manipulation of abstract data and animation of changing abstract data can be utilized to develop interactive applications that show algorithm animations. TRIP2a/3D is another algorithm animation system, which is a re- developed system for animating the execution of programs. This system has been extended to handle three-dimensional animations and a type of event-driven asynchronous animation.

With these tools, we can create applications that have a kind of direct manipulation user interface and animations by providing only declarative mapping rules. These tools can reduce the cost of implementing interactive user interfaces.

Methods for Visualizing Constraints

The third objective of this study was the development of a new method for visualizing constraints in the visual mapping rules. Although this method was developed mainly for debugging visual mapping rules that specify the translations among four data representations, it can be applied for visualization of general constraints. We have developed a visualization system using this method. This system visualizes constraints in two ways. The first is to visualize them as a three-dimensional graph structure, with nodes in the graph as constraints and constrained variables. Edges connect a constraint and the variables by which it is constrained. By showing a constraint system as a graph structure, the viewer can grasp the overall structure of the constraint system. It is also possible to

unfocus some constraint nodes in the graph so that its structure becomes simple. It helps the user to browse the constraint graph. The other uses our new method, which animates the degree of freedom in the under-constrained system. This method displays the target figure itself, and tries to 'move' each graphical object in the figure, which may have various constraints. If the object is sufficiently constrained, the system will not be able to move it, while if it is not constrained and has some degrees of freedom, the system will be able to move it. Therefore, when there are some degrees of freedom, the system shows an animation in which objects are moving in some way. Otherwise, the system shows an animation indicating that the object cannot move. This animation can help users to readily recognize any lack of constraints.

1.4 Overview of the Thesis

In this thesis, the framework for visualization, recognition, and animation is first presented. Second, several systems based on the framework are described, and the framework is verified based on knowledge acquired through construction of the above systems.

Chapter 2 surveys various work in related research areas; i.e., visualization, user interface toolkits, algorithm animation, and constraints. Chapter 3 describes the basic and extended frameworks. The basic framework handles the visualization of abstract data and recognition of visualized figures, and the extended framework handles animations. Chapter 4 describes TRIP2, an environment for building direct manipulation interfaces of abstract data. By providing only a set of visual and inverse visual mapping rules, the user can visualize abstract data, modify the visualized figure, and reflect the modified figure to abstract data. Chapter 5 presents TRIP2a, a system for animating the changing internal data during the execution of programs. The user can create animations by specifying a set of visual mapping rules and transition mapping rules. Various examples of animations created with TRIP2a are presented in this chapter. Chapter 5 also describes an extension of TRIP2a that is capable of handling asynchronous motion in animations. This was achieved by slightly modifying the framework to accommodate start and stop events of motion. In addition, as an appendix, we describe TRIP2a/3D; i.e., a tool for making three-dimensional algorithm animations. Several sample animations generated with TRIP2a/3D are also presented. Chapter 6 describes a new visualization method and the tool for visualizing constraints in the VSR generated by visual mapping. This is useful for debugging visual mapping rules used to generate the VSR from ASR data. Finally, Chapter 7 summarizes this thesis and its conclusions, and proposes some directions for future work in this field. The appendix also introduces the TRIP3 system and the IMAGE system; i.e., work on the interactive generation of visual mapping rules for TRIP systems, described here because it is closely related to the work discussed in this thesis.

Chapter 2

Related Work

2.1 Introduction

This chapter describes research into support software for building graphical user interfaces, especially GUIs that cannot be built only by composing WIMP widgets, and is organized into four sections: visualization, constraints, recognizing pictures, and GUI toolkits.

In the *Visualization* section, various areas of visualization software are summarized. The main concerns are first how to effectively present various types of information to the user, and second how to reduce the cost of implementing such visualization. This section includes sub-sections on, for example, information visualization, performance visualization, and algorithm animation. Graph-drawing systems are also described in this section.

Constraints are among the most important techniques in GUI software. The Constraints section describes user interface software systems using constraints, and constraint solvers for GUI software.

The *Recognizing Figures* section summarizes spatial parser generators and systems using spatial parsing, i.e., modules that parse drawn figures according to the specified syntax of two or three-dimensional figures. Although the range of figures that they can handle is limited, spatial parsers are useful for implementing applications that accept drawn figures as their input.

The section on *GUI Construction Tools* summarizes research on software and techniques that help in the development of various types of graphical user interface software. The main topics are Programming by Example-based systems and interface builders.

2.2 Visualization

2.2.1 Information Visualization

Information visualization is a research area that studies how various types of information can be presented effectively to the user. It facilitates the navigation of large information spaces and extraction of hidden structure and meaning. One of the most important issues in this area is scalability, that is, how to handle large amounts of information.

One approach to address this issue is to use three-dimensional representation. By using three-dimensional display, one more dimension can be exploited to devise the representation. For example, in the 3D/Rooms of Information Visualizer[19], the user can visit a number of rooms in which various types of three-dimensional figures are floating. For example, the Cone Tree[102] is a three-dimensional tree that represents hierarchical information. The user can click the node in the tree so that the tree is rotated and the clicked node is moved to the front. These rotations are animated

smoothly to prevent the user from losing track of the node. Subtrees can be displayed as translucent cones to hide unimportant nodes. The Perspective Wall is another figure in 3D/Rooms, which shows calendar-like information on a curved wall. The center of the wall is placed near the user, so it is magnified in the display, and the two sides are presented diagonally to the user, so areas closer to the ends of the wall are displayed as smaller on the screen.

Another approach to this issue is zooming. In zooming interfaces, everything is placed on one large plane, and the displayed view is a magnified, focused, area of the plane. The user navigates through the plane by zooming in/out with a moving focus point. The distortion and fisheye view techniques are similar to zooming. The basic idea of these techniques is that everything should be displayed on the screen, the focus point should be magnified, and other areas should appear diminished in size. As the focus point is always surrounded by the diminished area that serves as its context, the user can easily find and move to the target point in the search space. The most well known research into the use of the fisheye-view technique is the work of Furnas[43], who visualized hierarchical data structures using this method. Koike[71] presented a fractal view technique that can control objects to be displayed in average. Continuous zoom[7] is another zooming algorithm that handles a rectangular 2-D display space.

Another issue of information visualization is how to support design tasks for visualization. Visual design of effective information presentations is not easy for most users, and a great deal of work is required to support these tasks. For example, Mackinlay developed APT[79], which can automatically design effective graphical presentations such as bar charts and scatter plots of relational information. In this work, he codified the graphical presentation technique and developed *composition algebra* that can form complex presentations from that technique. SAGE[104] systems are tools that interactively support graphical design. SAGE can be used for designing according to partial user specifications that serve as design directives. It is also possible to draw and combine graphical elements interactively, and to browse and retrieve previously created pictures for use as design examples.

It is also important to enable users to interact with visualized information. Dynamic query interfaces[2] enable the user to control the range of information displayed with various tools. One such tool is a range slider that can specify the range of values easily, which is useful for filtering the information to be displayed. The results of making changes to the tools are instantly fed back to the display so that the user can easily access the desired information.

2.2.2   Visual Programming Languages

Most visual programming languages have a non-textual visual representation of programs; i.e., the programmer can *draw* programs in visual form. For example, Prograph[29] is a commercial dataflow-driven visual programming language. Its program is a diagram that consists of icons that correspond to operations and links that connect icons. CafePie[96] is a visual programming system for CafeOBJ[34]. It visualizes the program and its execution with animations. The user can customize its view interactively.

The merit of visual programming languages is that they can depict the overall structure of programs naturally. The relations among program modules become easier to understand, in the same way as CASE tools. However, they are not always suitable for programming small details because they require such a large space for drawing programs and it is often more tedious to draw programs with a mouse than it is to write using a keyboard.

2.2.3 Performance Visualization

Performance visualization aims to provide performance data regarding program execution in various ways. The user examines the visualized information for debugging and performance tuning.

Performance visualization is particularly important when using parallel programs, the behavior of which is difficult to understand. ParaGraph[78] is a system for performance visualization that handles parallel programs in a message-passing multi-computer architecture. ParaGraph visualizes trace data monitored during actual program execution. Program executions can be replayed pictorially based on trace data. In addition, various aspects of performance data can be extracted from trace data and visualized appropriately. Pablo[100] is a similar performance visualization system for massively parallel architecture, which has been extended to SvPablo[33]. Paradyn[86] is also a performance visualization system for parallel and distributed systems; it can identify possible performance bottlenecks in program execution.

2.2.4 Algorithm Animation

Algorithm animations are animations that depict the behavior of algorithms. The main purpose of algorithm animation is education. A previous study has shown that algorithm animations are useful for learning algorithms[118].

Sorting out sorting is one of the most impressive algorithm animations, which first explains the usefulness of algorithm animations. It animates several sorting algorithms simultaneously, which greatly assists the user to compare the behavior of each algorithm. The process of sorting is animated as moving points that correspond to the elements being sorted. The y-coordinate corresponds to the value of the element, and the x-coordinate corresponds to the position of the element. The points are finally placed in a line. The changes in the distribution of points indicate the characteristics of the sorting algorithm.

Most subsequent algorithm animation systems are environments to facilitate the creation of algorithm animations. The best-known systems in the early days were BALSA and BALSA-II[15]. These systems have been applied to many algorithms, such as sorting algorithms, bin-packing algorithms, and many types of graph algorithms, and they have been widely used by students and researchers. The concept of *interesting events*, which inform the animation system that an "interesting" operation has been executed, was introduced by the BALSA system. Zeus[16, 17] is the successor of BALSA II, which is implemented on Modula- 3. However, neither BALSA-II nor Zeus directly supports the creation of the graphics parts of animations in a declarative way.

TANGO[115, 116, 114] is based on a *Path-Transition* paradigm. In this model, there are four data types: location, image, path, and transition. An animation consists of transitions comprised of pairs of images and paths. An image is an object drawn on a screen, and a path describes a series of locations. TANGO also uses interesting events for calling these animations. The path can be generated by the user's demonstration using the Dance system[117]. However, we believe that it is more laborious to specify complex animations involving compound data structures with TANGO than it is with our system. Polka[61] is the successor of TANGO; it allows the creation of 3D animations of concurrent programs.

Pavane[28, 103] is a declarative animation system. This system visualizes programs written in the *Swarm* language, which is based on the state-transition model. Animation is specified by declarative mapping from *State* space, which is a set of tuples that represent the data of the program to be visualized, to *Animation* space, which contains graphical objects with time-dependent attributes. The difference between this and our model is that our method of specifying an animation is more abstract, and is thus more concise. For example, using our system, the programmer can specify

the motions of objects, including animation effects, by simple transitional operations. In addition, the layout of a picture can easily be specified using graphical constraints in our system, which also makes it easy to specify an animation.

Animus[36] uses *temporal constraints* for creating animations. These are specified in the same manner as with ThingLab[11] and ThingLabII[80]. In Animus, the appearance and structure of a picture, as well as how pictures evolve over time, are described by constraints. Dynamic mechanical and electrical circuit simulations, problems in operating systems, and geometric curve-drawing algorithms have been animated using Animus[36].

Pictorial Janus[63, 64] is a visual representation of Janus, a general-purpose concurrent constraint programming language. It is *complete*, which means that pictures are programs and represent screenshots of computation, and it completely corresponds to the text version of Janus. In addition, the execution of programs (pictures) can be animated.

Bentley and Kernighan developed ANIM[9, 10], which aims to allow easy creation of animations. In this system, an animation is made from a *script file*, which is output from a target application program that contains commands for drawing an image, such as box and line.

Some systems incorporate the techniques of cartoon animation[76]. The user interface of Self[20] utilizes such techniques for the movement of windows, and Artkit[108] provides animation abstraction, which also supports cartooning techniques.

2.2.5 Graph Drawing Systems

Graph-drawing produces a nice layout of a given graph $G = (V, E)$ (where V is a set of vertices and E is a set of edges). The problem is formalized as positioning vertices $p \in V$ by determining the mapping $\phi : V \to R^2$ (or R^3). Edges are straight lines among the positioned vertices. However, when applying them to real applications, vertices may have size and occupy some area, and edges may bend or curve.

Graph-drawing algorithms can be classified by the types of their target graphs,which have different criteria for appropriate layout. For trees, a number of algorithms to minimize the area required to draw a tree have been proposed. Algorithms for planer graphs draw figures without edge-crossing. Directed graphs are drawn so that the edges are oriented in one direction. Graph-drawing algorithms are summarized in the annotated bibliography of Tammasia [8].

General graph drawing is much more difficult than specific graph drawing because no assumptions about the structure are possible. In addition, simple criteria are often not enough to allow the drawing of general graphs. For example, there are cases in which minimizing edge-crossing is not natural, and it is necessary to maintain the balance of various criteria.

One of the most important approaches for general graph drawing is the force-directed approach in which nodes in a graph are treated as material points. Pulling and repulsing force are exerted on nodes. For example, in the spring embedder[37], each edge in a graph is treated as a spring. This system places nodes at random positions, and simulates their movement. As it is assumed that there is friction, they eventually reach the state of minimum energy. In Kamada and Kawai's algorithm[66], every pair of nodes in a graph is connected by a spring the length of which is equal to the graph- theoretical distance between the two nodes. The layout is obtained by moving each node one by one to minimize the overall energy of this dynamic system.

With these force-directed approaches, general undirected graphs can be drawn quite well. However, these force-directed algorithms are very slow and do not scale well. Drawing large-scale general graphs, such as graphs with over 1000 nodes and edges, is difficult.

The multi-scale approach recently proposed by Hadany and Harel[46] and in subsequent studies by Harel and Koren [47] promises to target large graphs such as those with over 3000 nodes. Their

key idea is a coarse-scale representation of the graph that represents its simplified structure. This representation consists of clusters of nodes and edges between clusters. The algorithm lays out fine-scale positions of nodes locally in a cluster, and then arranges coarse-scale positions of nodes globally. They claim that a graph with over 3000 nodes can be arranged in a few seconds using this system.

In real applications, it is often necessary to reflect the user's preferences. There are several graph-drawing tools in which the user can select an appropriate drawing algorithm. These include GraphEd[50], Gem[1], GraphViz[38], or Graph Editor/Layout Toolkit by Tom Sawyer Software[112].

Henry and Hudson[49] proposed a graph-drawing tool that can apply different algorithms to subgraphs in a graph. The subgraph is interactively selected by a specific graph-traversing algorithm. For example, the user can select nodes on the shortest path in a tree, and arrange them in a row. The other nodes in a tree are then drawn using an ordinary tree layout algorithm. The merit of this approach is that it is possible to handle very large graphs by selecting and drawing only subgraphs.

GLIDE[68] is another system in which the user can interactively lay out graphs. The user can specify various constraints among nodes. For example, the user can specify that selected nodes should be aligned horizontally or positioned circularly. All constraints in GLIDE are solved by a force-directed approach, and the process of solving constraints is animated so that the effects of constraints can be well understood by the user. Constraints are regarded as various 'springs' among nodes. Therefore, the user can specify conflicting constraints that can be solved approximately by balancing these constraints.

Masui[83] proposed a graph-drawing system based on a genetic algorithm. The user provides the system with several pairs of examples, and the system infers the evaluation functions using the genetic programming technique.

ILOG JViews[106] is set of Java components for building interactive graph-based user interfaces. It provides various components for cartography, charting, resource scheduling, and graph-layout. The method of visualization can be specified using XML style sheets.

2.3 Constraints

2.3.1 Constraint Solvers for GUI software

Constraint-solving problems are studied widely in various areas such as artificial intelligence and operations research. Various resource allocation problems and task scheduling problems are typical examples of constraint-solving problems. The merit of using constraints is that programmers have only to write the problem to be solved declaratively and do not need to specify how the problem should be solved. The constraint solver solves the specified problems automatically.

Constraint solvers have also been applied for construction of graphical user interface (GUI) software. For example, they are useful to determine the layout of parts of the GUI on the screen, and to reflect the application's internal changes/events on the screen. Several constraint solvers have been developed for GUI software. As they were developed for interactive GUI software, they have different characteristics from those developed for other areas. The characteristic functions required if constraint solvers are to achieve interactive systems are:

Preference A typical use of constraints for constructing GUI software is to specify and determine the layout of graphical parts. For example, the layout of a slider S placed at the right side of

a window W can be specified with a constraint:

$$S.left_x = W.right_x$$

A constraint specifying that the heights of the two windows are the same is:

$$S.height = W.height$$

$S.left_x$ is the x coordinate of the window S, and $W.right_x$ is the x coordinate of the window W. $S.height$ and $W.height$ are the heights of each window. Most constraints to determine the layout of widgets are required to be satisfied.

There are also esthetic criteria that must be considered when creating a picture. The programmer must manage screen space, and use constraints to place objects neatly. For example, any lines connecting objects should not cross each other. Graphical objects must not be too close, and the distances between objects should be even.

However, it is not always possible to satisfy such criteria, and it is not always necessary to satisfy all of these criteria at the same time; i.e., there are *preferences* for constraints. Therefore, it is necessary to provide a way to specify and handle such preferences.

In the TRIP system, two levels of constraints can be used: *pliable* and *rigid*. Rigid constraints are solved exactly, before pliable constraints are solved approximately, i.e., rigid constraints have precedence over pliable constraints.

In the IMAGE system, the DETAIL constraint solver is used. The DETAIL solver incorporates constraint hierarchies[12]. That is, each constraint has a strength value, and the solver satisfies as many of the higher valued constraints as possible. DETAIL can solve constraint hierarchies incrementally, using local propagation, and still handle cyclic constraints[55].

Constraint hierarchies are also useful for specifying the dynamic behavior of interactive systems. For example, the behavior of a rectangle when the user points to it, i.e., whether it should be resized or moved, can be specified by constraints on the points: the width and the height, each with a strength value. To move a rectangle, the value of the constraints that maintain its width and height should be increased. On the other hand, to resize a rectangle, the constraints that prevent a diagonal point from moving should be given higher values. The system only changes the values of constraints according to the operation specified by the user.

Furthermore, constraint hierarchies are useful for handling over/under-constraint systems. In general, for interactive applications, the values of constrained variables should always be determined. However, solving over/under-constraint systems is a difficult problem, and it is generally difficult for programmers to avoid such systems, i.e., to specify constraints on variables with sufficient precision to determine the values of the constrained variables. By using constraint hierarchies, the programmer is able to use weaker constraints to prevent under-constrained systems, thereby assuring that constraints are always solvable.

Performance Where constraints are used to determine the behavior of graphical components on the screen, as described above, they must be solved quickly, as this influences the smoothness of the interface, and affects its usability.

There are several ways to cope with this problem, one of which is to solve constraints *incrementally*. Constraints used in interactive systems may become numerous, but they tend to be clusters of independent constraints. Therefore, once all of the constraints are satisfied, constraint solvers only need to maintain those that are affected by changes made to the

constraints. DeltaBlue[41] is a well-known incremental constraint solver that uses local propagation, and solves non-cyclic linear equations with strength values. We have developed the DETAIL constraint solver[55], which is also an incremental constraint solver. DETAIL can have multiple types of sub-solvers, such as a simultaneous equation solver and a least squares method solver, and can solve multiple types of constraint.

HiRise[51] is another incremental constraint solver, which solves constraints utilizing LU decomposition so that it is able to solve thousands of constraints in a few milliseconds. HiRise can also handle inequality constraints.

Complexity Linear equations are not sufficient to describe the behavior of interactive systems. There are more complex types of constraint, which are useful for describing various behaviors in interactive systems. For example, a general constraint that two lines should be parallel is not described by linear constraints. A constraint that maintains the distance between two points is also not linear[1].

Inequality constraints are also useful for GUIs. For example, the constraints that some elements must be put inside a window are represented as conjunctions of several inequality constraints:

$$win.left_x \leq obj.left_x \qquad (2.1)$$

$$obj.right_x \leq win.right_x \qquad (2.2)$$

Cassowary and QOCA[13] can handle inequality constraints. They solve constraints as an optimization problem using the simplex method. In addition, constraints that objects do not overlap are represented as disjunctions of the above constraints. Such disjunctive constraints are also described by Kim Marriott et al. [69]

Chorus[53, 52] solves non-linear constraints using the optimization method. To cope with the local minimum problem, it also uses the genetic algorithm. Chorus is so expressive so that it can specify force-directed graph layout algorithms. In addition, the solver is fast enough to drag and modify the layout of a graph interactively.

However, in general, expressive constraints often require inefficient constraint-solving algorithms, such as the Newton-Rhapson method and the relaxation method. This is why most of the constraint solvers in use in current interactive systems solve only one-way constraints using local propagation. The DETAIL solver copes with this problem by dividing constraint systems into those sets of constraints that can be solved by local propagation and those that cannot be solved. DETAIL then applies inefficient constraint-solving algorithms only to those parts that actually require them. Such combinations of different constraint solvers are promising approaches for interactive systems.

2.3.2 Other Systems Using Constraints

Constraints are used not only to support the construction of GUI applications but also as interaction techniques in GUIs. Most drawing editors have alignment and grid functions that can be thought of as types of constraints. In some editors, users can use connection constraints in which a line connected to an object follows that object when it is moved. IntelliDraw[27] is a commercial drawing application that has many constraints functions. Various types of geometric constraints are provided

---

[1]Of course, it is possible to specify that a line should be parallel to an axis, or to specify that the horizontal/vertical distance between two points should be constant.

for the user. However, IntelliDraw has not met with much success; one reason is that it is difficult to specify constraints directly in this application; there are too many icons representing the types of constraints from which the user can select. It is also difficult to combine constraints and to predict the effects of doing so. Furthermore, ordinary drawing may not require complex constraints.

The key issue of using constraints in a GUI is providing an easy and intuitive way to specify constraints. *Inferring* is a common technique used to address this issue. The user does not specify constraints directly but shows the state that satisfies the constraints the user intended. The system then infers constraints from the presented state. For example, Chimera[72] infers constraints in a picture from multiple examples of pictures. Characteristics that are invariable among examples are regarded as constraints.

For example, the constraints of a hinge, two lines that are connected at one end of each line and can be rotated at that point, can be extracted from several examples of a hinge with different angles. Pegasus[56] is another pen-based drawing tool that infers constraints in a picture. In Pegasus, the user can draw only straight lines, but the application can recognize several relationships among lines such as parallel, perpendicular, same length, and symmetry. Pegasus infers lines to be drawn based on recognized constraints, and presents them to the user. The user can select one from this list or can neglect them by drawing another line.

2.3.3 GUI Toolkits Using Constraints

Constraints are incorporated into several GUI toolkits. Garnet[91] and its successor, Amulet[14], have a simple one-way local propagation-based constraint solver. The programmer can define constraints among variables in the program so that when the values of source variables are changed, the target variables are updated appropriately.

Various toolkits use constraints for the layout of components in GUIs. SubArctic[107] incorporated a one-way constraint solver for the layout of components in a GUI. *Open Inventor*[126], which is a three-dimensional graphics toolkit, can have nodes called *engines* in the scene graph, which also correspond to one-way constraints that connect attributes of other nodes. Recent GUI toolkits have a mechanism of layout management such as the layout managers in Tcl/Tk and Java AWT. These can be thought of as special constraint solvers that solve predefined layout constraints on target widgets.

2.4 Recognizing Figures

2.4.1 Spatial Parser Generators

Spatial parsing is a function that parses pictures drawn by the user according to the grammar specified by the programmer. The *spatial parser generator* generates *spatial parsers* from the grammar specified by the programmer. Ordinary drawing editors with spatial parsers can be used as interaction modules for visual language systems.

Spatial parsing has been studied by a number of researchers for a long time. For example, Chang proposed "picture-processing grammar" [22]. Lakin developed the PAM system and the VMACS editor [75]. VennLISP[73], a visual notation of LISP, is one application of these systems. Various grammars have been proposed. For example, Wittenburg proposed relational grammar[127] and bottom-up and Earley-style parsers. Rekers proposed the use of graph grammars for graphical languages[58], which were later refined to layered graph grammars[101]. Constraint multi-set grammar (CMG)[82] is another well-known example of grammar for parsing two-dimensional figures. The production rules of CMG have the form $P ::= P_1, P_2, ...P_n$ where C, which means that

$P_1, P_2, ..., P_n$ can be reduced to the non-terminal symbol P whenever the attributes of all symbols satisfy the constraint C. Terminal symbols represent primitive shapes such as lines and boxes. Eviss[5] and its successor Rainbow[62] is an example of a visual system that uses CMG.

2.4.2 Systems that Parse Figures

There are many systems that can recognize figures drawn by the user, and utilize them as input to the application. Most visual language systems have built-in parsers that interpret visual programs. An interface builder is a type of VL system. Programmers can lay out widgets visually on the screen, which are then compiled into the GUI modules in applications. Various CAD systems also interpret drawings, and these systems are used for mechanical simulations. Some CASE tools can recognize UML diagrams and convert them into textual programs. However, they usually do not utilize spatial parser generators. There are several reasons for this. First, the spatial parser generators available at present are not very popular. In addition, it is often difficult to understand complex figures in real applications.

Another type of built-in spatial parser can interpret freehand drawings. For example, the Electronic Cocktail Napkin project by Mark Gross[81] involved development of a pen-based system that can recognize hand-drawn sketches and diagrams. The ultimate aim of this project is the development of a pen and intelligent paper that can support the sketching of designs on the screen. Teddy[57] is another example that enables freehand drawing of three-dimensional polygon models. By only sketching outlines of models, users can create 3D polygon models, which can be used in the design of objects such as the stuffed toys after which it is named.

2.5 GUI Construction Tools

Writing a graphical user interface (GUI) used to be a complicated process, but a great deal of work has gone into reducing the cost of making a GUI. Various toolkits, libraries, and environments are now available for constructing GUI applications. However, most still provide only a basic set of widgets and primitive graphics libraries. This section summarizes the work that goes into making GUI applications, focusing on programming by example, programming by demonstration, and interface builders.

2.5.1 Programming by Example and Programming by Demonstration Approaches

Programming by example (PBE) and programming by demonstration (PBD) are techniques by which the user provides examples or demonstrations to the system, from which the system constructs "programs" for a given task. For example, EAGER[31] watches the user's operations in a GUI environment, and detects repeated patterns. When a repeated sequence of operations is detected, EAGER asks the user whether it should complete and repeat the sequence. This can be used as a "keyboard macro" in GUI environments.

These techniques have been applied to the construction of GUI applications. Peridot[90] is an early PBE-based user interface construction tool in which the user places graphical objects on a screen and can make them react to input. Peridot infers two types of constraints; i.e., *graphical constraints* that hold among graphical objects placed by the user, such as a constraint that the sizes of rectangles must remain the same, and *data constraints* that hold among graphical objects and the arguments of a function that the programmer is to construct. A *virtual mouse* is used to simulate the user's mouse input. DEMO, DEMOII[42], and Pavlov [128] are also based on PBD. The DEMO system has been used to make a rotating gauge that works together with a number panel, DEMO-II

to make xeyes-like applications by demonstration, and Pavlov to specify the user's control module in a car racing game.

2.5.2 Interface Builders

Most current software development environments, usually called Integrated Development Environments (IDEs), such as Microsoft's Visual Basic or Visual Studio, and Borland's JBuilder, have an *interface builder* that enables programmers to layout widgets interactively in a window and change their various properties such as colors and fonts. Some IDEs allow the programmer to specify call-back functions or event handler methods for widgets. This specified information is compiled and linked with other programs to build applications.

The NextStep Development Kit [95] is a well-known early IDE with an interface builder. The user can not only lay out the widgets interactively on the screen, but can also connect a widget to an object via a line that specifies the flow of events at the widgets. In this environment, the user can immediately try the interface without compiling the application.

SILK[60] is a user interface prototyping tool in which the programmer roughly draws a sketch of the user interface with a mouse or a pen. The drawn sketches are recognized by SILK, and the corresponding widgets are arranged automatically. In addition, the user can operate the rough sketch of the interface, i.e., the user can move the drawn scrollbar or push the sketched buttons.

Chapter 3

The Bi-Directional Translation Model

This chapter describes two models: the basic model of visualization and parsing pictorial data, and the extended model for animation. These models form the basis of this thesis, and have been applied to the systems described in the following chapters. Note that , as described in Chapter 1, our model is extended from the framework of TRIP[67] which only handles visualization of abstract data. We integrated parsing pictorial data and animations into the framework.

3.1   Overview of the Bi-Directional Translation Model

The objective of this model is to depict the standard structure of implementations that provide two functions: visualization of application data and interpretation of figures. They are important functions when building direct manipulation interfaces, but are not fully supported in most GUI construction tools. This model aids in building direct manipulation interfaces whose domain is abstract data structures such as trees and graphs.
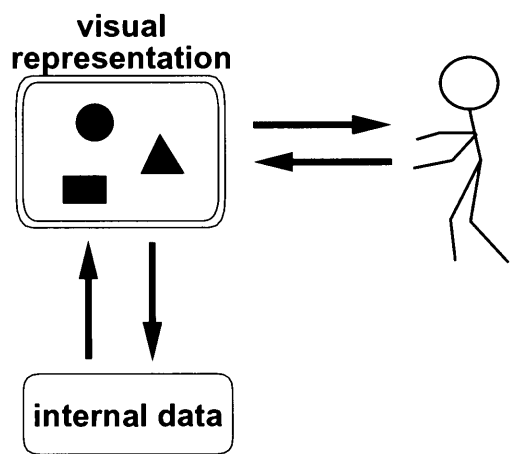
Figure 3.1: The idea of direct manipulation interfaces.

Figure 3.1 outlines the idea of direct manipulation interfaces for the data visually represented on the screen. The diagram on the screen visually represent internal data of the system. The user recognizes the diagram as internal data, and operates on the diagram. These operations on the diagram are interpreted as operations on internal data. That is, the system *visualizes* application data to the diagram on the screen, and *interprets* the diagram on the screen to application data. In our model,

visualization is thought of as translation from source data representation into target pictorial data representation, and interpreting of pictorial data is thought of as the converse, i.e., translation from pictorial data representation into target data representation. To build direct manipulation interfaces, these two functions, visualization of internal data and interpretation of operations on the visual representation, must be implemented. Thus, two translators are required that can perform translations between pictures and internal application data.

Usually, programmers must use canned translators or write such translators from scratch, but such translators resemble each other in their translation process. Our model decomposes the bi-directional translation process into several sub-translations by incorporating intermediate data representations. It enables systematic implementation of bi-directional translations for direct manipulation interfaces. In fact, we have enabled such translations by a set of declarative rules, designated as a *visual (and inverse visual) mapping rule set*. To visualize abstract data, programmers have only to write a visual mapping rule set that maps the ASR of the original abstract data into the VSR of the target picture. In the same way, to interpret visualized pictures, the programmer has only to write an inverse visual mapping rule set. This topic is described in more detail in the following sections.

3.2   Data Representations

There are many types of both data representations to be visualized, and of target pictorial data representations. As the translation process depends on these source and target data representations, it is difficult to reuse the translators if translations are performed between these individual representations. Therefore, two intermediate representations are incorporated into the process of translations; i.e., abstract structure representation (**ASR**), and visual structure representation (**VSR**). ASR consists of a set of ground compound terms. ASR represents the structure of abstract data, and is used as a proxy for the application data in the model. VSR is a high-level representation of a picture, and consists of a set of graphical objects and graphical constraints. By using these two intermediate representations, programmers can specify the mapping between abstract data and a picture independently of each specific representation.

With these two intermediate data representations, our models has four data representations as follows:

Application Data Representation (AR)   Any application data can be the source for visualization. For example, data may be the result of a simulation, internal data generated during some computation, or logs of communications between processes. However, their primary purpose is not necessarily visualization. In fact, they usually include data that are not used for visualization or may not have a form appropriate for visualization. To visualize such data, necessary information must be extracted and transformed appropriately for visualization. Similarly, the data obtained by interpreting pictures must be converted for application data representation before they are passed to the application.

Abstract Structure Representation (ASR)   Abstract structure representation (ASR) is the user-defined data representation for representing application data in our model. In contrast to application data representation described above, the purpose of ASR is to represent direct data corresponding to the target picture. Therefore, ASRs should be designed so that they do not include data that are unnecessary for visualization, and that they are structurally similar to the corresponding picture.

ASR data are essentially a set of facts. The syntax of facts is the same as that used in Prolog,

i.e.,

$$predicate.$$

or

$$predicate(arg_1, arg_2, ..., arg_n).$$

where $arg_1, ..., arg_n$ are symbols, numbers, lists, or terms. Variables cannot be used as arguments. This syntax is primarily intended to represent symbolic structural and relational data, because, although this model is also applicable to other types of data such as numerical data, we mainly handle abstract structural and relational data such as hierarchical tree structures and network graph structures.

The user determines what kinds of facts are required to represent application data as ASR. For example, if the application has network data, a graph structure representation should be used as the ASR. A simple graph structure can be represented by two types of facts; node/1[1] and edge/2. For example, a complete graph with three nodes (a, b and c) can be represented as follows:

```
node(a).
node(b).
node(c).
edge(a, b).
edge(b, c).
edge(c, a).
```

For real applications, the definition of the ASR will be more complex. For example, it may have more types of nodes and edges, and nodes and edges may have more attributes to represent more data in the application. Therefore, the corresponding picture can be more complex.

Visual Structure Representation (VSR) Visual structure representation (VSR) is a high-level representation of pictures. It is high-level in the sense that it describes pictures with graphical objects and graphical relations among graphical objects. The VSR does not contain absolute coordinate values, but represents the structures of pictures directly. That is, graphical objects in VSR have their size information, but do not have positional information. Graphical relations represent relative positional relations among graphical objects, but do not specify their absolute positions. For example, the *horizontal* relation put the graphical objects horizontal. However, it does not specify each graphical object's x- and y-coordinate. This is because such graphical relationships are expected to be closely related to the meaning of the picture that represents an abstract relational structure, and thus the mappings between ASR and VSR become simple.

Graphical objects, which represent primitive shapes in a picture, are written as follows:

$$type(id, arg_1, arg_2, ..., arg_n, options).$$

The type describes the type of the graphical object, e.g., line, circle, or rectangle. Arguments describe the attributes of graphical objects, such as width and height. The number of arguments differs among the different types of graphical objects. Each graphical object has a unique identifier(ID) that is used to distinguish it from other graphical objects.

Graphical relations are written as follows:

$$type([id_1, id_2, ..., id_m], arg_1, arg_2, ..., arg_n, options).$$

---

[1]The syntax *predicate*/N represents a predicate with N arguments. For example, node/1 is a predicate named node with one argument.

Graphical relations also have a type and several attributes. Their first argument is a list of the IDs of graphical objects related to them. Other arguments are specific to each type of relation. Note that graphical relations do not have their own ID. Graphical relations that have the same type and attributes are treated as one relation.

The following is an example of VSR data, indicating two rectangles that are horizontally aligned:

```
rectangle(r1, 30, 30, []).
rectangle(r2, 30, 30, []).
horizontal([r1, r2], []).
x_order([r1, r2], 10, []).
```

The first and second lines are graphical objects with the IDs r1 and r2. Atoms, ground terms, and lists can be used as IDs of graphical objects. The third and fourth lines are graphical relations that indicate that r1 and r2 are arranged horizontally and that the distance between the x-coordinates of r1 and r2 is 10. As shown in the above example, they have an empty list of options as the last argument. The lists of options contain attributes such as colors of shapes and the method of alignment for geometric constraints. An empty list means that they use default values as attributes.

The available set of graphical objects and relations varies depending on the implementation of the geometric constraint solver in the system. For example, TRIP systems use the COOL system[67]. It has a linear equation solver and an undirected graph layout module, which are convenient for specifying the layout of hierarchical tree structures and network graph structures. The details of VSR implemented in our systems are described in the following chapters.

Picture Data Representations (PR)   Any pictorial data representation that is designed to display or print, such as various bitmap representations, saved files for drawing editors, PostScript data, or VRML, can be a target of visualization and a source for spatial parsing. Pictorial data representations are presented directly to the user, and the user interacts with displayed data. Thus, they should have concrete information regarding the coordinates of objects in the picture, and should be ready to be displayed. Conversely, differently from VSR, their structures may not be obvious. It is often necessary to devise a way to parse such representations and extract their structures.

3.3   Various Functions in the Model

Several functions related to visualization of abstract data and interpretation of pictures can be regarded as a combination of translations among the data representations described in the previous section. Figure 3.2 shows an example translation process in the model. Using this figure, this section introduces the interpretations of several functions in this model.

Visualizing Application Data   Visualization is a process of generating pictorial representations from source information. This process is modeled as shown by steps (1), (2), and (3) in Figure 3.2, which is same as the model of TRIP[67]. The first step of visualization is the extraction of ASR data from application data. The purpose of this translation is first to convert the syntax of data into the format of ASR, in order to prepare for the next translation process. Second, as described in the previous section, the application data are usually not prepared for visualization. Data that are unnecessary for visualization are often included, and data that are necessary may be distributed over several parts of the application data. This step extracts and aggregates necessary information from application data, and converts it into ASR. In Figure 3.2, translation (1) corresponds to this step. Here, two sentences, "X consists of p, q, and r." and "X, p, q, and r

```
"X consists of p,q,and r"     AR
"X, p, q, and r are objects"

    (1)          (6)

consists_of(x,[p,q,r]).
object(x). object(q).        ASR
object(p). object(r).

         | Visual
    (2)  | Mapping  (5)
         | Rules
above([x],[p,q,r],Ygap,[]).
hor-listing([p,q,r],Xgap,[]).
mconnect([x],[p,q,r],bottom,  VSR
    top,[solid,orthogonal]).
boxwithlabel(x,Width,Height,x,
    [visible]).

    (3)      (4)

              x


      p       q       r        PR
```
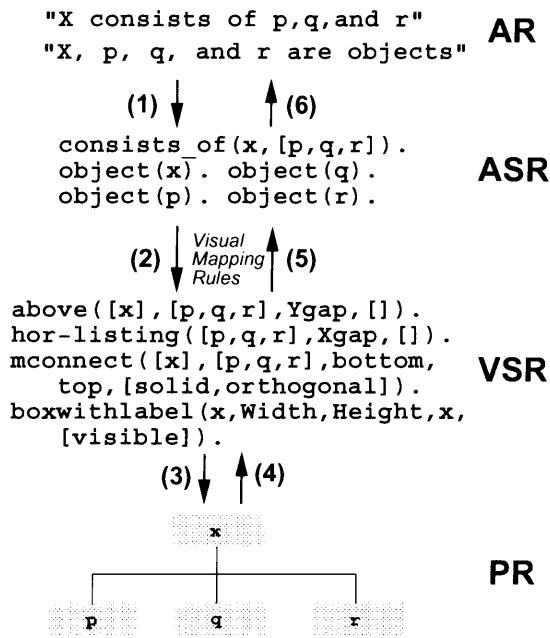
Figure 3.2: A translation example in the bi-directional translation model.

are objects," are translated into five terms. The term "consists_of(x,[p,q,r])" implies that there is a "consists_of" relationship between x and the list p, q, and r. The other four terms imply that x, p, q, and r are "objects". With this translation step, the structure of the application data has become more explicit. In other words, the programmer has extracted the structure to be visualized from the application data.

The second step is the translation from ASR into VSR, which we call *visual mapping*. In this translation, the visual characteristics of the target picture are determined according to visual mapping rules. By changing visual mapping rules, various pictures can be generated from the same ASR data. However, as the picture that we wish to generate by visualizing abstract data is a representation of abstract data from which the user can recognize its structure and meaning, there should be some correspondence between the structure of abstract data and the structure of the visualized picture. The programmer should specify visual mapping rules with this point in view. In addition, if they are structurally similar, visual mapping rules are expected to be simple. Abstract objects and relations in ASR are simply translated into graphical objects and relations in VSR. In addition, the expressiveness of VSR, that is, how many and strong graphical relations are supported by the implementation of VSR is related to the simpleness of visual mapping rules. If enough expressive graphical relations are provided in VSR that are convenient to depict abstract relations, visual mapping rules will be simple. If not, the programmer must combine complex constraints to specify the intended graphical relations.

In Figure 3.2, each object is mapped to a graphical object. For example, object(p) is mapped to

```
boxwithlabel(p, Width, Height, p, [visible])
```

This is a graphical object representing a box labeled with a string whose ID is p. The relation consists_of(x, [p,q,r]) is mapped to three graphical relations:

```
above([x],[p,q,r],Ygap,[])
```

```
hor-listing([p,q,r]),Xgap,[])
mconnect([x],[p,q,r],bottom,top,[solid,orthogonal])}
```

These relations lay out the objects so that (1) the object x is placed above the other objects, p, q, and r, (2) objects p, q, and r are arranged horizontally, and (3) object x is connected by a line to the other objects, p, q, and r. These mappings are specified by the visual mapping rules written by the programmer. Figure 3.3 shows the visual mapping rules used for the mappings in Figure 3.2. Each rule maps an ASR term on the left-hand side of a rule to VSR terms on the right-hand side. The syntax of rules varies depending on the implementation. The rules shown in Figure 3.3 are those for the TRIP2 system described in Chapter 4.

```
object(X) :-
      boxwithlabel(X,Width,Height,X,[visible]).
consists_of(X,L) :-
      above([X],L,Ygap,[]),
      hor-listing(L,Xgap,[]),
      mconnect([X],L,bottom,top,[solid,orthogonal]).
```

Figure 3.3: An example of visual mapping rules.

In the last step of visualization, PR data — a picture in the form of a target representation is generated from VSR. VSR data are a set of high-level graphical objects and geometric relations, and do not have absolute positional information. The layout of the picture are specified by the graphical relations among graphical objects in VSR. In order to generate PR data, positional information of the picture must be obtained. Therefore, in this step of translating VSR into PR, first, absolute coordinates of the elements in a picture are calculated by solving geometric constraints of graphical relations in VSR data. Second, target picture data are generated by substituting the calculated results for the coordinate variables of graphical objects. Although this step depends on the representation of target picture data, the module that solves constraints of graphical relations in VSR can be used for all translators for this step. In Figure 3.2, this step is shown as translation (3). The result of translation is the tree shown at the bottom of Figure 3.2.

In summary, visualization of abstract data is a process of translating AR into PR via ASR and VSR. It is a process of extracting the structure of the data to be visualized from the application data, mapping abstract objects and relations into graphical objects and relations, and calculating the layout of elements in the picture according to the corresponding graphical relations in VSR.

Interpretation of a Picture   A figure that represents abstract data is a proxy for that data. Thus, it can be used not only for viewing data, but also for handling abstract data. People draw and modify a figure when they want to create or change abstract data. Thus, we must provide not only a means to visualize abstract data, but also a means to reflect the changes to application data. In our model, this function is regarded as an inverse translation from pictorial data to application data (Figure 3.2, steps (4), (5), and (6)).

First, pictorial data are translated into VSR data, a process referred to as *spatial parsing* (Figure 3.2(4)). In this translation, geometric objects and relations in the picture are recognized by parsing the picture, and output in the format of VSR. The translator for this step must be constructed for each pictorial representation. The purpose of this translation is similar to the translation from AR to ASR. That is, the translation converts various PR syntax into the format of VSR in preparation for the next translation. In addition, the structure of pictorial data is not usually obvious. In

this translation from PR into VSR, the important graphical relations in the picture that related to its 'meaning' is extracted from the picture, and explicitly represented as VSR data. Note that there are often multiple ways to interpret pictures, and there are so many graphical relations unrelated to the interpretation of the picture. This translation step extracts only meaningful graphical relations from a picture in reference to the next translation step.

The second step of the inverse translation is the translation from VSR into ASR, which we call *inverse visual mapping* (Figure 3.2(5)). In this step, the "meaning" of visual information is determined; i.e., graphical objects and graphical relations in VSR are mapped to abstract objects and abstract relations in ASR. As this translation is the inverse of the translation from ASR to VSR, the visual mapping rule set that maps ASR into VSR can be applied inversely for this inverse translation. However, for translation between abstract data and pictures, there are some differences between visual mapping and inverse visual mapping. Visual mapping *adds* geometric information to abstract data in order to visualize them, because abstract data do not have their own shape or layout. On the other hand, inverse visual mapping must *eliminate* layout information, and extract only the structure of the picture. Therefore, mapping rules are applied differently in each direction of translation between ASR and VSR. When mapping inversely from VSR to ASR, some geometric constraints in visual mapping rules are often ignored. The system tolerates some error of geometric constraints. For example, when visualizing a node in a graph as a box, it is necessary to determine its size precisely with visual mapping rules. However, when recognizing a box as a node, all boxes of approximately the same size should be mapped to nodes to deal with pictures drawn not so precisely by the user. In Figure 3.2, the graphical object

```
boxwithlabel(x, Width, Height, x, [visible])
```

is translated into the object `object(x)` in ASR. The three graphical relations:

```
above([x],[p,q,r],Ygap,[])
hor-listing([p,q,r]),Xgap,[])
mconnect([x],[p,q,r],bottom,top,[solid,orthogonal])}
```

are translated into the abstract relation `consists_of(x,[p,q,r])` in ASR.

In the last step of inverse translation (Figure 3.2(6)), ASR data are translated to application data. The purpose of this step is mainly conversion of the data representation. ASR data are converted to various types of application data, such as data structures in C/C++ and S-expressions in LISP. The converted data are passed to the application.

In summary, interpretation of a picture is a process of translating PR into AR via VSR and ASR. It is a process of translations that extract important graphical relations from pictures, map graphical relations and objects in VSR into abstract relations and objects in ASR, and compose the application data from the ASR data.

Direct Manipulation of Abstract Structure Data   Using the bi-directional translation between application data (AR) and picture data (PR) as described above, a kind of direct manipulation interface, especially direct manipulation user interfaces for handling relational structure data, can be interpreted in our model.

As an example, here we consider the construction of a user interface for the application that handles entity-relationship (ER) diagrams as input. In this application, an ER diagram is displayed on the screen. The user draws a diagram, or modifies the diagram displayed by the system. The application has the schema data, upon which it computes, internally. The schema data are visualized and displayed to the user, and the user manipulates the visualized diagram. (Figure 3.4).
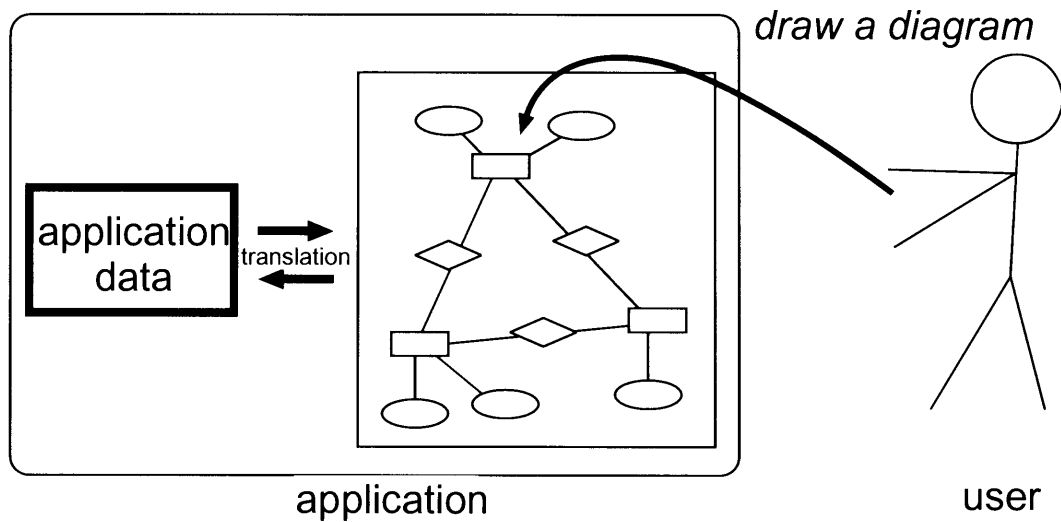
Figure 3.4: Architecture of ER Diagram Editor Application.

To construct the user interface part of this application, the following modules have to be implemented:

1. A module that draws a diagram that reflect the state of the internal data of the application.

2. A module that enables the user to draw and modify the diagram on the screen.

3. A module that recognizes the user's input, and sends it to the application.

The function of the first module is visualization, so that, in our model, it can be thought of as performing the translation from application data to pictorial data. On the other hand, the function of the third module is interpretation of a picture. That is, it corresponds to translation from pictorial data to application data. Therefore, these modules can be realized by bi-directional translations between application data and pictorial data, which are specified by visual mapping rules. The function of the second module is to allow users to draw pictures as the input to the third module. A general drawing editor module, such as Adobe Illustrator[2] and TGIF[3], is sufficient for this purpose, and can be shared among various applications.

In general, direct manipulation-style (DM-style) user interfaces can be modeled as two parts; the bi-directional translator between application data and pictures, and an interaction module for manipulating visualized pictures. TRIP2 is a system that aids in building these parts. With TRIP2, only a set of visual mapping rules is necessary to build DM-style user interfaces. Chapter 4 describes the TRIP2 system in more detail.

In summary, in our model, direct manipulation is modeled as a bi-directional translation between application data and pictorial data. Application data (AR) and picture data (PR) have extra information. ASR data and VSR data are extracted from them to represent the important structure of the data required for the bi-directional translation. With these intermediate data representations, the programmer can specify visual mapping rules that define the mapping between ASR and VSR.

---

[2]A powerful vector graphics tool developed by Adobe Systems Incorporated.
(http://www.adobe.com/products/illustrator/main.html)
[3]An interactive 2-D drawing tool under X11 on various UNIX platforms.
(http://bourbon.usc.edu:8001/tgif/index.html)

The translation between ASR and VSR is the essential translation in the sense that it couples the abstract concept and the visual representation. Note that the other translations — translation between AR and ASR and translation between PR and VSR — are also important for completing the whole translation. They should be also supported in the development environment of the bi-directional translations. The former is related to the application interface of the user interface part developed by the system. The latter is the problem of geometric constraint solver and spatial parser of the development system. Their strength and usability have influence on the cost of developing direct manipulation interfaces.

3.4 Extended Model for Animation

Animating application behavior is important and useful for various areas, such as education, debugging, and performance tuning. In this section, we extend the basic model described in the previous sections to accommodate the creation of animations in the model.

Basic Idea Application programs compute on their internal state during their execution. That is, execution of an application is a sequence of operations and states. When an operation is executed, the application's state is changed, and goes to the next state in that sequence. Animations can also be viewed as a sequence of pictures. If the target of an animation is a program execution, each picture in the animation represents the application's internal state, and the picture changes represent operations in the execution.

We have extended the bi-directional translation model for the creation of animations as shown in Figure 3.5. The model is composed of a series of data for each representation. In addition, four types of "operations" are incorporated into each representation in the model. In Figure 3.5, the vertical sequence $\langle AR(t) \rightarrow ASR(t) \rightarrow VSR(t) \rightarrow PR(t)\rangle(0 \leq t \leq N)$ corresponds to the process of translating application data at time t to picture data, as described in the previous section. The time t is discrete, and starts from 0 when the application is first invoked. It is incremented when an operation is executed in the application. The time is N when the application is terminated.

On the other hand, the four horizontal sequences in Figure 3.5 represent changing data and operations on each representation. For example, the horizontal sequence at the top of Figure 3.5, $\langle ... \rightarrow AR(t) \rightarrow AR(t+1) \rightarrow AR(t+2) \rightarrow ...\rangle$, is a series of application data and operations. At time t, an operation is executed and the state is changed from $AR(t)$ to $AR(t+1)$. Then, another operation is executed and the state is changed to $AR(t+2)$. This sequence continues until the application is terminated. The other three horizontal sequences are similar to the sequence of $AR(t)$. In the sequence of $XR(t)$, an operation on XR changes $XR(t)$ to the next state $XR(t+1)$ where $XR = AR, ASR, VSR,$ or PR.

Specifying Transitions In the extended model, an animation corresponds to the horizontal sequence of $PR(t)$ and operations on PR. In fact, an operation on PR is the short animation $A(t)$, the steps of which are connected to form a whole animation A; i.e., $A = A(0) + A(1) + ... + A(N-1)$. The short animation $A(t)$ is generated by interpolating two successive pictures $PR(t)$ and $PR(t+1)$. Each picture $PR(t)$ can be generated by translating application data $AR(t)$.

In summary, using the extended model, we can create an animation by (a) obtaining the application's internal data during its execution, (b) translating the sequence of application data to a sequence of pictures, and (c) generating a sequence of short animations by interpolating the sequence of visualized pictures. As the translation from application data to a picture (i.e., (b)) is achieved only by specifying visual mapping rules, as described in Section 3.3, it is possible to specify an animation
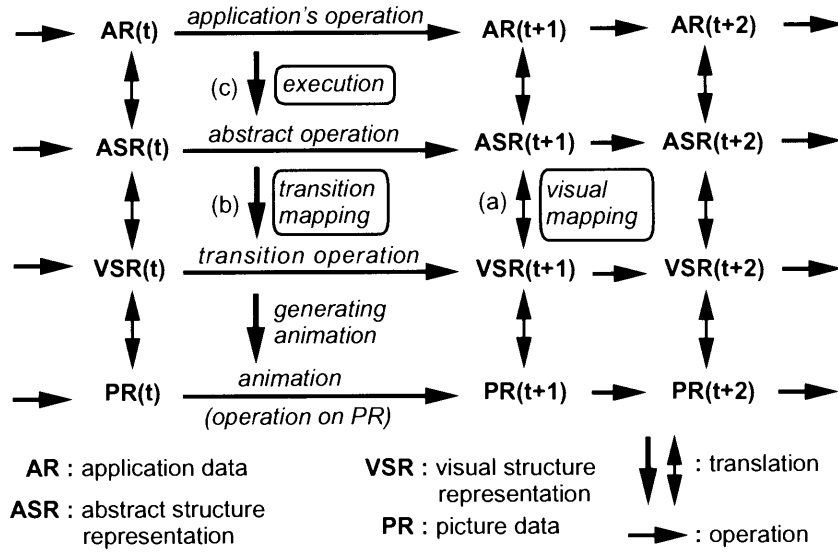
Figure 3.5: The extended bi-directional translation model that represents the general architecture of applications that generates animations of program execution.

using only visual mapping rules. However, since there are many possible ways to interpolate two pictures, it would be better to provide a method for their specification.

In our model, the method of interpolating pictures is regarded as a *transition operation* $to(t)$, which is an operation on VSR that changes $VSR(t)$ to $VSR(t + 1)$. This is a method of moving, scaling, and rotating the objects in one picture to those in another picture. In addition, a transition operation is translated from an application's operation via an *abstract operation* $ao(t)$ on ASR, which is defined as an operation that changes $ASR(t)$ into $ASR(t + 1)$. The translation from an abstract operation to a transition operation is specified by mapping rules, designated as *transition mapping rules*.

The process of translating application operations to an animation is illustrated in Figure 3.5. First, the application executes an operation; in our model, it is an operation on AR. Then, an abstract operation defined on ASR is invoked. As it is almost impossible to automatically recognize meaningful operations for animations from an application program, programmers themselves must define abstract operations; that is, they must insert code into the program to define abstract operations. Abstract operations correspond to *interesting events* in BALSA-II[15] or Zeus[16].

Then, abstract operations are translated to transition operations by transition mapping. As a transition operation is specified for one graphical object in VSR, some graphical objects in VSR are associated with transition operations specified by the mapping rules. Other objects are associated with default transition operations. Finally, the system generates a short animation by inbetweening the previous and the new picture according to the transition operations.

For example, suppose that the programmer is to make an animation that depicts the insertion sort algorithm. It shows the process of sorting bars of different lengths in ascending order from left to right. The insertion sort algorithm sorts bars by inserting a new bar into a row of bars already sorted. Figure 3.6 is an example animation that a bar is inserted at the appropriate place of the sorted row of bars. To move a bar in a circular motion, as in Figure 3.6, the following transition mapping rule is specified:

```
insert(X)  :- move(n(X), [clockwise]).
```
This mapping rule maps the operation insert (X) on ASR to the operation

move(X, [clockwise]) on VSR. Here, insert(X) represents an operation in the sorting program. X is a string or integer that represents the index corresponding to the inserted object in ASR, and n(X) is the name of the inserted graphical object in VSR[4]. The transition operation move(n(X), [clockwise]) specifies the method for moving the graphical object n(X) as clockwise. Thus, when insert(X) is executed, the graphical object named n(X) moves clockwise. The system moves the rest of the graphical objects, which are not governed by any transition operations, using the default transition operations. The default transition operation for moving a graphical object is move(Obj, [straight]), which moves the graphical object Obj in a straight line. In Figure 3.6, two small bars are shown moving straight to the left to make room for the bar being inserted.
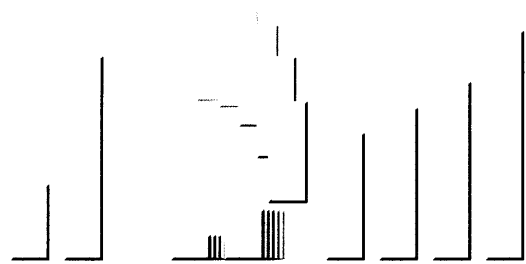
Figure 3.6: Circular movement of a bar in a sorting animation.

---

[4]The name of a graphical object is specified in a visual mapping rule.

Chapter 4

TRIP2 — A System For Constructing Direct Manipulation Interfaces

4.1 Introduction

Graphical and interactive user interfaces based on *direct manipulation*(DM)[109], allowing the users to point to, grasp, and drag objects on the screen, are now standard. To ease the high creation cost, we employ User Interface Management Systems (UIMSs) and user interface toolkits such as NextStep[95], MacApp[3], InterViews[113], GTK+[98], and Java Swing set[125]. However DM style interfaces are still costly, because UIMSs and toolkits usually only provide 'canned' abstractions of behaviors for a fixed set of simple graphical interface objects such as buttons and scrollbars. Such a fixed set of abstractions may be sufficient for simple, dialogue-style interfaces, but for cases where the semantics of complex application data must be reflected and be manipulatable on the display, appropriate graphical and input abstractions may not be available. This is primarily because current models of user interfaces are usually the models of the interaction architecture, and lack the support for consistent framework that allow visualization and manipulation of high-level abstract data, i.e., semantics of applications.

In this chapter, we show an implementation of a framework for bi-directional translation between application's data representations and pictorial representations of the user interface. With this framework, the feedback of user's manipulation can be generally achieved via inverse translation from picture data into abstract data. The specification of inverse translation is achieved with little extra cost, because (almost) identical rules can be used for mappings in both directions.

We have implemented TRIP2, a prototype system based on this model. In TRIP2, direct manipulation interfaces for various types of data can be achieved just by specifying declarative mapping rules.

4.2 The TRIP2 System

TRIP2 is a prototype system based on our bi-directional translation model, which is implemented on the NeXT computer. This section describes the implementation of TRIP2.

4.2.1 System Overview

Figure 4.1 shows an overview of the TRIP2 system. TRIP2 evolved from TRIP (TRanslation Into Pictures)[65, 67], which handled only one-way translation from abstract data into picture data. The
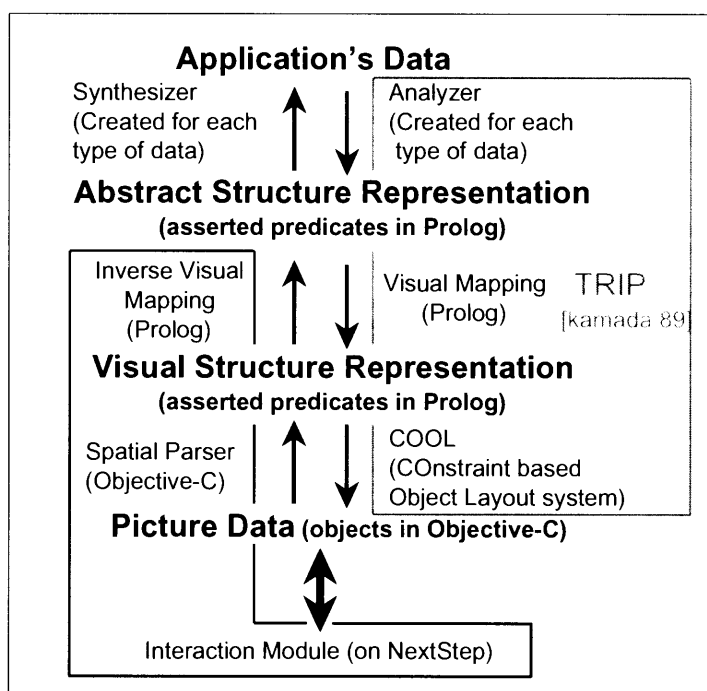
Figure 4.1: The architecture of the TRIP2 system.

new features of TRIP2 that realize our bi-directional translation model are as follows:

1. The **interaction module** which allows the users to manipulate the pictures directly with input devices such as a mouse. Currently, this module provides the users with an object-oriented freehand drawing interface similar to MacDraw[1]. This module is implemented on NextStep[2] by using Objective-C; thus, low-level graphical interactions, such as selection of a picture object with a mouse, are treated as method invocation of Objective-C objects that represent the picture.

2. The **spatial parser** which translates picture data into its visual structure representation. In TRIP2, picture objects are mapped directly to graphical objects in its visual structure representation, and this spatial parser extracts graphical relations from the picture data.

3. The **inverse visual mapping module** that translates visual structure representation into abstract structure representation. This translation is specified by inverse mapping rules given as a set of Prolog predicates.

In TRIP2, since the visual mapping and the inverse visual mapping are executed in Prolog[3], we are using Prolog's asserted predicates for ASR and VSR data. The spatial parser is integrated with the interaction module in the current implementation. The details of these modules are described in the following sections.

The users interact with TRIP2 in the following way:

1. Compile the mapping rules, and read them into the rule database.

---

[1] MacDraw is a trademark of Claris, Inc.
[2] NextStep is the operating system developed by NeXT Computer.
[3] SB-Prolog version 3.1, Copyright (C) 1986 at Stony Brook; 1987 University of Arizona.

2. Read the ASR data from a file into the *data* window.

3. Execute TRIP to translate the data into a picture.

4. Edit the picture in an object-oriented freehand manner.

5. Execute the *flush* command. This command first translates VSR data into ASR data, and then re-executes TRIP, i.e., translates the new ASR data into a picture. As a consequence, application semantics-directed beautification is achieved.

6. Return to 4.

Here we show an example in order to give some guidelines for using TRIP2 in Figure 4.2. Each screen dump of left-right windows shows the data of a Japanese family as ASR and its corresponding picture as a kinship diagram (Figure 4.2(a)). The left window is the *Data Window* which displays ASR data, and the right window is the *ObjView Window* which displays its pictorial representation. The user is free to modify either data at any time. (He interacts only with the pictures in the *ObjView Window* in this example.) Now, suppose that a new member, *Midori*, joins the family by marrying *Shinichi*. To express this fact, we draw a rectangle (Figure 4.2(b)), write the label *Midori*(Figure 4.2(c)), and connect it to the rectangle labeled with *Shinichi* by a line(Figure 4.2(d)). (Note that this label *Midori* was not produced by an explicit command to label the rectangle; rather, it is a textual picture object that has an equal status to the rectangles and lines just drawn.) Then, we execute the *flush* command. These modifications are translated back into ASR data, and the kinship diagram is redrawn from the new data (Figure 4.2(e)). Now a baby, *Tomomi*, is born: We draw a circle between two rectangles, draw a new rectangle, write the name, and connect the circle and the rectangle (Figure 4.2(f)). Finally, we execute the *flush* command again. These modifications are fed back, and the diagram is regenerated (Figure 4.2(g)).

4.2.2 The TRIP Module

The TRIP[65, 67] module is responsible for the translation from abstract data into concrete pictures. As illustrated in Figure 4.1, the TRIP module consists of three submodules (1) application's data analyzer, (2) the visual mapper, and (3) the layout system called COOL (COnstraint-based Object Layout system).

The analyzer translates application's data into ASR data, which is a set of Prolog predicates. In [67], examples of several types of application's data are presented, such as English sentences representing kinship relations, S-expressions in Lisp, and C programs. Those analyzers are not an integral part of TRIP2, but TRIP2 provides an editing window for ASR data as shown in Figure 4.2[4]. Thus, the users can interactively manipulate both ends of application-independent representations at the same time.

ASR data are then translated into VSR data by the visual mapper written in Prolog. ASR data and the mapping rules are read into the database, after which the visual mapper is invoked.

Finally, VSR data are translated into PR data by COOL. In COOL, VSR data are declared as special predicates that generate (a) drawing data, and (b) constraints on the geometric attributes in the drawing data. For example, the predicate box generates (a) data for drawing a rectangle, and (b) constraints on coordinate variables of the rectangle. The predicate horizontal generates (b) constraints on y-coordinate variables of picture objects corresponding to the graphical objects listed

---

[4]In fact, the original TRIP system was interfaced with a simple natural language parser; as a result, the user could input sentences describing a scenery, and TRIP would automatically visualize it[67]. This parser can easily be interfaced with TRIP2 as well.
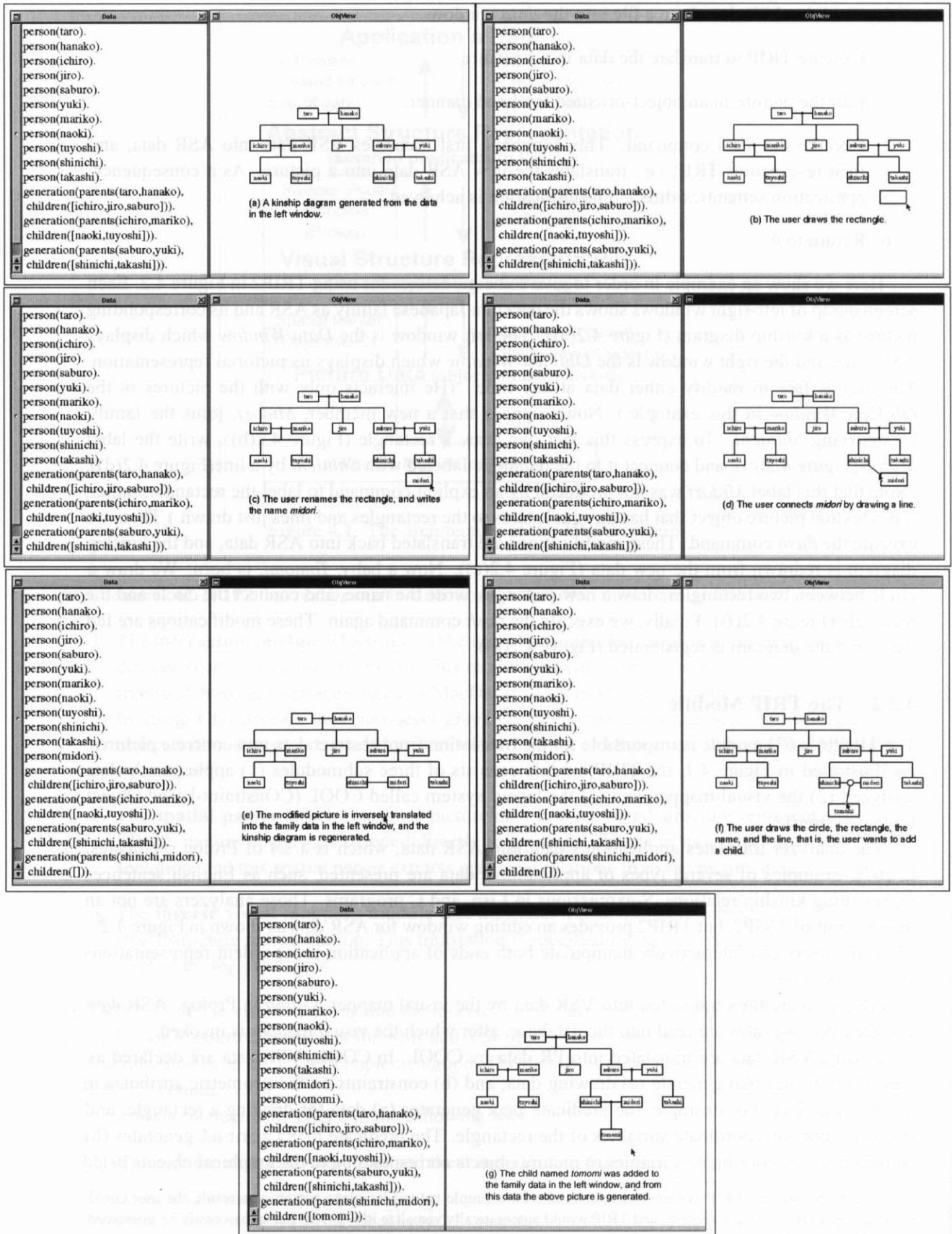
Figure 4.2: A kinship diagram.

in its argument. These constraints are solved by the constraint solvers that are parts of COOL. The constraint solvers can handle two types of constraints: 'rigid' constraints, which must be solved exactly, and 'pliable' constraints, which may be solved approximately by the least square method. The purpose of pliable constraints is to allow COOL to handle over-constrained systems. In addition, COOL has special predicates and another constraint solver for visualizing undirected graph structures as balanced graphs. The constraint solving algorithms of COOL are described in [65, 67]. The solutions computed by the constraint solvers instantiate the coordinate variables in the drawing data. They are then passed to the interaction module to generate the Objective-C picture objects.

For example, given the ASR data:

```
item('AR').
item('ASR').
item('VSR').
item('PR').
order(['AR','ASR','VSR','PR']).
```

and the visual mapping rules (Figure 4.3), the VSR data in Figure 4.4 are generated.

```
objectmap([itemmapping]).
relationmap([ordermapping]).
itemmapping :-
      item(NAME),                        % gets 'item' data from database
      itemmap(NAME),                     % maps 'item' data
      fail.                              % iterates to the next data
itemmap(NAME) :-                         % item(NAME) is mapped to a box with a la-
bel
      box(NAME, 200, 100, []),           % box (width=200, height=100)
      label(l(NAME), NAME, []),          % label 'NAME'.
      contain(NAME, l(NAME), 0, []).     % a box contains a label

ordermapping :-
      order(ILIST),                      % gets 'order' data from database
      ordermap(ILIST).                   % maps 'order' data
ordermap(ILIST) :-                       % items are horizontally arranged.
      horizontallisting(ILIST, 20, []).
```

Figure 4.3: Visual mapping rules example.

To execute the visual mapping, TRIP calls the predicates listed in the argument of predicates `objectmap` and `relationmap`. The predicates of the `objectmap` (`itemmaping`, in Figure 4.3) maps abstract objects to graphical objects, and the predicates of the `relationmap` (`ordermapping`) maps abstract relations to graphical relations. When the predicate `itemmaping` is called by the system (Figure 4.5 (1)), `item(NAME)` gets one ASR data (`item('AR')`) with unification (Figure 4.5 (2)). Then `itemmap(NAME)` unifies the head of the other rule (Figure 4.5 (3)), and the three VSR predicates output VSR data, `box`, `label`, and `contain` (Figure 4.5 (4)). The special predicate `fail` forces backtracking, so this process is repeated with the other `item` until all `item`s are mapped to VSR. Similarly, the predicate `ordermapping` applies the predicate `ordermap` to the abstract relation `order(ILIST)`, and it maps `order(ILIST)` to graphical relation `horizontallisting`. Thus the 'item's are layed out horizontally. Next, COOL is invoked and the PR shown in Figure 4.6 is generated from the VSR.

% Graphical Objects and Relations mapped from 'item'

```
box('AR', 200,100,[]). label(l('AR'), 'AR', []).
contain('AR', l('AR'), 0,[]).
box('ASR',200,100,[]). label(l('ASR'),'ASR',[]).
contain('ASR',l('ASR'),0,[]).
box('VSR',200,100,[]). label(l('VSR'),'VSR',[]).
contain('VSR',l('VSR'),0,[]).
box('PR', 200,100,[]). label(l('PR'), 'PR', []).
contain('PR', l('PR'), 0,[]).
```
% Graphical Relation mapped from 'order'
```
horizontallisting(['AR','ASR','VSR','PR'], 20, []).
```

Figure 4.4: Generated VSR data.

4.2.3 The Interaction Module

The users of TRIP2 manipulate pictures directly via the interaction module, which provides the generic lowest-level editing facility. They can create, move, and delete picture objects such as rectangles, circles, etc. We stress that the resulting interface is not a structured editor for some specific data — rather, the users simply draw rectangles or circles as if they were using object-oriented freehand drawing systems such as MacDraw. This is unlike graphical editors which are customized for each application, such as Unidraw[124].

A picture in the interaction module is represented as a set of picture objects in Objective-C. Picture objects are the instances of the subclasses of the Shape class (see 4.6), in which the usual generic methods, such as drawing, moving, changing the size/place of the object, getting the attributes of the object, and testing whether a given point is within the object, are defined. Several coordinates describing the picture objects, such as the center (cx, cy), the top left corner (lx, ty), and the bottom right corner (rx, by), are declared as instance variables of the Shape class. These coordinates are used to test whether geometric relations among the objects hold. For example, when two objects are tested to be horizontal in the *top_align* mode, the *ty* coordinates of these objects are compared.

As mentioned before, VSR data is a set of graphical relations among graphical objects expressed as Prolog predicates. However, since the translation between PR and VSR is executed in the interaction module written in Objective-C, the interaction module must also have the data corresponding to VSR. Therefore, the graphical objects in VSR are mapped idempotently to the picture objects in the interaction module, i.e., the picture objects serve as the dual role of also being the VSR graphical objects. Graphical relations in VSR are represented as a list of Objective-C graphical relation objects, which are the instances of the subclasses of the Relation class (see 4.6), which supports the following methods:

- A method for calculating the error of a graphical relation. This method is used for testing whether the graphical relation holds or not.

- A method for testing whether a graphical object is 'addable' to a graphical relation, i.e., testing whether the graphical relation holds when the graphical object is added. This method is used for translating PR data into VSR data.
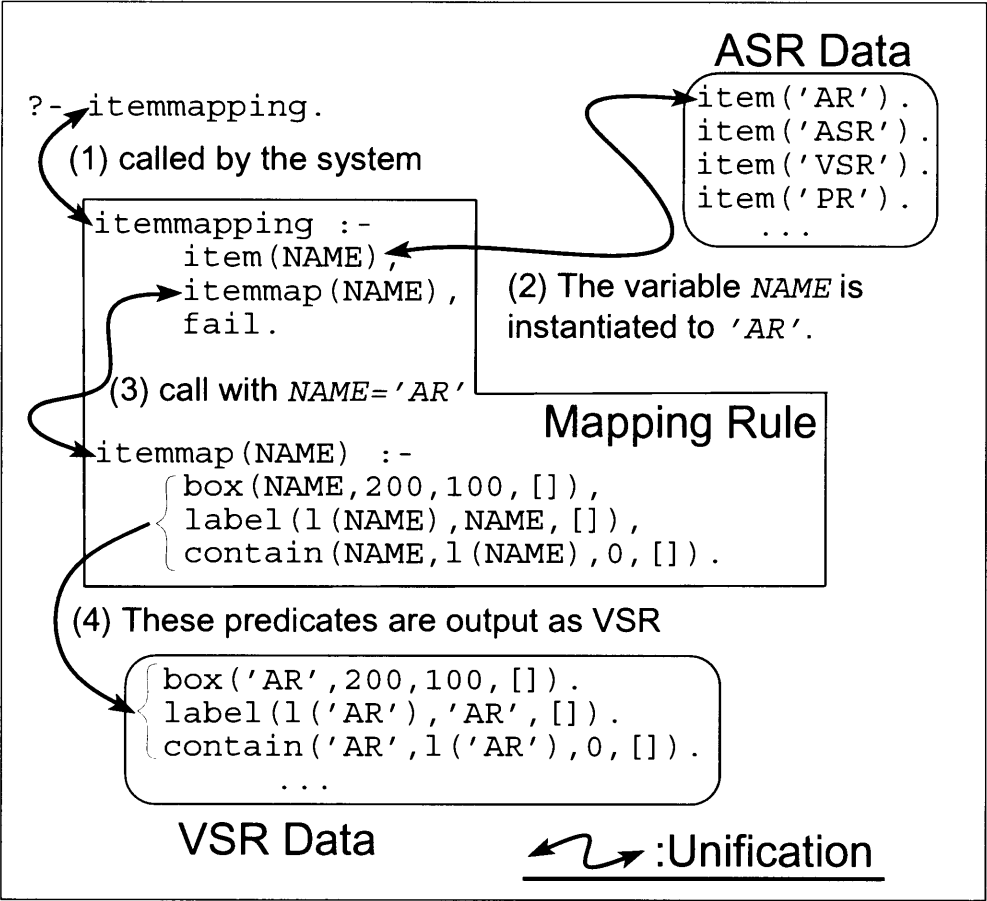
Figure 4.5: The execution of mapping rule.

Figure 4.6: The generated picture.

4.2.4   Implementation of the Inverse Translation

The inverse translation involves translation from PR into VSR, followed by translation from VSR into ASR.

Translation from PR into VSR — Spatial Parsing   Here, we refer to translation from PR into VSR as *spatial parsing*. Although spatial parsing has been previously studied (for example, [45, 75, 74]), its application to GUI poses difficulties due to the requirement of real-time use, because there are a great number of graphical relations with no a priori information in a picture.

In our model, spatial parsing has a more restricted role; it extracts a set of graphical relations among graphical objects. Since pictorial objects are recognized as graphical objects in the interaction module, the spatial parser in TRIP2 needs only to extract the graphical relations.

To realize real-time responsiveness, we use an *incremental* spatial parser, which retains most of the graphical relations and updates only its modified portions. Our incremental spatial parser is invoked in the following situations:

- **When a picture object is deleted:** The graphical object (in VSR) corresponding to the deleted picture object (in PR) is removed. In addition, it is removed from each (existing) graphical relation containing it. For example, when the rectangle with a label 'PR' in Figure 4.6 is deleted, 'PR' is removed from the `horizontallisting` relation, that is,

  ```
  horizontallisting(['AR','ASR','VSR','PR'], 20, [])
  ```

  is modified to

  ```
  horizontallisting(['AR','ASR','VSR'], 20, []).
  ```

  To achieve such modifications efficiently, the interaction module maintains a table that maps each graphical object to a list of graphical relations containing it.

- **When a new picture object (obj) is created:** The message `addWithTest:obj` is sent to each existing graphical relation object. The invoked method tests whether the `obj` is 'addable' to the graphical relation, and if so, adds it to the graphical relation.

  For example, consider the situation depicted in Figure 4.7. In Figure 4.7, four boxes are constrained by the graphical relation `x_order([a,b,c,d],10)` so that they are placed at even intervals[5]. When a user draws a new rectangle with a label e on the right of rectangle d, `addWithTest:e` message is sent to the graphical relation object, `x_order([a,b,c,d],10)`. That is, the system checks whether e can be added to the graphical relation or not. In this case, the distance w between d and e is nearly equal to 10; as a result, the new rectangle object is added to the `x_order` relation. The spatial parser of our system permits a certain amount of error, because the user usually cannot put new objects at the exact place that satisfy existing graphical relations.

- **When a picture object is moved:** The spatial parser is also invoked when a picture object is moved; this change is treated as a combination of deletion and addition.

---

[5]It constrains only each box's x-coordinate. Their y-coordinates do not have any meaning in this figure.
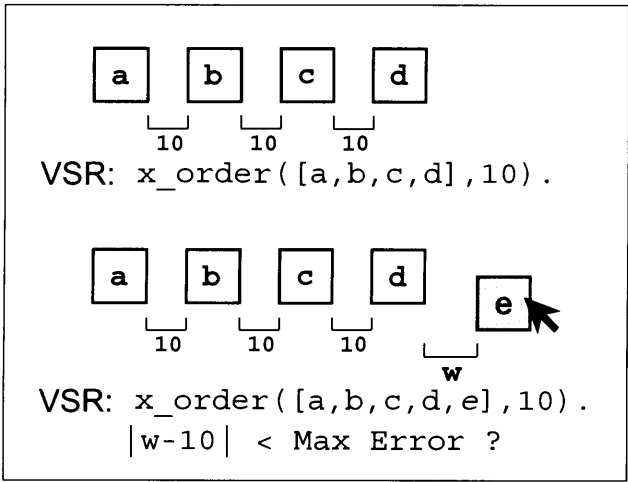
Figure 4.7: Adding a new graphical object to a relation.
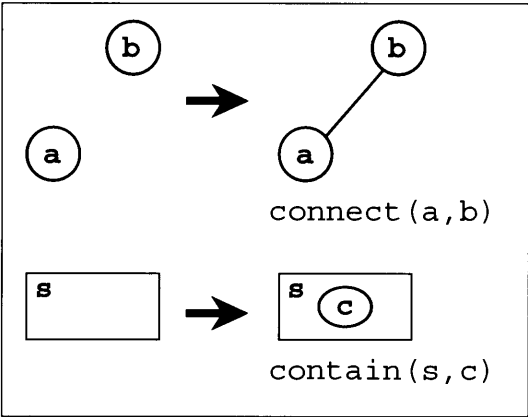
Figure 4.8: Creating a new graphical relation.

- **Adding a new graphical relation:** Although the above scheme can handle the modification of existing graphical relations with an arbitrary number of graphical objects, the shortcoming is that it cannot handle the creation of new graphical relations. For example, the `connect` relation, which represents two graphical objects connected by a line, must be created when a user draws a line that connects two picture objects that were previously unrelated. We alleviate this by associating generations of new graphical relations only with drawings of the picture objects, that is, when a certain kind of an object is drawn, the system checks whether the object has established new relations or not. For example, as shown in Figure 4.8, when a line connecting two objects is drawn, a `connect` relation is created, and when an object is drawn within the boundary of another object, a `contain` relation is created. We are now working to provide alternative means of generating other types of graphical relations.

In TRIP2, all graphical relations are inferred by the spatial parser. One problem of this method is that users cannot see that the system is inferring correctly or not. It is preferable that inferred constraints are displayed visually. In addition, there may be cases that explicit specification of constraints is preferable. These problem has been solved by our next system TRIP3[88].

Translation from VSR into ASR   Translation from VSR into ASR (inverse visual mapping) is
executed in Prolog. The interaction module outputs VSR data to a Unix pipe, which are read into
the Prolog database. VSR data are translated into ASR data according to the inverse mapping rules.

Ideally, we should be able to use identical rules for inverse visual mapping as well as for visual
mapping, because declarative rules do not have 'directional' information. However, this is not the
case due to the following reason: If some rule is used naively for mappings in both directions, the
pictorial editing performed by the user must exactly satisfy the layout constraints imposed by the
visual mapping rules. The resulting interface would be extremely rigid and difficult to use — for the
example, in Figure 4.3, the user would be required to draw a rectangle of the exact size as specified
by the mapping rule ($width = 200, height = 100$) for his input to be recognized.

Instead, we must relax the constraints so that system would be tolerant of the non well-formed
operations of the user. In TRIP2, this is currently achieved by having the user provide a separate
set of corresponding inverse visual mapping rules which would be 'less specific' compared to its
visual mapping counterparts. For example, in Figure 4.9, the term box has several of its variables
unspecified (this is indicated by underscores ('_'), which serves as *anonymous variables* in Prolog).
As a result, the height and the width of the user-drawn box are ignored, allowing the user to draw
a box of arbitrary size[6]. Fortunately, the structure of the mapping rules are essentially identical;
thus, an experienced user will be able to derive one from the other with ease. Moreover, automatic
derivation of inverse mapping rules from the other mapping rules can be possible.

```
invomap([inv_itemmapping]).
invrmap([inv_ordermapping]).
inv_itemmapping :-
      inv_itemmap(NAME),          %finds 'item' data
      asr(item(NAME)),            % outputs 'item' data
      fail.                       % tries for another data

inv_itemmap(NAME) :-             % 'item' data is a
      box(A, _, _, _),           % box with its name 'A'
      label(B, NAME, _),         % and the label 'B' whose string is 'NAME'
      contain(A, B, _, _).       % box 'A' contains label 'B'

inv_ordermapping :-
      inv_ordermap(ILIST),        %finds 'order' data
      asr(order(ILIST)),          % outputs 'order' data

inv_ordermap(ILIST) :-           % 'order' data is a
      horizontallisting(ILIST, _, _).  % horizontal ordering
```

Figure 4.9: Inverse visual mapping rules example.

For example, the inverse visual mapping rules that correspond to the visual mapping rules in
Figure 4.3 are shown in Figure 4.9. They are executed as follows.

1. The predicate inv_itemmaping is called (Figure 4.10 (1)).

2. The term inv_itemmap(NAME) is unified with the head of another rule (Figure 4.10 (2)).

---

[6]This is one of the reasons why we are able to first do the inverse mapping, and then perform the visualization again
to achieve beautification; the beautified PR is in effect the 'exact' visualization of a given ASR data, whereas the edited
picture just prior to beautification is not.

```
                      ?- inv itemmapping.

   (1)called by the system          inverse mapping rules

                         inv itemmapping :-
                            inv itemmap(NAME),
                            asr(item(NAME))
   (2)call with NAME unbounded    fail.                   (4)NAME is bound to 'AR'.


                         inv itemmap(NAME):-
                          { box(A, _, _, _),                (5)asr() outputs
   (3)The variables A,B,NAME are    label(B, NAME, _),       item('AR') as ASR data.
   instantiated to oid1,oid2,       contain(A, B, _, _)
   'AR', respectively.
                                                 VSR data
                          { box(oid1,200,100,[]).                 ASR data
                            label(oid2,'AR',[]).            { item('AR')
                            contain(oid1,oid2,0,[])

         : unification
```

Figure 4.10: The execution of inverse mapping rules that translate a set of of VSR data (box, label, and contain) into ASR data (item).

3. The three terms (box, label, and contain) are unified with a set of VSR data in database, and the variables A, B, and NAME are instantiated (Figure 4.10 (3)).

4. The variable NAME is bound to 'AR' (Figure 4.10 (4)).

5. The predicate asr/1 generates item('AR') (Figure 4.10 (5)).

Since the predicate fail forces backtracking, all combinations of VSR data (box, label, and contain) are tested whether they match with the body of the rule, thus all ASR data are generated by the inverse visual mapping. Similarly, order(ILIST) is derived from the horizontallisting relation in the VSR database, and output in a certain format. Compared to the visual mapping rules (Figure 4.3), we can see that each clause corresponds one-to-one, and the terms in the clause body and their ordering are almost the same. Derivation of other inverse mapping rules is similar.

Often diagram interpretation involves a multiplicity of interpretations. In TRIP2, diagram is interpreted with inverse mapping rules. Thus the problem of multiple interpretations can be solved by writing the rules carefully. When the mapping rule allows multiple interpretations, the ASR data of all interpretations are usually generated, because most inverse mapping rules search with backtracking like the rules in Figure 4.9.

As mentioned earlier, the predicate inv_itemmap has essentially the same structure as itemmap, but the width and height of the box are ignored. If one desires to restrict the size of boxes within a certain range, only the predicates that specify the range of the width and height need to be added as follows:

```
inv_itemmap(NAME) :-
    box(A, W, H, _), W < 300, H < 200,
    label(B, NAME, _),
    contain(A, B, _,  _).
```

4.3   Examples

In this section, we show the examples of TRIP2 in action. All the figures in this section show the screen dumps of two windows of TRIP2. In the following examples, we focus on the feedback of user's operations on the PR data. Appendix C shows visual and inverse visual mapping rules for the examples in this section.

4.3.1   A Simple Graph Editor

This example is a simple graph editor on TRIP2. Appeindix C.1 shows the visual mapping rules for this example: there, nodes are mapped to circles, and edges are represented as straight lines. The user of this editor simply draws circles, lines, and labels. These pictures are recognized as a graph, and translated into nodes and edges. The nodes in this example are arranged by the adjacent relation which is an entry to the graph layout constraint solver in TRIP2. The graph drawing algorithm used here is described in [65, 66].
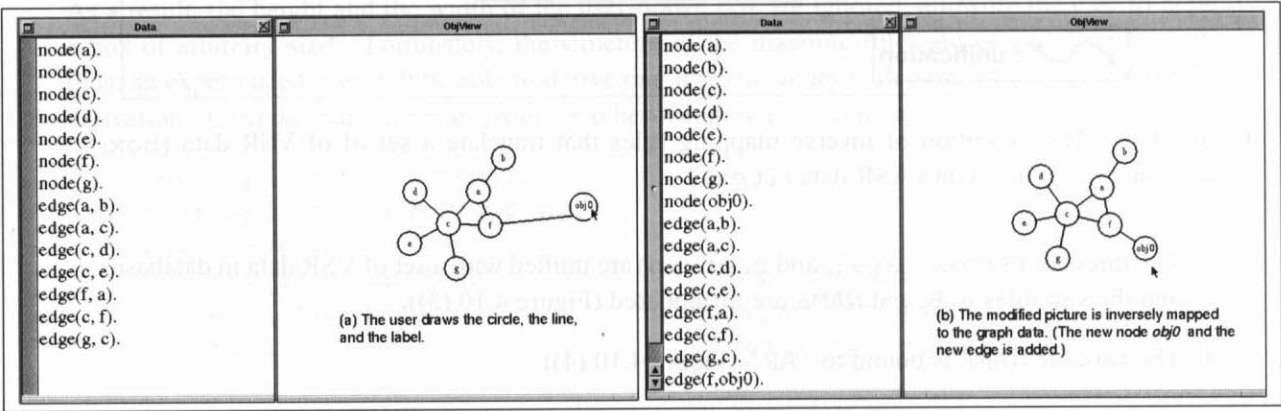
Figure 4.11: A simple graph editor.

4.3.2   A Simple E-R Diagram Editor.

Figure 4.12 shows the screen snapshots of the windows of an Entity-Relationship Diagram editor. The upper-left window shows the E-R database schema in a textual form, and the right window displays its E-R diagram. The user extends the schema by adding a new entity *course* (Figure 4.12(b)) with two attributes (*number* and *name*) (Figure 4.12(d)), and new relationships (*S-C* and *I-C*) (Figure 4.12(c)). The user then performs the *flush* command. The modified picture in the right window is translated back to the E-R schema data in ASR, and these data are translated into the corresponding beautified E-R diagram (Figure 4.12(e)).

In this example, entities (represented by rectangles), and attributes (represented by ellipses) are arranged by the graph layout constraint solver as is with the previous graph editor example.

4.3.3   A Small Othello Game

This example provides an interface for a small Othello game application. Figure 4.13 shows the screen dumps of the windows during the game, which display the board data (ASR) and its pictorial representation. The Othello application itself is written in Prolog, and uses this ASR data directly as its representation of the board (therefore, there is no translation process between AR and ASR). It

(a) The E-R database schema data in the left window
are translated into the picture in the right window.

(b) The user draws an *entity* (a rectangle
with a label).

(c) Two *relationships* (diamonds) are added by the user.

(d) Two *attributes* (ellipses) are added.

(e) The user executes the *flush* command. The updated
schema is generated, and the E-R diagram is beautified.

Figure 4.12: A simple E-R diagram editor.

has no knowledge of the interface of the game. The board data consists of six rows of six characters (Figure 4.13(a)). The character 'e' means that the position is empty. The characters 'b' and 'w' represent black and white stones respectively.

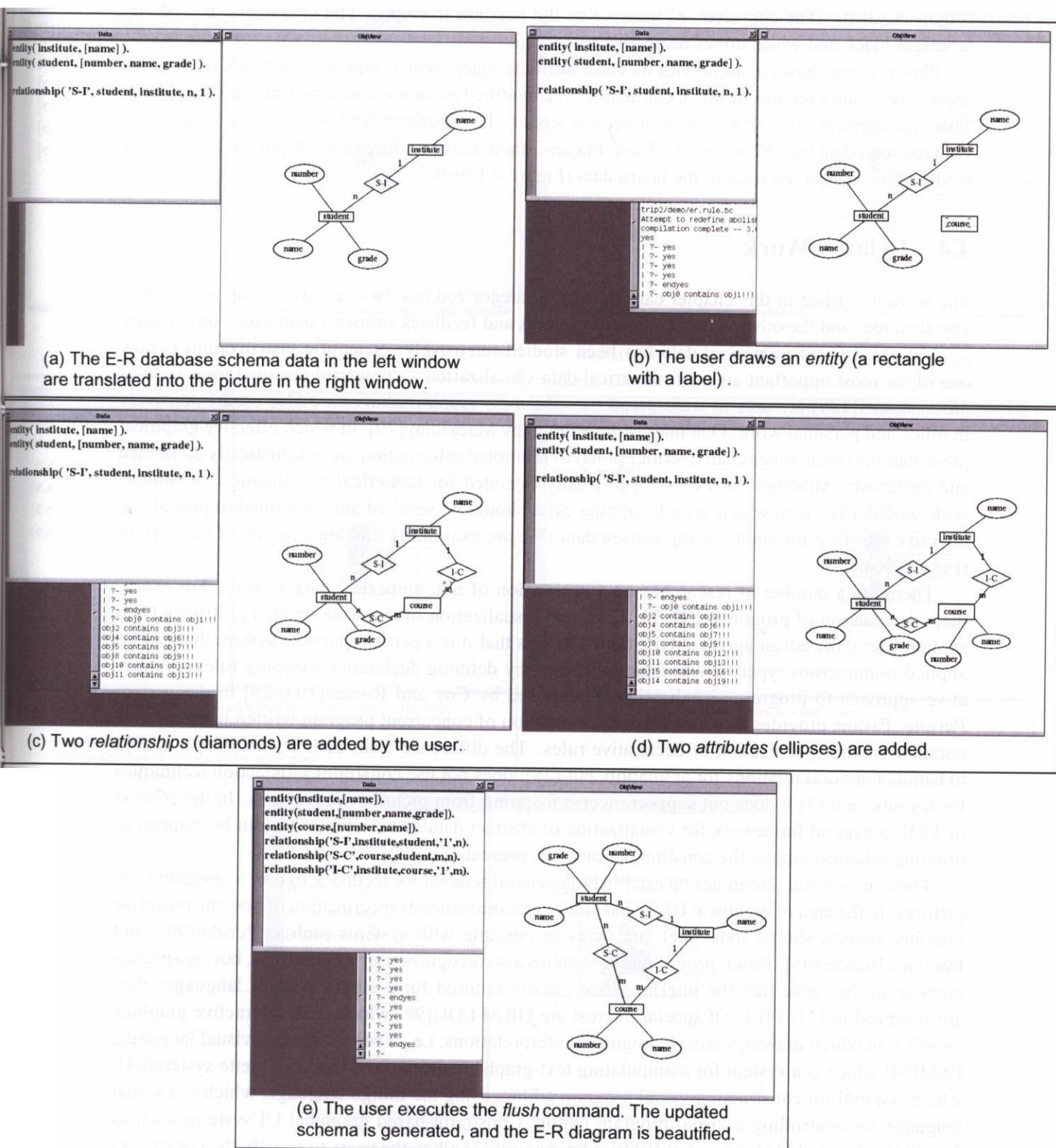Player-1 first draws a circle with its color attribute black, which represents a black stone (Figure 4.13(b)), and executes the *flush* command. The modified picture is translated into the board data. Then, the application reverses the intermediate stones. The resultant ASR of the board is again visualized to update the PR (Figure 4.13(c)). Player-2 then draws a white stone (Figure 4.13(d)). This modification is also fed back to the board data (Figure 4.13(e)).

4.4   Related Work

The work described in this chapter can be largely categorized into two research areas: one is data visualization, and the other is recognition of pictures and feedback of user's operations on pictures.

Visualization of *numerical* data has been studied intensively: Scientific visualization is now one of the most important areas of numerical data visualization — powerful personal tools such as *Mathematica*[129] are now in widespread use. Business graph packages are also used extensively in office and personal work. One interesting work is by Mackinlay[79], in which effective graphical presentations (such as bar charts, scatter plots) of relational information are automatically combined and designed. Although TRIP2 is not primarily intended for numerical visualization[^7], a similar issue would arise as to which visual mapping rules should be selected and combined to provide an effective interface for complex application data. We are examining this topic as one of the current research issues.

There are a number of researches on visualization of non-numerical data as well. For example, visualization of programs [70, 93, 123], and visualization of database schema [130] has been studied. Here, the advantage of TRIP2 (and TRIP) is that it is a general-purpose system that can be applied to numerous types of visualization simply by defining declarative mapping rules. Declarative approach to program visualization is proposed by Cox and Roman[103, 28] in the system *Pavane*. Pavane provides three-dimensional animation of concurrent program written in the *Swarm* notation via the specification of declarative rules. The difference from TRIP2 is that (1) it is able to handle temporal relations for animation, but (2) it does not use constraint satisfaction techniques for layouts, and (3) it does not support inverse mapping from pictures to programs. In the context of TRIP, a general framework for visualization of abstract data and relations that can be mapped to ordering relations among the constituent objects is presented in [92].

There are several researches on establishing general schema for feedback of user's operations on pictures. In the area of graphical UI, visual and/or demonstrational specification of how the primitive interface objects should look, feel, and work is possible with systems such as Peridot[90], and Interface Builder[95]. *Visual programming* systems also recognize user's operations, but are special-purpose in the sense that the interfaces are usually tailored for a single (visual) language; they are surveyed in [21, 110]. Of special interest are GRAFLOG[99] which is an interactive graphics interface in which drawings receive linguistic interpretations, i.e., a certain kind of visual language, PAM[74] which is a system for manipulating text-graphic patterns, and Golin's Palette system[44], which is a tool for constructing visual program editors using the Indigo language, which is a visual language for controlling a visual program editor. Constraint-based graphical UI systems such as Juno[94], ThingLab[11], ThingLabII[80], Metamouse[32] allow the users to specify the constraints among the picture objects, but do not explicitly support the specifications of mappings between application and picture data. Advanced UIMSs such as Garnet[91] and Unidraw[124] provide the

[^7]: It is of course possible and useful to apply TRIP2 to visualization of numerical data.

(a) The board-data in the left window are
mapped to the picture in the right window.

(b) The user draws a black circle. (note that certain margins are
allowed for the size and the position of the piece the user draws.)

(c) The modified picture is inversely mapped to the board-data
and then the middle white stone is reversed by the Othello
application written in Prolog. The picture in the right window is
generated from the data in the application.

(d) The user draws a white circle.

(e) The inverse mapping has been executed, and
the intermediate black stone is reversed. This data
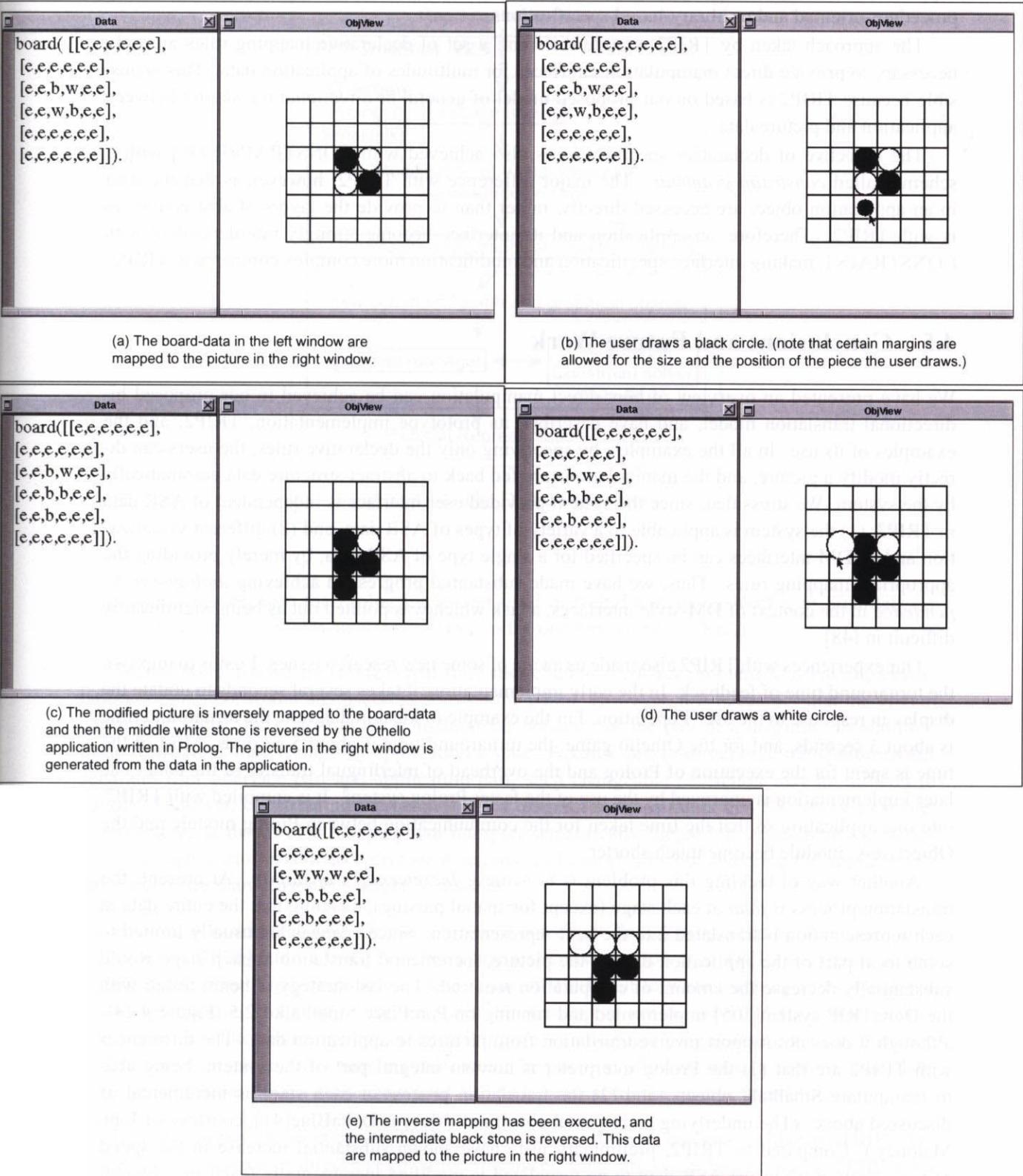are mapped to the picture in the right window.

Figure 4.13: TRIP2 othello.

support for visual interfaces which reflect application semantics, but the schema includes more procedure-oriented and/or library-based specifications.

The approach taken by TRIP2 is quite different; a set of declarative mapping rules are only necessary to provide direct manipulation interfaces for multitudes of application data. This is possible because TRIP2 is based on our proposed model of general bi-directional translation between application and picture data.

The objective of declarative specification is also achieved with CONSTRAINT[131] with a schema called *constraint grammar*. The major difference with TRIP2, however, is that the data in an application object are accessed directly, rather than to provide the layers of abstractions as is with TRIP2. Therefore, an application and its interface become strongly interdependent with CONSTRAINT, making interface specification and modification more complex compared to TRIP2.

4.5   Conclusions and Future Work

We have presented an overview of how direct manipulation can be achieved in our proposed bi-directional translation model, and have described its prototype implementation, TRIP2, and the examples of its use. In all the examples, by specifying only the declarative rules, the users can directly modify a picture, and the manipulations are fed back to abstract structure data automatically by the system. We stress that, since the system-provided user interface is independent of ASR data in TRIP2, (1) the system is applicable to a variety of types of ASR data, and (2) different visualization and its DM interfaces can be specified for a single type of ASR data, by merely providing the appropriate mapping rules. Thus, we have made substantial progress in achieving *dialogue independence* in the context of DM-style interfaces, a task which was pointed out as being significantly difficult in [48].

Our experiences with TRIP2 also made us aware of some new research issues. First is to improve the turnaround time of feedback. In the early implementation, it takes several seconds to update the display in response to the user's operation. For the example of kinship diagram, the turnaround time is about 3 seconds, and for the Othello game, the turnaround time is about 5 seconds. Most of the time is spent for the execution of Prolog and the overhead of interlingual transfer of the data. The later implementation is improved by the use of the faster Prolog system[8]. It is compiled with TRIP2 into one application so that the time taken for the communication between Prolog module and the Objective-C module become much shorter.

Another way of tackling this problem is to achieve *Incremental translation*. At present, the translation process is *total* at each stage (except for spatial parsing), meaning that the entire data in each representation is translated into the next representation. Since changes are usually limited to some local part of the application data or the picture, incremental translation at each stage would substantially decrease the amount of computation required. The last strategy is being tested with the DeltaTRIP system[105] implemented and running on ParcPlace Smalltalk v2.5 (Figure 4.14), although it does not support inverse translation from pictures to application data. The differences with TRIP2 are that (1) the Prolog interpreter is now an integral part of the system, being able to manipulate Smalltalk objects, and (2) the translation process at each stage is incremental as discussed above. (The underlying incremental constraint solver is DeltaBlue[41], courtesy of John Maloney.) Compared to TRIP2, preliminary tests have shown substantial increase in the speed of visualization when the ASR data to be visualized is modified incrementally (each incremental translation process takes less than a second). The speed improvement achieved in DeltaTRIP is
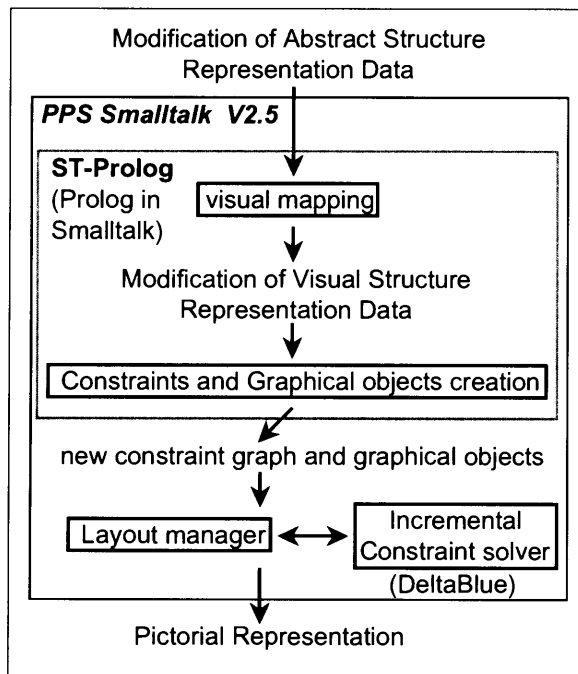
---

[8]We used SICStus prolog[111].
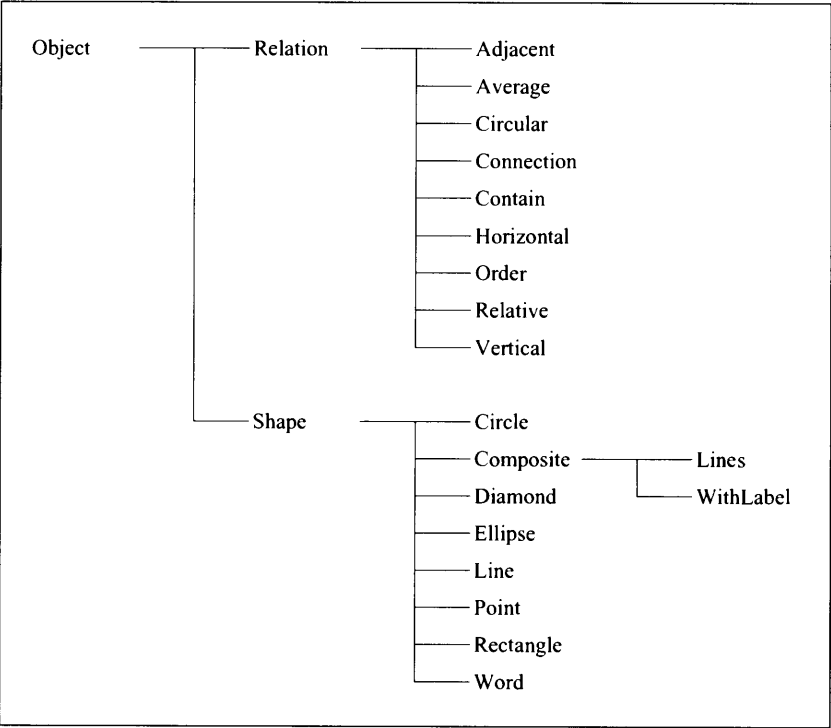
Figure 4.14: The DeltaTRIP system.

very promising, because the Prolog interpreter used in DeltaTRIP, being written in Smalltalk, is even an order of magnitudes slower compared to the one employed in TRIP2.

Second issue that we are facing is real-time and semantic feedback problem. Current implementation does not constraint any users' actions on the editor. However, it is necessary that the system can restrict some users' actions, or constraint the behavior of the part of a diagram. In addition, these restrictions may change by application's state. Because our bi-directional translation model separates user interface part and application part, such changeable restriction may be difficult to handle. However, constraining the behavior of a diagram by constraints in VSR can be possible. Furthermore, as the inverse mapping can interpret semantically illegal but syntactically legal actions, the application is able to detect such actions, and cancel it by not changing the application's data. There are many cases that might be enough.

Another way of extending TRIP2 is to use a constraint (logic) programming language, such as PROLOG III[25] and CLP(R)[59]. The numerical constraint solver in TRIP2 is external to the Prolog in the current implementation. As a result, it is difficult to write mapping rules which the high-level mapping rules are re-applied when constraints can not be solved in the lower translation stages. By integrating the constraint solving stages into logic programming with a constraint programming language, such mapping rules would be possible, allowing specification of more elaborate visualization. Such a language would also be beneficial for the inverse translation.

It is also important to integrate TRIP2 with existing UIMS or toolkits. Although TRIP2 has realized easy creation of direct manipulation interfaces, there are many functions that TRIP2 does not support. For example, TRIP2 does not intended to support ordinary widgets, such as menus, scroll-bars, and buttons. Business graphs and scientific visualization are not within the scope of TRIP2, too. TRIP2 must be integrated as one of the tools for creating user interfaces.

4.6   Classes in TRIP2

```
Object ──────┬──── Relation ──────┬──── Adjacent
             │                    ├──── Average
             │                    ├──── Circular
             │                    ├──── Connection
             │                    ├──── Contain
             │                    ├──── Horizontal
             │                    ├──── Order
             │                    ├──── Relative
             │                    └──── Vertical
             │
             └──── Shape ─────────┬──── Circle
                                  ├──── Composite ──┬──── Lines
                                  ├──── Diamond     └──── WithLabel
                                  ├──── Ellipse
                                  ├──── Line
                                  ├──── Point
                                  ├──── Rectangle
                                  └──── Word
```

Classes for graphical relations are derived from the *Relation* class, and classes for picture objects are derived from the *Shape* class. Picture objects (in PR) are also used as graphical objects (in VSR) in the interaction module.

Chapter 5

TRIP2a — Constructing Algorithm Animations

5.1 Introduction

Animation is an effective technique for use in a graphical user interface (GUI), and is useful for indicating various changes, processes, and movement of data or execution of application programs. Animation is widely used in entertainment, with video games and cartoons making effective use of computer animation. The results of computations in various scientific fields, such as simulations of fluid mechanics, are often visualized and animated. Algorithm animation, which illustrates how a target algorithm works, is a useful tool for teaching algorithms, as well as for creating and debugging algorithms.

Nevertheless, these advantages are not fully utilized in current GUIs, because of the time and cost associated with creating animations despite the many tools available. The primary reason for the time and cost is that the process of creating animations is still procedural, and the programmer has to be concerned with various details of object movements, etc. As a result, although computer and graphics performance have improved dramatically, we cannot fully utilize this performance for visualization and animation in GUIs.

This chapter describes a tool for creating animations in a fully declarative framework to reduce the cost of creating animations and to facilitate their use. However, we do not attempt to cover all conceivable kinds of animations here. Our target is *abstract* animations; that is, we do not intend to deal with animations that illustrate "real" shapes of some real objects, nor the movements based on certain physical phenomena. For example, animations that represent the results of simulations of hydromechanics, animations of the human body, or animations of dinosaurs are outside the scope of this thesis. To be more precise, we will discuss animations that describe the process of changing abstract data or relations, such as algorithm animations or program visualizations.

Our target animations have two features in common:

1. The abstract data and the relations to be animated do not have their own intrinsic "shapes." How to visualize and animate them should be specified in some way so that they can be understood easily. This is in marked contrast to animations of real objects, which have their own shapes, and usually images of such objects should resemble their real shapes.

2. The abstract data are stored within the application program, and change during its execution. To create an animation of an application program, we have to obtain up-to-date information regarding the status of the application program during its execution.

TRIP2 described in Chapter 4 enables programmers to easily create direct manipulation-style interfaces for manipulating abstract data. In this model, visualization of application data and recognition of pictures are regarded as translations from/to application data to/from pictorial data via abstract structure representation (ASR) and visual structure representation (VSR). This translation is specified by a declarative mapping rule between ASR and VSR (*Visual Mapping Rule*). We have extended this bi-directional translation model to handle animations. The extended model includes the notion of time; i.e., it incorporates a series of changing data. Animations are achieved by interpolating a series of pictures *translated* from a series of application data. In addition, we define *operations* on four data representations in our model. These operations change their data to the next state. The picture changes, animations, are also regarded as operations on pictures, and how an animation works can be altered by providing a *transition mapping rule* that maps *abstract operation* on ASR to *transitional operation* on VSR. Using this model, only two sets of mapping rules, i.e., visual mapping rules and transitional mapping rules, are necessary to create an animation. Complex animation techniques, such as slow-in-slow-out and squash-and-stretch[76], can be specified easily in our model.

This chapter describes a prototype system based on this extended model. This system uses a visual mapping module similar to TRIP[65, 67], TRIP2[85, 120], and TRIP3[88, 87]. Thus, the same visual mapping rules can be used in our animation system. In particular, using the TRIP3 system, users can easily create visual mapping rules by example for this animation system.

The rest of this chapter is organized as follows. Section5.2 explains how the programmer makes an animation, and Section5.3 describes the implementation details of our system. Section5.4 shows various examples of use of our prototype system, and conclusions are presented in Section5.6.

5.2   How to Construct an Animation — Insertion Sort Example

To create an algorithm animation with TRIP2a, programmers have to write the following:

Visual mapping rules Visual mapping rules are used for translating ASR to VSR. They specify how to visualize application data; i.e., they specify how to lay out graphical objects in the target animation (Figure 3.5(a)).

Transition mapping rules Transition mapping rules are used for mapping abstract operations to transitional operations. They specify how to move graphical objects in an animation (Figure 3.5(b)).

Annotations Annotations output abstract operations when an application executes its corresponding operations. They are used for extracting information from the executed application program, and can be regarded as specifications of translations from applications' operations to abstract operations (Figure 3.5(c)).

For example, to create an animation of an insertion sort algorithm, a programmer must write an insertion sort program, and annotate it so that the information on the execution of that program is obtained. Figure 5.1 shows an insertion sort program written in Prolog. Annotations are inserted at points where the application executes an interesting algorithm operation. In this example, after performing `insert/3`, `write_asr/2` outputs ASR data such as:

```
numlist([3,8,4,9],[1,2,5,6,7]).
```

The first argument of `numlist` represents numbers to be inserted, and the second argument represents already sorted (= inserted) numbers. At the same time, the predicate `write_ao/1` outputs an

```
% For demo
go :- tell('sort.anim'),
        % Write Initial State
        write_asr([5,2,6,7,1,9,4,8,3],[]),
        write('%'),nl,
        % Start Sorting
        sort([3,8,4,9,1,7,6,2,5], L),
        told.
% Sorting Program – Insertion Sort
sort(I, O) :- sort1([], I, O).
sort1(_, [], []).
sort1(L, [IH|IT], O) :-
        % Sort IT
        sort1([IH|L], IT, TMP),
        % and Insert a number to the sorted numbers
        insert(IH, TMP, O),
        % Output ASR data and Abstract Operations
        write_asr(L, O), write_ao(IH).
% Insert N to the sorted list
insert(N, [], [N]).
insert(N, [F|S], [N, F|S]) :- N =< F.
insert(N, [F|S], [F|L]) :- insert(N, S, L).
% Program for output ASR and AO
write_asr(L1, L2) :-
        reverse(L1, L3),
        write(numlist(L3,L2)), write('.'), nl.
write_ao(O) :-
        write(insert(O)), write('.'), nl,
        write('%'), nl.
```

Figure 5.1: Annotated insertion sort program.

abstract operation `insert(N)`, which means that the program inserts the number N. The character `%` output just after `insert/1` separates these data from those generated at the next execution of `insert/1`. This output is logged to a file (Figure 5.2), and used to create an animation of this execution.

Then, the programmer writes the visual mapping rule shown in Figure 5.3 to visualize the ASR data `numlist()`. There are two parts to this rule. In the first part, the predicate 'o' translates all the numbers in `numlist/2` into `box/4` using `num2barmap/1`, and in the second part, the predicate 'r' translates `numlist/2` into three `horizontallisting/3`, which arranges boxes horizontally with appropriate intervals.

Last, the transition mapping rule shown in Figure 5.4, which specifies that the inserted `object(X)` should move clockwise, is provided.

```
numlist([3,8,4,9,1,7,6,2,5], []).
%
insert(5).
numlist([3,8,4,9,1,7,6,2], [5]).
%
insert(2).
numlist([3,8,4,9,1,7,6], [2,5]).
%
insert(6).
numlist([3,8,4,9,1,7], [2,5,6]).
%
..... omitted .....
%
insert(3).
numlist([],[1,2,3,4,5,6,7,8,9]).
%
```

Figure 5.2: Log file of the insertion sort program.

```
% Mapping Numbers to Graphical Objects
objectmap([o]).
o  :- numlist(L1, L2),    % Get numlist/2
       num2barmap(L1),     % Map numbers to boxes
       num2barmap(L2).     % Map numbers to boxes
num2barmap([]).
num2barmap([B|L]) :-
       Y is B * 20,
       % a box with width=30 and height=Y
       box(B, 30, Y, [shaded]),
       num2barmap(L).
relationmap([r]).
r  :- numlist([H|L1], L2),    % Get numlist/2
       do_layout([H|L1], L2),% Layout numbers
       place(H,80,100,[]).
% Layout numbers horizontally
do_layout(L1, []) :-
       horizontallisting(L1, 12, [bottom_align]).
do_layout([], L2) :-
       horizontallisting(L2, 12, [bottom_align]).
do_layout(L1, L2) :-
       horizontallisting(L1, 12, [bottom_align]),
       horizontallisting(L2, 12, [bottom_align]),
       tail(L1, T), L2 = [H|_],
       horizontallisting([T, H], 50, [bottom_align]).
```

Figure 5.3: A visual mapping rule for insertion sort animation.

```
transitionmap([t]).
% insert(X) => move(X,[clockwise])
t  :- insert(X),              % get insert()
       move(X, [clockwise]).  % output move()
```

Figure 5.4: A transition mapping rule for insertion sort animation.

Figure 5.5 shows the resulting animation created with the above mapping rules[1]. Note that not only the inserted bar is moving; the other two bars are also being moved to make room for the inserted bar. The programmer did not describe such movements, but the system automatically created the transitions for the movements of the two bars (See Section 5.3.1).

Figure 5.5: Animation of an insertion sort algorithm.

5.3 Implementation of TRIP2a

We have implemented a prototype system based on the extended bi-directional translation model. This system is integrated with TRIP2 [85, 120] and shares the TRIP module and the interaction module. This system was written in Objective-C, and was created on NextStep[2] using the NextStep

---

[1]Our system uses color, but most screenshots in this thesis have been converted to gray-scale pictures. In addition, some figures show trails of objects, which are displayed only to illustrate the animation and are not actually presented in the real system.

[2]NextStep is the operating system developed by NeXT computer.

Application Kit[3] and Interface Builders. This section describes how we generate pictures intermediate between those generated from the application data, what kinds of transitional operations (including various animation effects) have been built, and how we obtain data from a running application program.

5.3.1 Implementation of Animations

In our implementation, a picture is represented as a set of picture objects in Objective-C. Thus, an animation is a set of transitions from the objects in one picture to those in another picture. An animation from picture A to picture B is created in the following manner:

1. Find correspondences between the objects in picture A and the objects in picture B, and examine how the object in picture A is changed to the corresponding object in picture B (Figure 5.6(1)).

2. Make instance objects of **Transition** class with the above pairs of corresponding objects, which will change the object in A to that in B (Figure 5.6(2)).

If the corresponding object is not found, the system will make a transition from/to the object to/from a *null* object (an object whose size is zero).
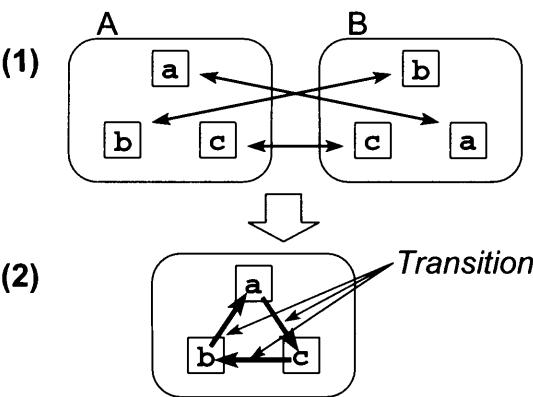

Figure 5.6: Making an animation.

Correspondences between objects in two pictures are searched for using the names of the objects. For example, if objX in picture A and objY in picture B have the same name, it is assumed that they are the same object, and that objX is moved and/or resized to objY. This means that to animate an object, the programmer must maintain its name in a sequence of pictures; i.e., the programmer has to name objects in a consistent way. The names of objects are specified in visual mapping rules.

Other animation systems often require explicit specification of transformations of graphical objects in an animation. For example, to generate animations using Pavane[103, 28], the programmer has to write a rule that maps data in the old and current states of a program into graphical objects, which change over time. In our system, the transformations of graphical objects are detected automatically, and do not need to be specified. Programmers have only to write a rule to layout abstract data, and *alter* the method of transformation if the default method is not sufficient.

---

[3]NextStep application kit is the Objective-C class library for developing various software on NextStep.

One interesting feature of our method of creating animations is that the structure of the moving sub- picture is often preserved automatically. For example, in Figure 5.7, which shows an animation of a list structure, the layout of the sub-structure ((b) c d) is maintained during its movement[4]. This is because the transitions of these cells and arrows are created using the same (**straight**) transitional operation, and the starting and the ending layouts of the sub-tree are the same. This kind of movement would be cumbersome to specify if the programmer had to specify the movements of every graphical object in the animation.
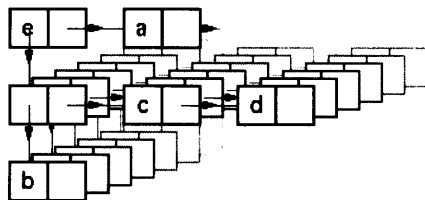

Figure 5.7: A cons-cell animation.

5.3.2 Specifying Transitional Operations

In our system, the motions of graphical objects are specified with transition mapping rules that translate abstract operations on ASR into transitional operations on VSR. These transitional operations are attached to each graphical object in the picture translated from ASR, and are used to make object transitions.

There are several ways to specify transitional operations:

1. *No specification.* When the programmer does not specify transition mapping rules, transitional operations for the graphical objects are determined automatically. For example, if a movement of a graphical object is not specified by the programmer, the transitional operations that move the object in a straight line and that scale linearly are used by default. Therefore, at the beginning of the construction of an animation, the programmer does not necessarily have to provide a transition mapping rule. The programmer can execute an animation without transition mapping rules, and may later improve it by specifying transitional operations.

2. *Using predefined transitional operations.* The following operations are provided:

   move Specifies how to move a graphical object. The following operations are provided:

   straight A straight path is created. The graphical object to which **straight** is attached moves in a straight line from the starting location to the ending location. This operation is used by default when no transitional operation is assigned to a moving object.

   clockwise, counterclockwise A path that goes clockwise/counterclockwise to the destination is created.

   lazy, immediate The **lazy** operation keeps an object from moving until the end of this transition. The **immediate** operation moves an object immediately at the beginning of this transition, to the final position and size.

   up, down, left, right These operations specify tangent vectors at the start and the end of a movement of a graphical object.

---

[4]Here, the list (e a (b) c d) is changed to (((b) c d) a).

rotate Specifies how to rotate an object during movement. The direction and the number of rotations can be changed.

scale Specifies how to scale an object. immediate, lazy, etc., are provided.

color Specifies how to change the color of an object.

Furthermore, several decorative animation effects can be specified in the same way. The techniques of cartoon animation[76] are also important in graphical user interfaces[24]. The following are examples of these techniques:

blink The object blinks during the movement, or blinks before and after its movement.

shake The object trembles before moving so that the start of its movement can be easily recognized.

slow-in-and-slow-out This operation slows down a graphical object near the start and the end of its movement, which makes the movement seem more natural.

squash-and-stretch This operation squashes or stretches a graphical object as it moves. This effect can depict rigidity of an object, and makes the object more vivid.

Figure 5.8 is a screenshot of an animation of the tower of Hanoi. It shows the trajectory of the disk moved from the left tower to the right tower. This is an example of use of slow-in-and-slow-out and squash-and-stretch. The moving disk in the figure is specified by the following simple transitional operations:

```
move(X, [up,squash-and-stretch]).
move(X, [slow-in-and-slow-out]).
```

The disk moves upward at the start. The movement of the disk is slowed down near the start and the end, and the disk is stretched as it moves[5].

Figure 5.8: Slow-In-Slow-Out and Squash-and-Stretch.

3. *Defining transitional operations.* It may be convenient if the programmer can define transitional operations by writing functions that describe the motion of a graphical object. We are planning to implement this feature by using a fast constraint solver to evaluate these functions.

5.3.3 Application Interface

As mentioned above, the layout and the movement of graphical objects can be determined by specifying only two mapping rules; i.e., a visual mapping rule and a transitional mapping rule. However, to obtain information from the running application, some programs must be inserted into the target application program. At present, there are three ways to do this:

---

[5]Here, the object is stretched in proportion to its acceleration.

- Writing the operations and the data in the application to a log file as ASRs and AOs. This can be used for almost all programming languages as it uses only file operations.

- Sending abstract operations and data using the ToTripSpeaker class, which is a subclass of the Speaker class in the NextStep Application Kit. In this case, the execution of the application and its animation run concurrently. However, the application program must be written in Objective-C.

- Inserting the predicate intrEvent into the application program on ASR[6].

All of the above methods require modification of the application program, which is obviously undesirable. We are planning to use the technique employed in debuggers to obtain operations and data in the application without modification of the application program.

5.4 Examples

This section describes several examples using our system. Our system uses color, but all screenshots of the system in this chapter have been converted to gray-scale pictures. In addition, some figures in this chapter show tracks of objects, which are displayed to illustrate an animation but do not actually appear on screen. The mapping rules for all the examples in this chapter are listed in Appendix D.

5.4.1 Animations of Data Structures

Animating Graph Structures

As mentioned above, we have a special constraint solver for the layout of undirected graphs. Using this constraint solver, a graph structure can be visualized easily by providing only its nodes and edges[7].

In the animation shown in Figure 5.9, a new edge between b and d is added to the graph in Figure 5.9(a). In this transition, due to an error with our graph constraint solver, the positions of nodes e and g are reversed in Figure 5.9(c). However, using an animation, users would easily be able to see that the positions of the nodes are exchanged[8].

Animating List Structures

List structures are often illustrated as box-and-pointer diagrams. Using almost the same visual mapping rules listed in [65, 67], we have animated list structures.

Figure 5.10 shows an animation depicting the allocation of a new cell. The boxes that represent the new cell (c) and the arrow pointing to the cell are newly generated in this transition. They gradually become larger, increasing eventually to the size of the cell. We can also easily move the cells in the list together (Figure 5.11). In this animation, the list (e a (b) c d) is changed to (((b) c d) a). The cells corresponding to ((b) c d) are moved together to the destination.

---

[6]ASR is implemented as clauses in Prolog. To use this method, the program has to be written in Prolog.

[7]It is also possible to specify relative lengths of edges.

[8]The exchange of nodes is caused by the differences in the initial conditions of the graph constraint solver. Of course, these sudden changes of an image are undesirable. We are currently working to improve the algorithm of the graph constraint solver.
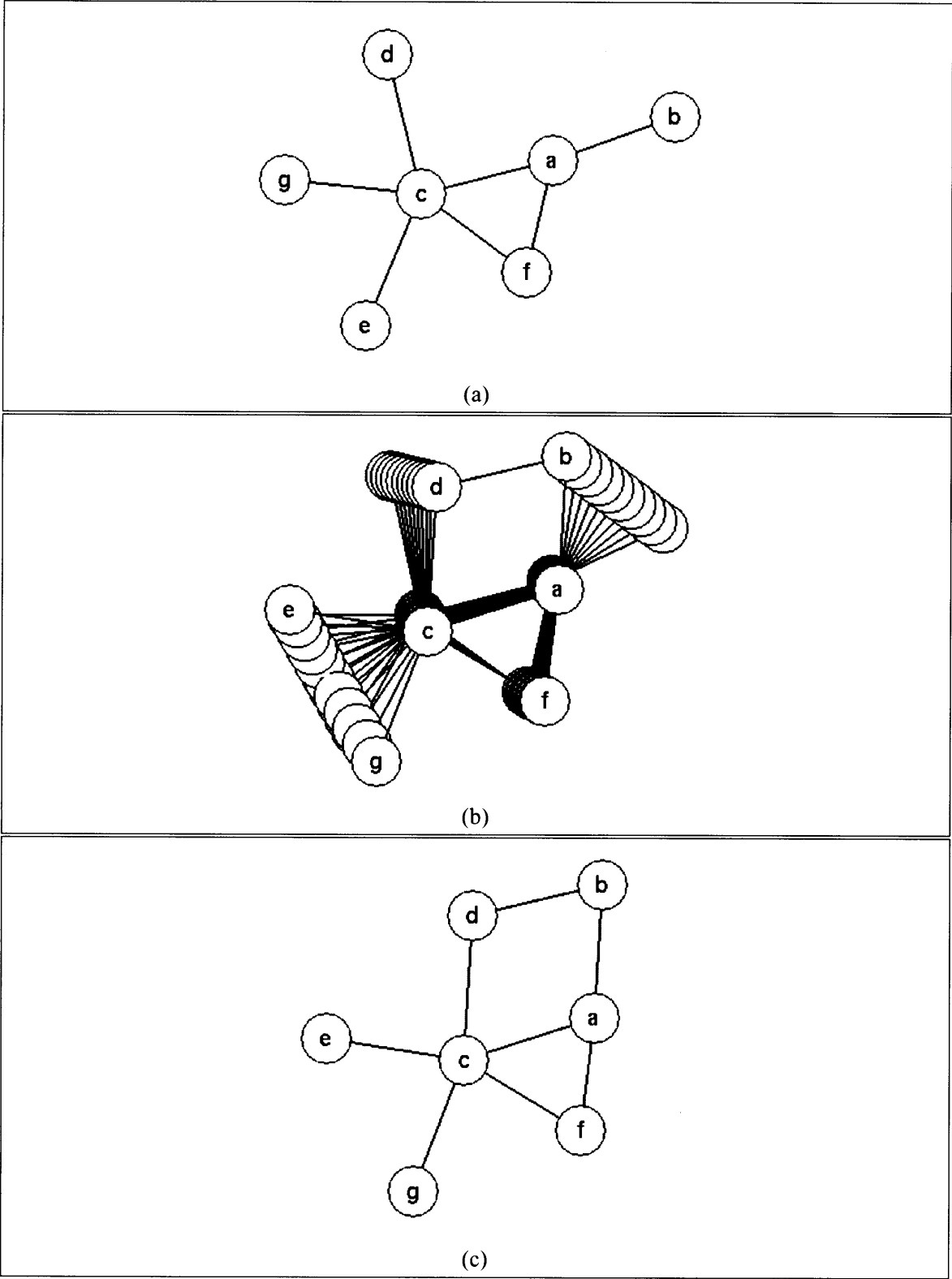
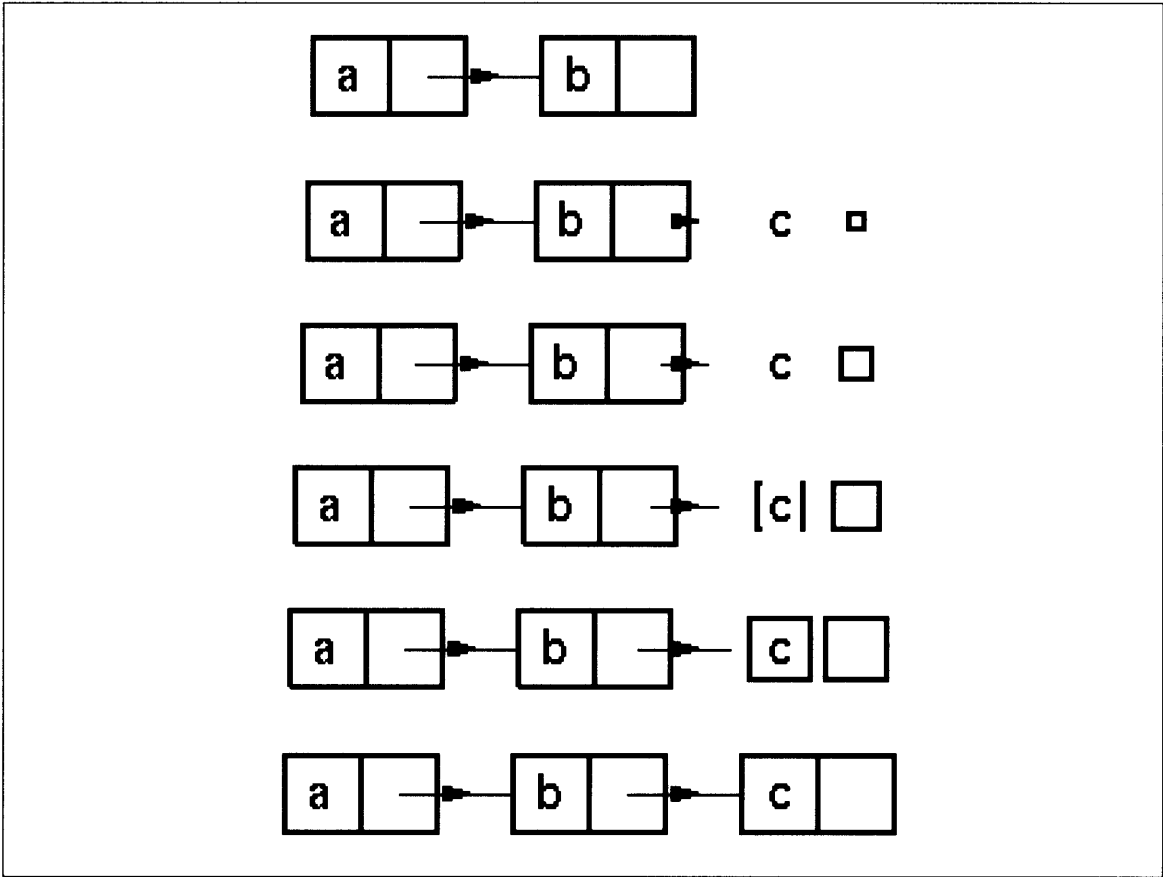Figure 5.9: Animation of a graph structure.
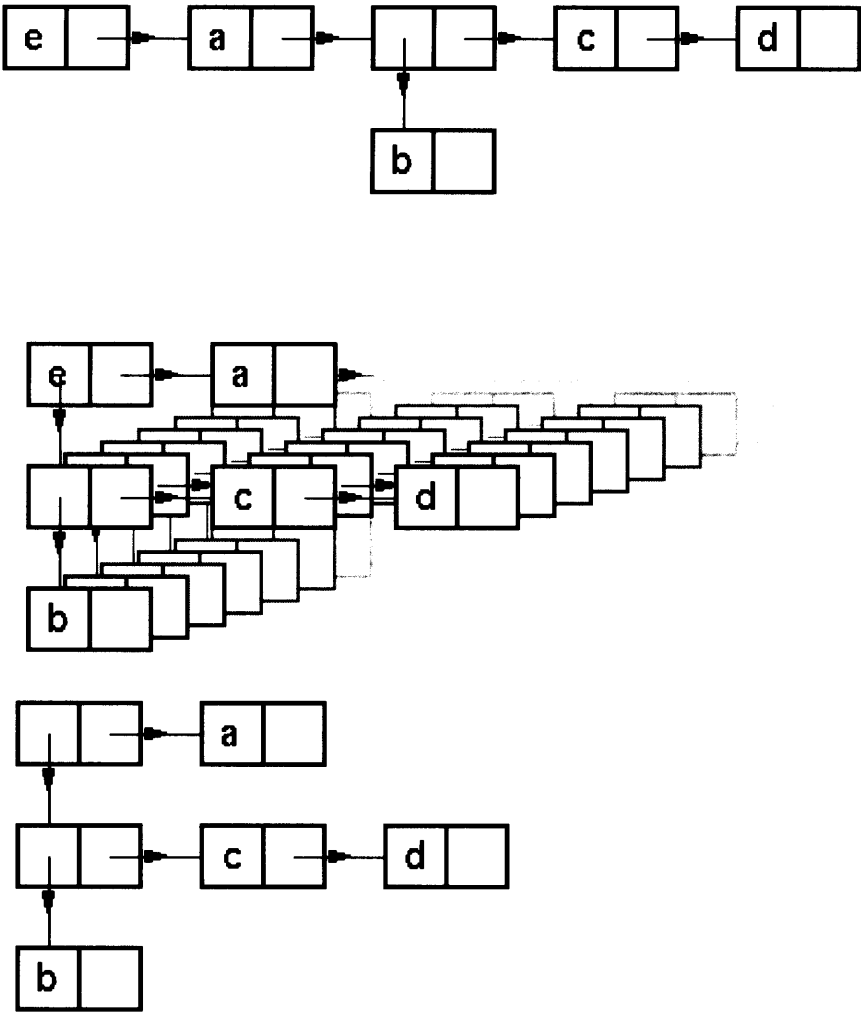
Figure 5.10: Allocation of a new cell.

Figure 5.11: Moving the substructure of a list

5.4.2   Sorting Algorithms

Sorting algorithms are often used for algorithm animation systems. The previous section already described the insertion sort animation. This section describes a bubble sort animation, a quicksort animation, a merge sort animation, and a heap sort animation.

Bubble Sort

The bubble sort algorithm is a naive sorting algorithm, which searches for a maximum number in order. This algorithm is thus named because the process of searching looks like a rising bubble. Figure 5.12 shows animations that represent a "rising" number. To move bars linearly, the default transitional operation `straight` is used in this animation. Therefore, this animation is created using only a simple visual mapping rule that lays out bars horizontally.

Quicksort

Quicksort is based on the divide-and-conquer paradigm. The numbers to be sorted are put in an array, and divided into two parts so that the numbers in one part are smaller than a *key* and those in the other part are larger than the key. Then, the two parts are sorted recursively.

The dividing process is performed using two indices. In our quicksort animation, they are represented as two small black rectangles (Figure 5.13(a)). The bars between the two indices are to be partitioned. The number at the right end of the sorting region is used as a key, which is drawn in black.

The process of partition proceeds as follows. First, the left index searches for a number larger than the key to the right, which is represented in the animation as its movement to the right, and then the right index searches for a number smaller than the key to the left, which is represented in the animation as its movement to the left. When these two numbers are found (Figure 5.13(b)), they are exchanged (Figure 5.13(c)). Here, the `clockwise` transitional operations are attached to the two moving bars. This process is repeated until the indices meet each other. In this partition, as only one number is larger than the key, no more exchange is necessary. At the end, the key is exchanged with the number at the boundary of the partition (Figure 5.13(d)). Then, the smaller and the larger part are each sorted in order. Figure 5.13(e) shows a screenshot at the beginning of the sorting of the smaller part.

Merge Sort

Merge sort is also based on the divide-and-conquer paradigm. First, the numbers are divided into halves. They are sorted severally, and then *merged*.

The numbers in the two parts before merging are drawn as white rectangles bounded with black or gray (Figure 5.14 (a)). Merging is achieved by taking the larger of the two numbers at the right end of the two parts, and arranging it in a different place. In Figure 5.14(a), the number at the right end of the left part is larger, so it is moved above (Figure 5.14(b)). Next, as the number of the left part is larger, it is moved above and arranged at the left of the previous bar (Figure 5.14(c)). The bars that are moved above are represented as gray shaded rectangles. When merging is finished, the merged numbers are moved down (Figure 5.14(d)), and the next merging process is started. Figures 5.14(e) and (f) show the final merging process.

In this example, no transitional operation is specified. Thus, the transitional operation `straight` is automatically used, and the rectangles move linearly.
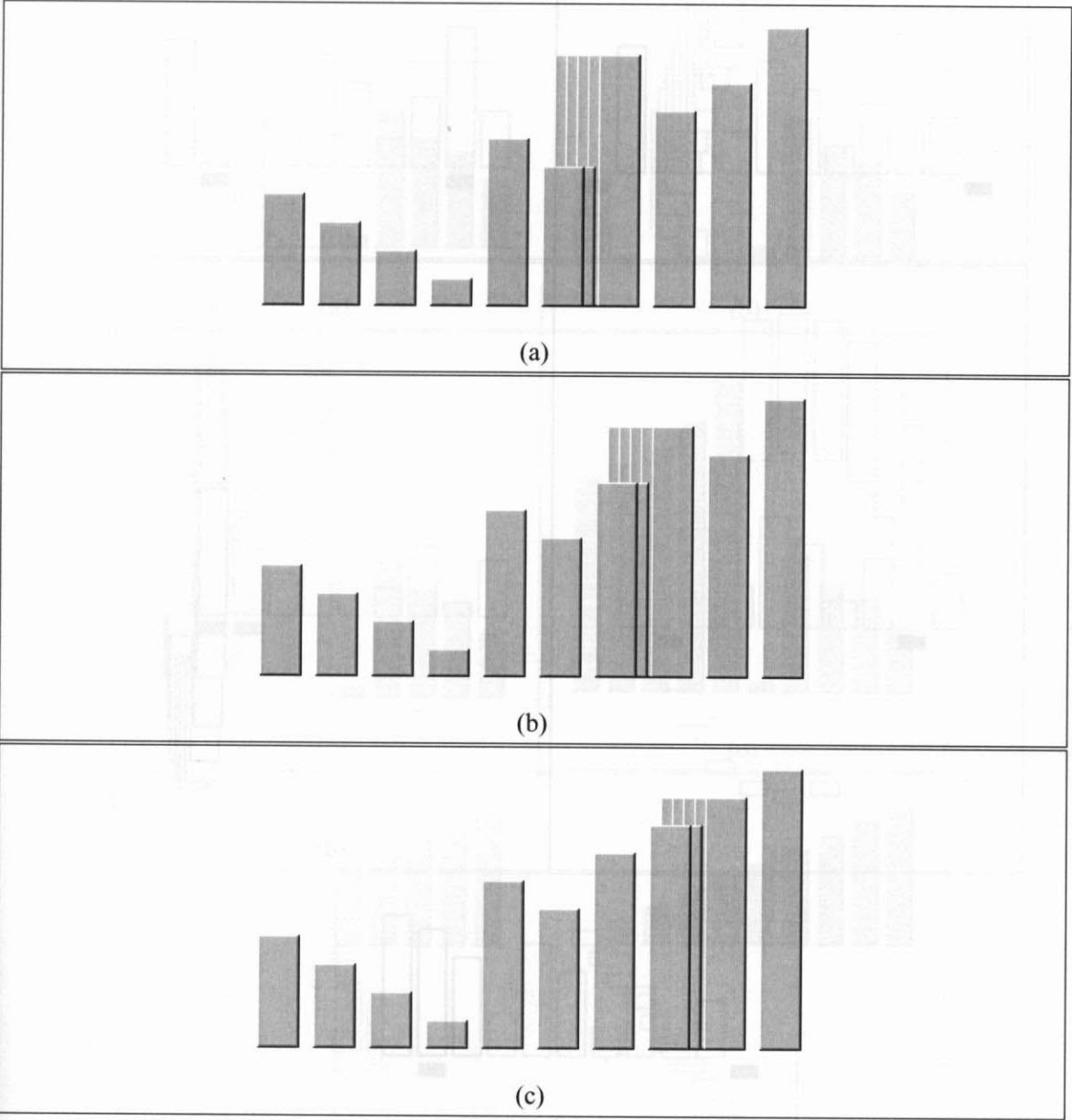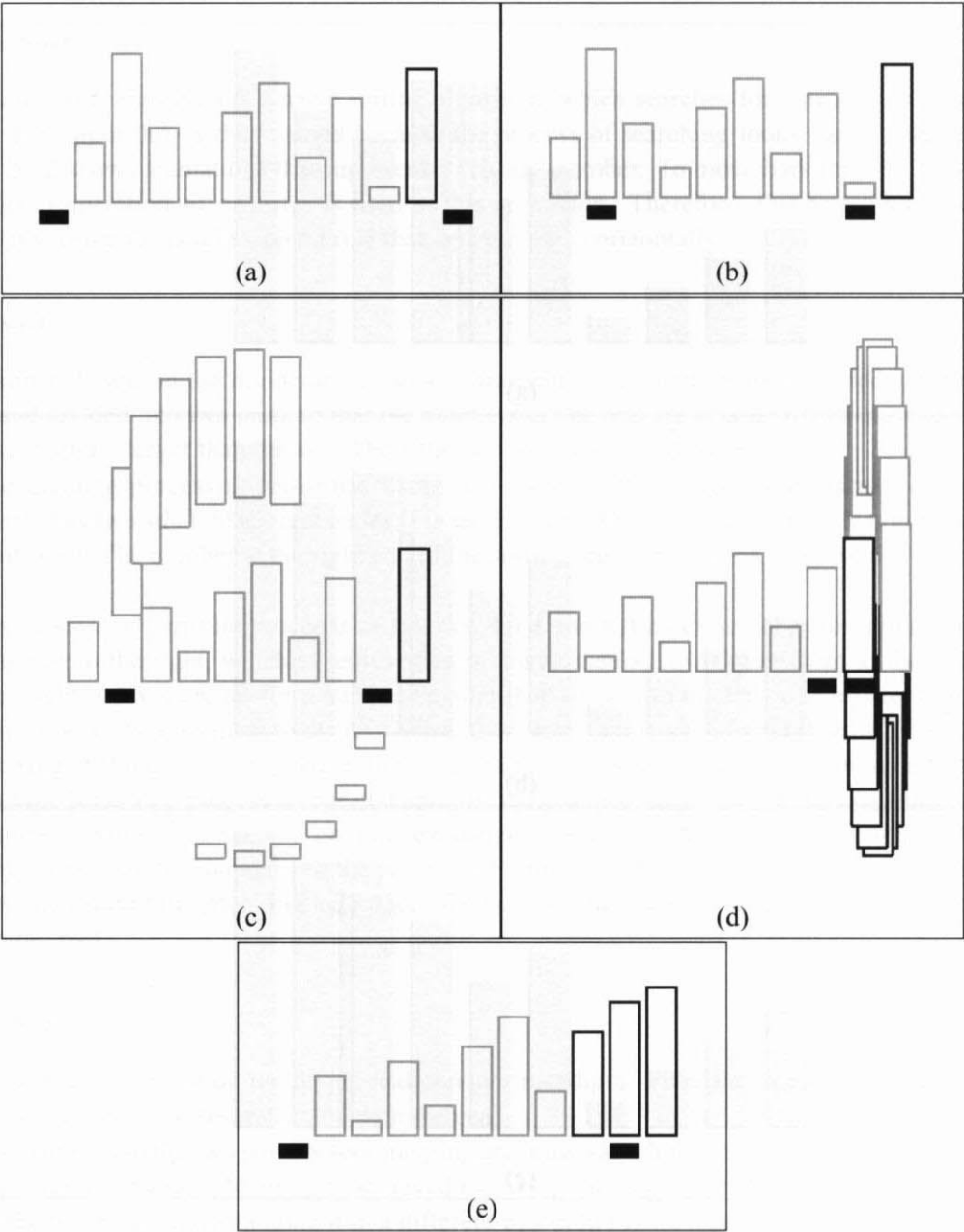
Figure 5.12: Bubble sort animation.
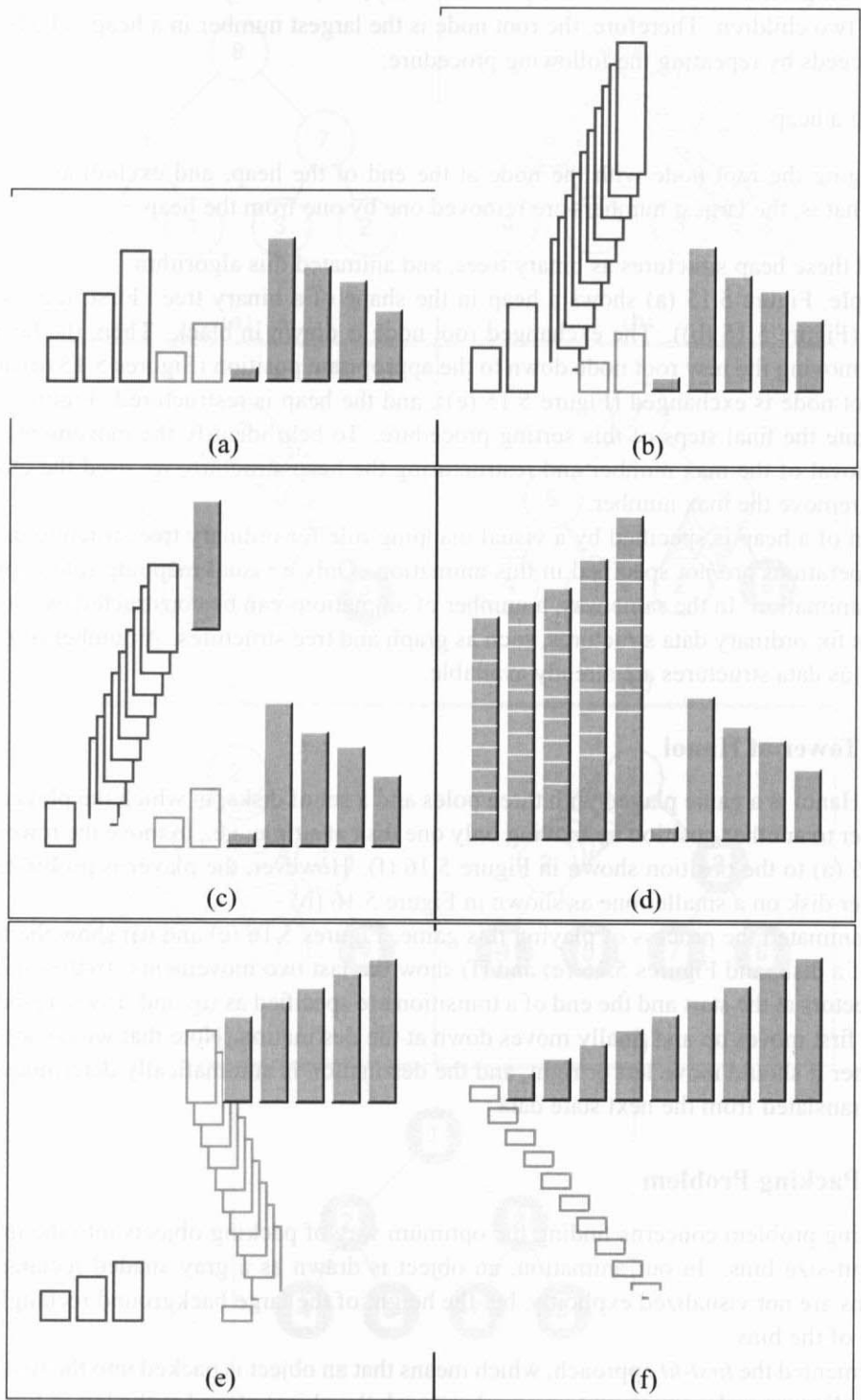
Figure 5.13: Quicksort animation.

Figure 5.14: Merge sort animation.

Heapsort

In the heap sort algorithm, the numbers are put into a *heap*, i.e., a binary tree in which each node is larger than its two children. Therefore, the root node is the largest number in a heap. The heap sort algorithm proceeds by repeating the following procedure:

1. Creating a heap.

2. Exchanging the root node with the node at the end of the heap, and excluding it from the heap. That is, the largest numbers are removed one by one from the heap.

We visualized these heap structures as binary trees, and animated this algorithm.

For example, Figure 5.15 (a) shows a heap in the shape of a binary tree. First, the root node is exchanged (Figure 5.15 (b)). The exchanged root node is drawn in black. Then, the heap is restructured by moving the new root node down to the appropriate position (Figures 5.15 (c) and (d)). Again, the root node is exchanged (Figure 5.15 (e)), and the heap is restructured. Figures 5.15 (f) and (g) illustrate the final steps of this sorting procedure. To help identify the movements associated with removal of the max number and restructuring the heap structure, we used the clockwise movement to remove the max number.

The layout of a heap is specified by a visual mapping rule for ordinary tree structure data. The transitional operations are not specified in this animation. Only a visual mapping rule is provided to create this animation. In the same way, a number of animations can be constructed by modifying mapping rules for ordinary data structures, such as graph and tree structures. A number of mapping rules for various data structures are already available.

5.4.3   The Tower of Hanoi

The tower of Hanoi is a game played with three poles and a set of disks, in which the player tries to move the tower to another position by moving only one disk at a time; i.e., to move the tower shown in Figure 5.16 (a) to the position shown in Figure 5.16 (f). However, the player is prohibited from putting a larger disk on a smaller one as shown in Figure 5.16 (b).

We have animated the process of playing this game. Figures 5.16 (c) and (d) show the first two movements of a disk, and Figures 5.16 (e) and (f) show the last two movements. In this animation, the tangent vectors at the start and the end of a transition are specified as up and down, respectively, so that a disk first moves up and finally moves down at the destination. Note that we do not need to specify whether it should move left or right, and the destination is automatically determined by the next picture translated from the next state data.

5.4.4   Bin-Packing Problem

The bin-packing problem concerns finding the optimum way of packing objects into the minimum number of unit-size bins. In our animation, an object is drawn as a gray shaded rectangle (Figure 5.17). Bins are not visualized explicitly, but the height of the large background rectangle represents the size of the bins.

We implemented the *first-fit* approach, which means that an object is packed into the first bin that can contain it. For example, an attempt was made to pack the object placed at the left in Figure 5.17 (a) into the first bin (Figure 5.17 (b)), but this failed. This process was repeated at the next bin, and was successful (Figure 5.17 (c)).

As in the animation of the tower of Hanoi, the transitional operations up and down are specified for the tangent vectors of the transition of this animation.
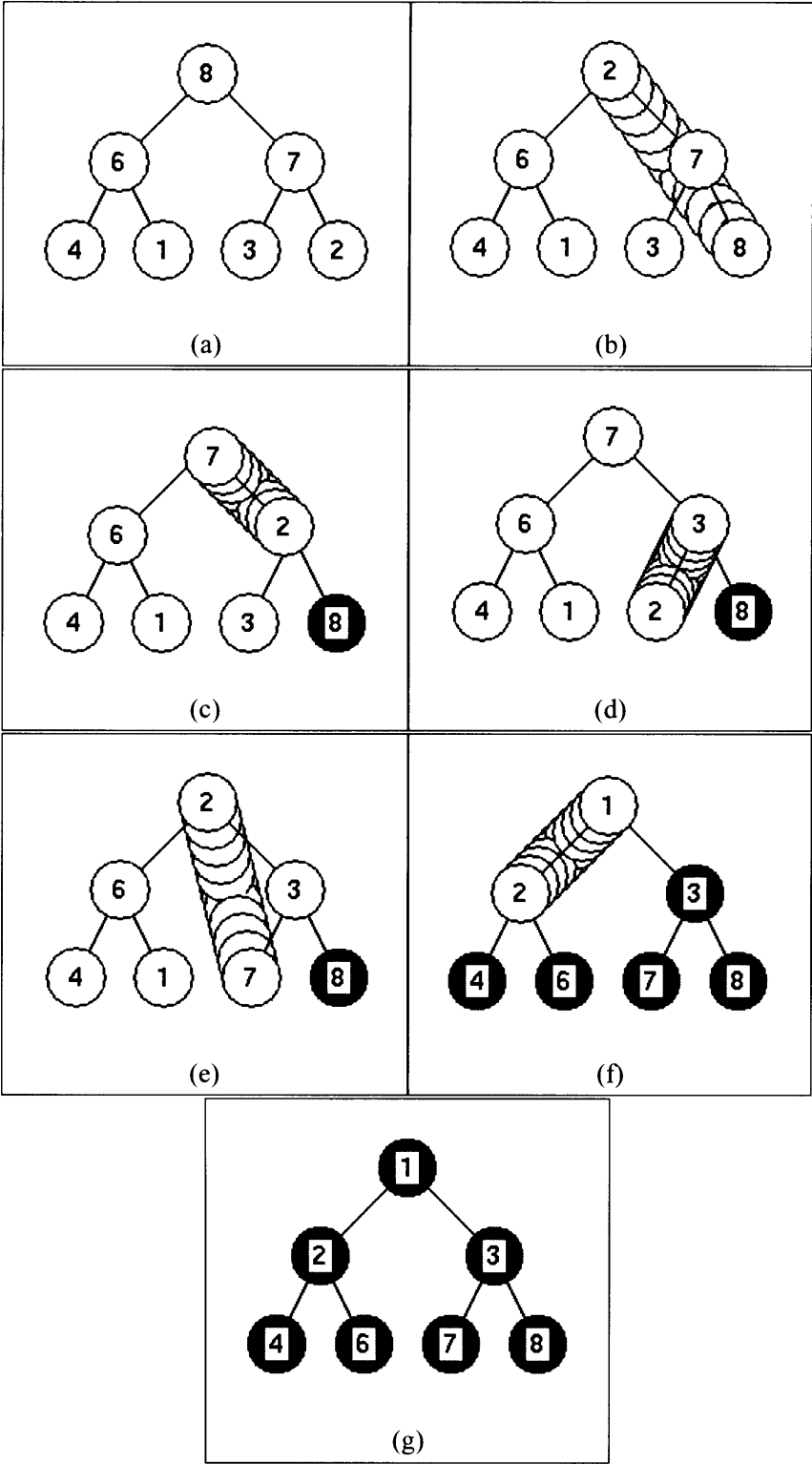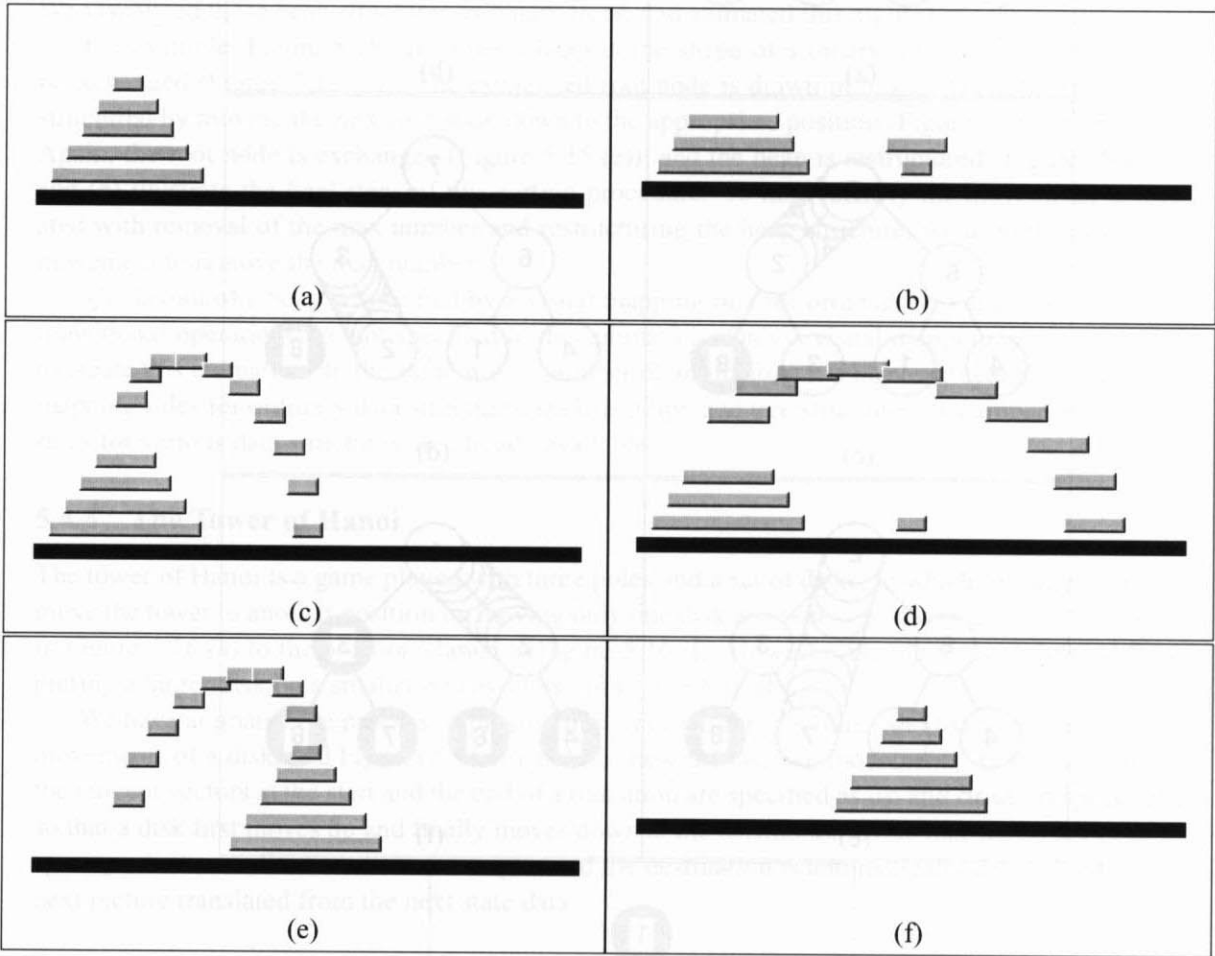
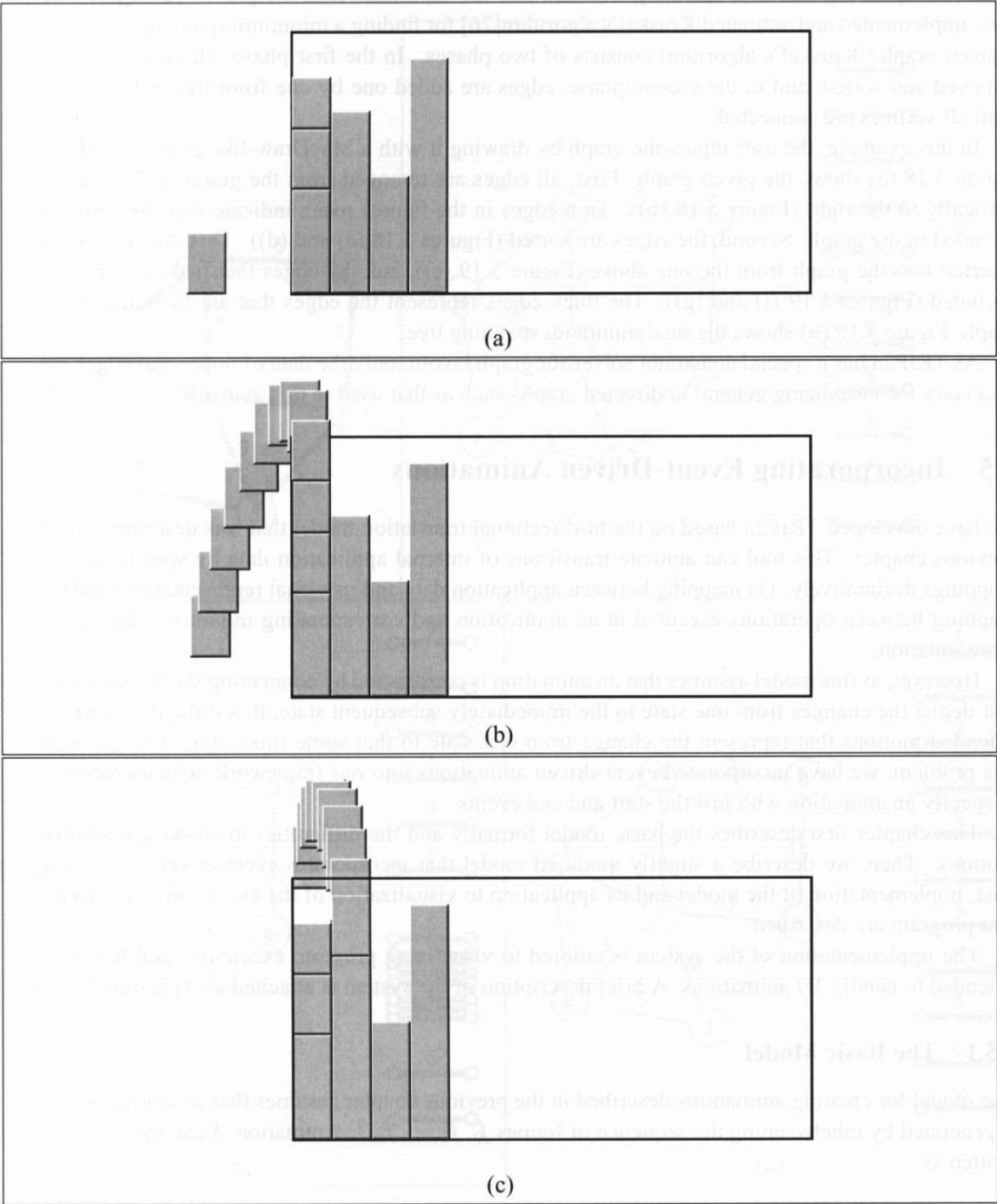Figure 5.15: Heap sort animation.

Figure 5.16: The tower of Hanoi.

(a)

(b)

(c)

Figure 5.17: Bin-packing algorithm.

5.4.5   Finding a Minimum Spanning Tree

A spanning tree is a subset of an undirected graph that is acyclic and connects all the vertices. A minimum spanning tree is a spanning tree that has the minimum total weight of its edges. We have implemented and animated Kruskal's algorithm[26] for finding a minimum spanning tree from a given graph. Kruskal's algorithm consists of two phases. In the first phase, all edges are first removed and sorted, and in the second phase, edges are added one by one from lighter to heavier until all vertices are connected.

In this example, the user inputs the graph by drawing it with a MacDraw-like graphical editor. Figure 5.18 (a) shows the given graph. First, all edges are removed from the graph, and arranged vertically to the right (Figure 5.18 (b)). Thin edges in the figures mean indicate that they are not included in the graph. Second, the edges are sorted (Figures 5.18 (c) and (d)). Then, the edges are inserted into the graph from the one above (Figure 5.19 (e)), but any edges that make a cycle are excluded (Figures 5.19 (f) and (g)). The thick edges represent the edges that are included in the graph. Figure 5.19 (h) shows the final minimum spanning tree.

As TRIP2a has a special constraint solver for graph layout, only the data of nodes and edges are necessary for visualizing general undirected graphs such as that used in this animation.

5.5   Incorporating Event-Driven Animations

We have developed TRIP2a based on the bi-directional translation model that was described in the previous chapter. This tool can animate transitions of internal application data by specifying two mappings declaratively: (1) mapping between application data and its visual representation; and (2) mapping between operations executed in an application and corresponding motions in the visual representation.

However, as this model assumes that an animation is constructed by connecting short animations that depict the changes from one state to the immediately subsequent state, it is difficult to display extended motions that represent the change from one state to that some time later. To cope with this problem, we have incorporated event-driven animations into our framework, which enable us to specify an animation with just the start and end events.

This chapter first describes the basic model formally and the difficulties in showing extended motions. Then, we describe a slightly modified model that incorporates event-driven animation. Last, implementation of the model and its application to visualization of the execution of a concurrent program are described.

The implementation of the system is tailored to visualizing program executions and has been extended to handle 3D animations. A brief description of the system is attached as Appendix B.

5.5.1   The Basic Model

The model for creating animations described in the previous chapter assumes that an animation, A, is generated by inbetweening the sequence of frames $f_0, f_1, ..., f_n$ [9]. Animation A can therefore be written as:

$$A = I(f_0, f_1) + I(f_1, f_2) + ... + I(f_{n-2}, f_{n-1}) + I(f_{n-1}, f_n) , \qquad (5.1)$$

where $I(f_i, f_{i+1})$ is a short animation generated by inbetweening frames f_i and f_{i+1}, and '+' indicates the concatenation of two animations. The frame f_i is a visualized picture translated from the

---

[9]Strictly speaking, we should refer to these as *keyframes*. However, as we do not refer to the frames generated by inbetweening keyframes, the word *frame* refers to *keyframe* in this chapter.
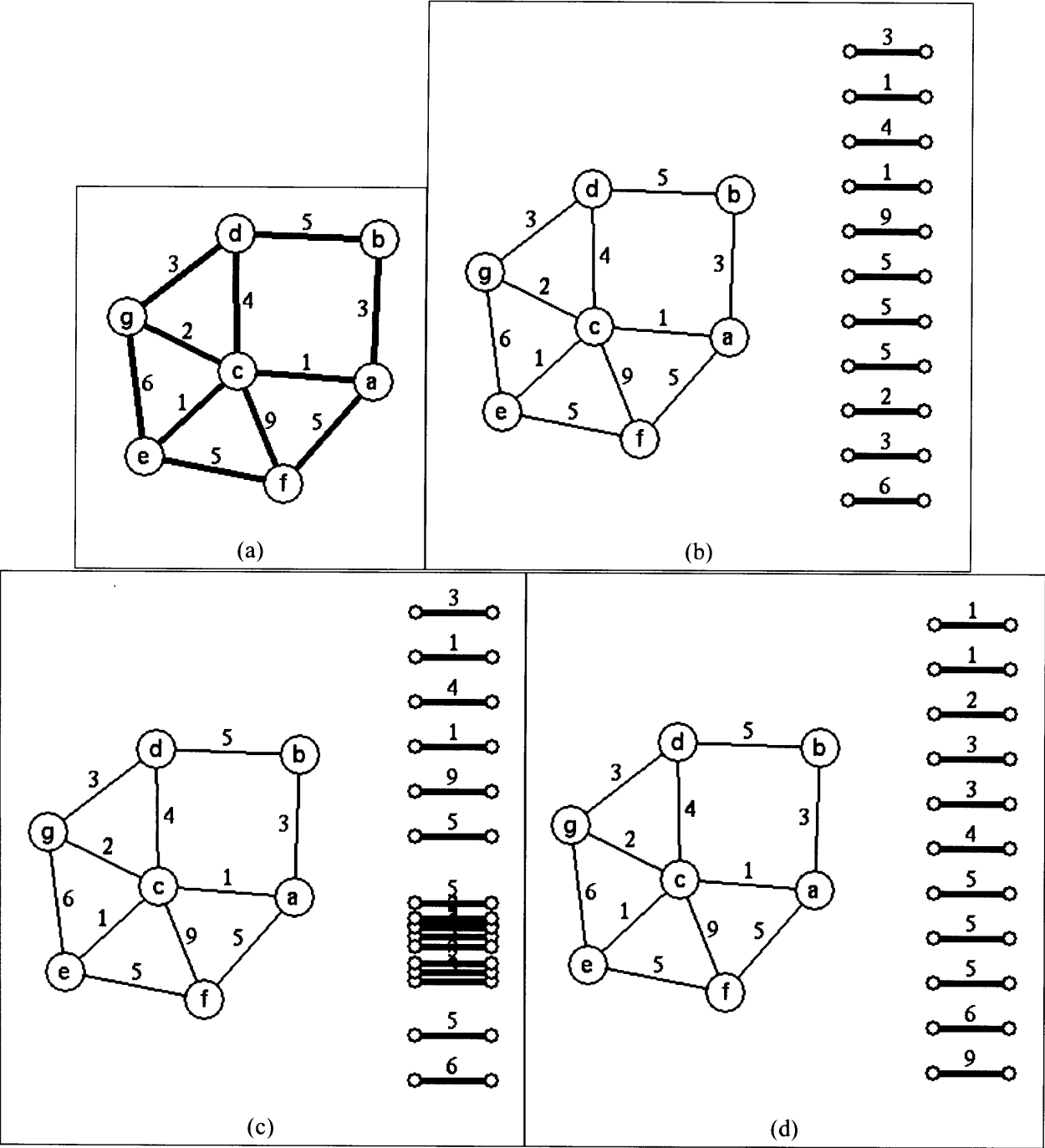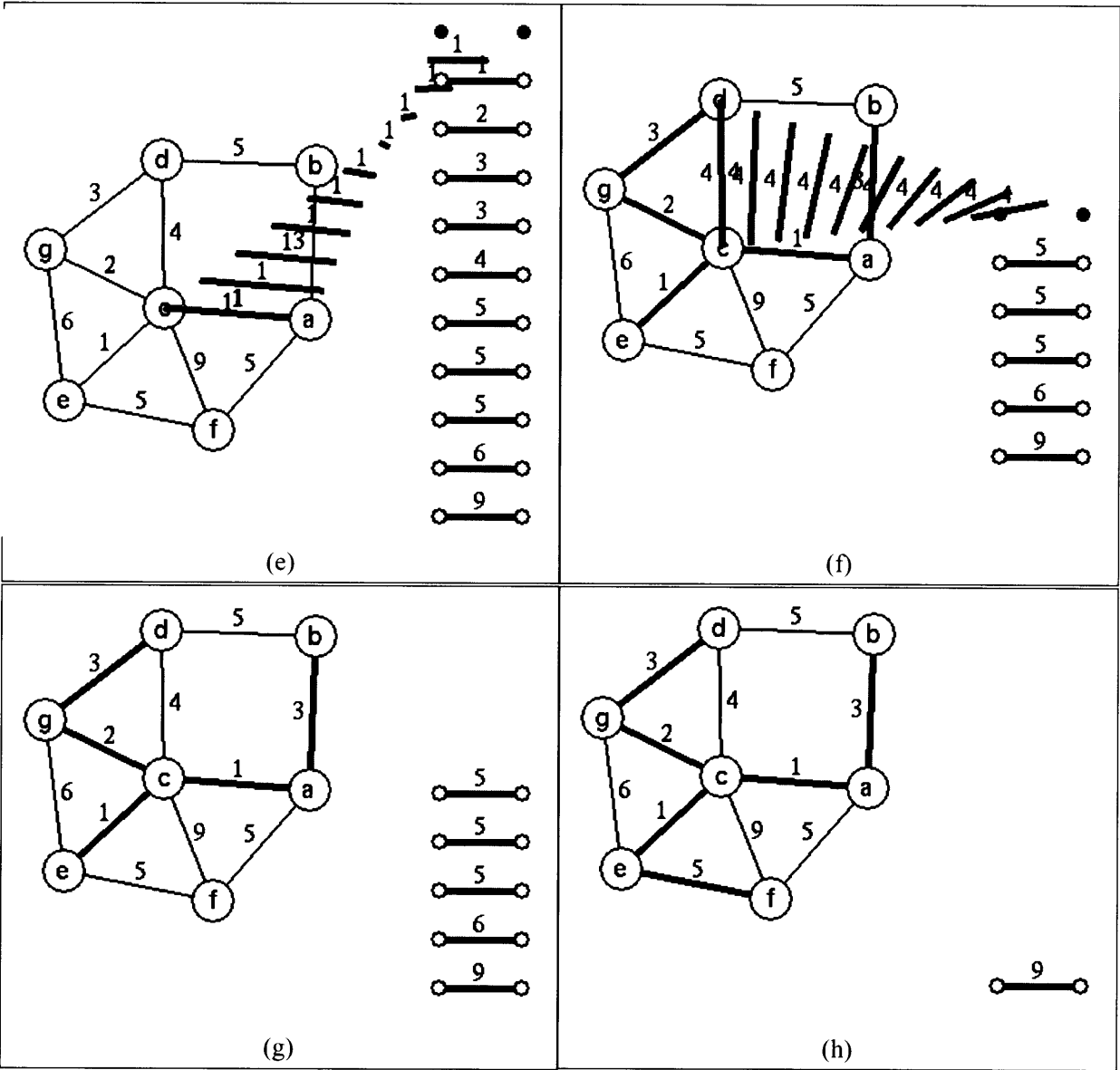
Figure 5.18: Minimum spanning tree algorithm (1).

Figure 5.19: Minimum spanning tree algorithm (2).

internal state s_i by the visual mapping function $vm : S \rightarrow F$.

$$f_i = vm(s_i) \,. \tag{5.2}$$

This model assumes that the execution of a target application is a sequence of operations, op, that change the internal data of the application.

$$\ldots \rightarrow s_i \xrightarrow{op_i} s_{i+1} \xrightarrow{op_{i+1}} s_{i+2} \rightarrow \ldots \tag{5.3}$$

As described above, the animation is a concatenated sequence of short animations $I(f, f_{i+1})$(Eqn 5.1). That is, each short animation $I(f_i, f_{i+1})$ represents the operation op_i.

5.5.2 The Problem

There are situations that cannot be represented well with the above model. As an example, let us consider the message-passing animation in which two nodes, A and B, send messages to each other. We represent these nodes as two boxes, and messages as spheres (Figure 5.20).
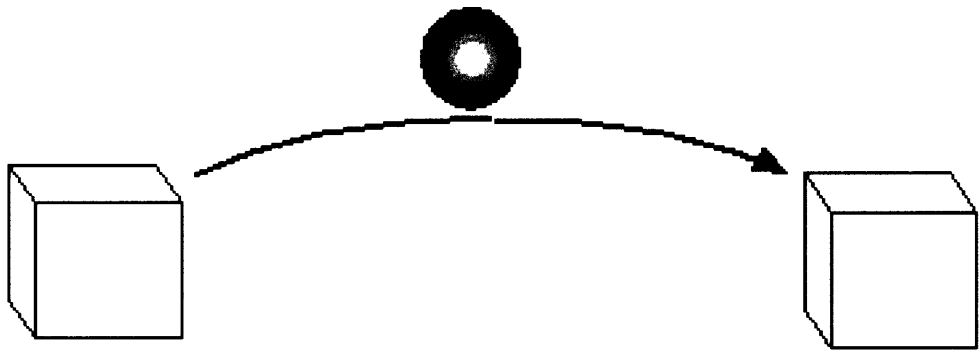
Figure 5.20: Sending a message.

With our model, the animation that shows a message m_1 being sent from node A to node B is constructed by concatenating three units of animation that interpolate four successive states. Each state is described by the state of the message. There are three types of state: (1) no message exists; (2) a message exists at node A; (3) a message exists at node B. The following are the four successive states that occur during message sending:
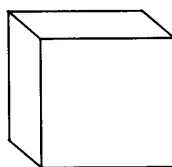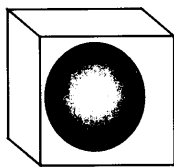
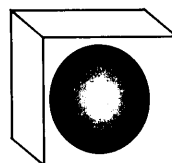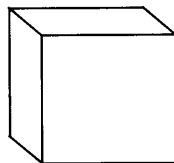1. There are two nodes in the state.

   ```
   a, b : node
   ```

2. The message m_1 is located at node A.

   ```
   a, b : node
   m1(a) : message
   ```
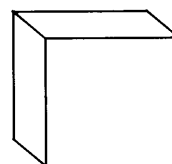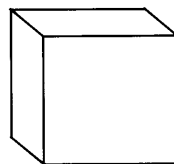
3. The message is located at node B.

```
a, b : node
m1(b) : message
```


4. There are two nodes in the state.

```
a, b : node
```


The three units of animation that interpolate the above states are:

$I(1,2)$: The animation that shows the appearance of the message object.

$I(2,3)$: The animation that shows the movement of the message object from node A to node B.

$I(3,4)$: The animation that shows the disappearance of the message object.

The entire animation A_1 is a concatenation of the above animations, i.e., $A_1 = I(1,2) + I(2,3) + I(3,4)$.

Next, consider the case in which two messages m_1 and m_2 are sent from node A to node B. This is simple if the second message is sent after the arrival of the first at node B, because it is achieved by concatenating the above message-passing animation twice.

$$A_{2s} = A_1 + A_1 \qquad (5.4)$$

However, the case in which the second message m_2 is sent before receipt of the first message is difficult to handle with the model. In this case, we cannot construct the animation by only concatenating the unit animations: $I(1,2)$, $I(2,3)$, $I(3,4)$.

To make the animation of this execution, these units of animation must be *overlapped*. However, this is difficult with our basic model, because it is necessary to determine the state when the message is still in transit and generate a keyframe by visualizing it. For example, when the second message m_2 starts moving, the first message m_1 is still in transit to node B. As the first message is at neither node A nor node B, the position of the message cannot be specified statically.

That is, in our model, (1) animations are created by concatenating short animation units, and (2) the programmer must specify the location and size of all objects at the start and end frames of each animation unit. It is difficult to satisfy these constraints if multiple objects start and stop moving asynchronously, because it is difficult to know the positions of moving objects as the system calculates them by interpolating frames.

5.5.3 Approach

To cope with this problem, we modified the model to enable specification of the ends of unit animations. The unit animations are constructed with respect to each object in keyframes. In the original model, it was assumed that unit animations of all objects are generated by interpolating the current and the next keyframe. The new model makes it possible to make unit animations of some objects by interpolating the current and distant keyframes. When interpolating distant keyframes, the in-between keyframes are ignored so that the programmer does not need to determine the positions of moving objects, which allows overlapping of unit animations.

More formally, the original and the new model are described as follows. Each frame f_i is a set of graphical objects.

$$f_i = \{obj_j^i | 0 \le j \le m\}$$

A unit animation of an object is written as $I_{obj}(obj_j^k, obj_j^l)$ which is a motion from obj_j^k (the object obj_j in the frame f_k) to obj_j^l. The animation in the original model is written as follows.

$$I(f_i, f_{i+1}) = \bigoplus_j I_{obj}(obj_j^i, obj_j^{i+1}) \tag{5.5}$$

$$A = \sum_{i=0}^{n-1} I(f_i, f_{i+1}) . \tag{5.6}$$

In the above equation, '$\oplus$' means parallel concatenation of motions of objects in which element animations may overlap, and 'Σ' means sequential concatenation of unit animations. On the other hand, the animation in the new model is written as follows.

$$A_{new} = \bigoplus_{j,k} I_{obj}(obj_j^{e_j(k)}, obj_j^{e_j(k+1)}) . \tag{5.7}$$

Here, $e_j(k)$ is the kth number of a monotone increasing subsequence of $0...n$. Figure 5.21 shows both the original model and the new model. To differentiate the new type of unit animations, we
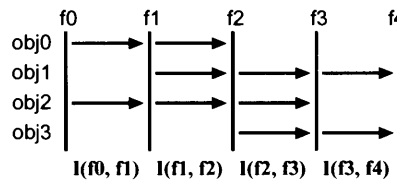
Figure 5.21: The original & new models.

call $I_{obj}(obj_j^k, obj_j^l)$ *long actions* if $k + 1 < l$. On the other hand, we call $I_{obj}(obj_j^k, obj_j^{k+1})$ *instant*

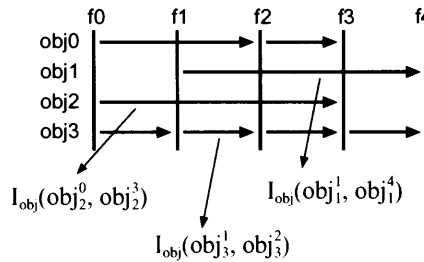*actions.* For example, in Figure 5.21, $I_{obj}(obj_2^0, obj_2^3)$ is a long action, and $I_{obj}(obj_3^1, obj_3^2)$ is an instant action. The actions in the original model (Figure 5.21(a)) are all instant actions.

To make the animation $I_{obj}(obj_2^0, obj_2^3)$, only the keyframes f_0 and f_3 are required, and f_1 and f_2 are not necessary. Similarly, to make $I_{obj}(obj_1^1, obj_1^4)$, the keyframes f_1 and f_4 are required, and f_2 and f_3 are not necessary. Therefore, the programmer does not need to specify the positions of moving objects. For example, the frame f_1 is specified only by the positions of obj_1 and obj_3, and the frame f_2 is specified only by the positions of obj_0 and obj_3. This enables the overlapping of unit animations, such as $I_{obj}(obj_2^0, obj_2^3)$ and $I_{obj}(obj_1^1, obj_1^4)$.

5.5.4   Implementation

Based on the new model described in the previous section, we have built the TRIP2a/3D system to handle long actions. This section describes the implementation details of the new TRIP2a/3D system.

Specifying Long Actions

A long action is specified by a pair of transitional operations. Usually, each operation corresponds to an event during the execution of an application. The programmer modifies the program so that it writes out the events and the current state together to a log file when such events occur during its execution.

Events are treated as abstract operations on ASR. Similar to abstract operations, the programmer defines events as terms, but each must have an ID number. The ID number of an event is necessary to allow searching for a pair of events. Two events with the same ID are considered to be a pair of events that represent a long action. The event data are translated to special transitional operations of type move/2, which indicate long actions. The translation from events to transitional operations is specified by mapping rules written by the programmer.

As an example, consider the message-passing animation example described above. In this case, sending a message corresponds to the start event of a long action, and receiving a message corresponds to the end event. First, the programmer defines these events: snd(Obj, From, Id) and rcv(Obj, To, Id). This means that the message Obj is sent from From to To. The third variable, Id, is the ID number of the pair of events.

Then, the programmer defines mapping rules that map these two events to special transitional operations. These mapping rules are written as follows:

```
rule(snd(Obj, From, Id), Result) :-
    Result = [move(Obj, [from(From,Id)])].
rule(rcv(Obj, To, Id), Result) :-
    Result = [move(Obj, [to(To,Id)])].
```

Thus, two types of events, snd(Obj, From, Id) and rcv(Obj, To, Id), are mapped to the transitional operations move(Obj, [from(From,Id)]) and move(Obj, [to(To,Id)]), respectively.

The variables Obj, From, To in the arguments move/2 are IDs of graphical objects. The two moves that have the same ID[10] are regarded as a pair, and are thought to represent the long action in which Obj moves from the position of From to the position of To.

---

[10]This ID is unique to this type of transitional operation, and is not concerned with the IDs of graphical objects.

Generating an Animation for a Long Action

Chapter 5 described how animation is generated by interpolating pairs of keyframes that are translated from application data. In our new model, animations that represent long actions are generated in the same way, but this leads to a problem in implementing the system. In the previous model, the two keyframes to be interpolated were adjacent, and a unit of animation was generated by interpolating two adjacent frames. In the new model, the two keyframes that correspond to the start and end events may be distant. Even when the system finds a frame that corresponds to the start frame of a long action, it cannot know when the end frame will appear in the execution log. As both events are necessary to generate an animation for a long action, the system will have to search for another event for the rest of the log file once one of them is found.

There is no problem if the animation is generated after the execution is finished, because the entire log can be obtained and searched. However, if the animation must be displayed parallel to the execution of the target application, the start of the animation for a long action must be delayed until the end event occurs. Currently, we only generate animations after the execution has finished.

Architecture

Figure 5.22 shows the architecture of TRIP2a/3D. The programmer creates the hatched area in Figure 5.22 for each target application. The programmer first instruments the application to output execution logs. Unlike TRIP2a, TRIP2a/3D reads only from text log files. The application must output its internal data and events into a text file, and then TRIP2a/3D reads and visualizes them. When it is difficult to output data directly in the format of ASR terms, especially when they are distributed applications, the programmer must write a program to collect log files from each executed application and compile them into the ASR data file.

The programmer then writes visual mapping rules, and builds a *mapper* module for each type of execution log. Visual mapping rules are compiled and linked with other libraries. They are written in KLIC [23], a concurrent logic programming language. The resulting module reads the input ASR data and translates them into animation scripts for the viewers.
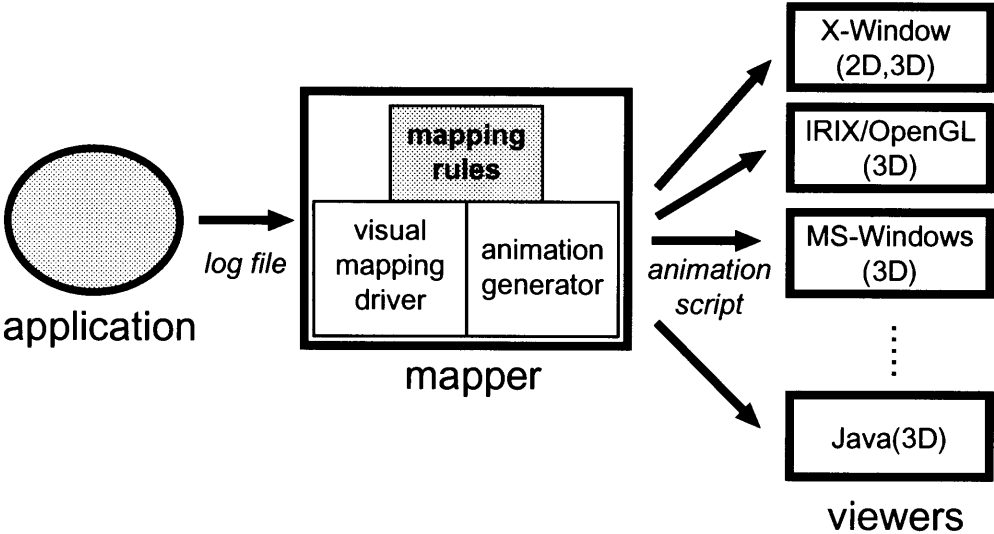

Figure 5.22: Architecture of TRIP2a/3D.

The major difference from the implementation of TRIP2a is that the viewers and the visual

mapping module are separate, and several viewers are provided for different platforms. It is also possible to view animations in different ways, such as two-dimensionally or three-dimensionally, by changing the viewer. The three-dimensional viewer is implemented with Open Inventor[126] and OpenGL, and the two-dimensional viewer is implemented with Tcl/Tk and STk. We have also implemented the viewer with Java using Java3D. The multiple viewers can be easily implemented, because we defined a simple *animation script*, which is a series of low-level operations on graphical objects such as move objects and change size. The viewers are interpreters that interpret the animation scripts. Their syntax is described in Appendix B.

Figure 5.23 shows a screenshot of the viewer implemented with Java. The viewer has two modes. In animation mode, the animation is displayed in the window. By pressing the button at the lower-left corner of the window, the user can switch the viewer to the step mode. In this mode, the user can freely wind the animation forward or backward using the slider. The window displays the still image at the moment corresponding to the slider position. The 3D viewing position can be changed with a mouse and a keyboard. The user can rotate the displayed objects by dragging on the window, and can shift the view or zoom-in/out by using the cursor keys and Page-Up/Down keys, respectively.
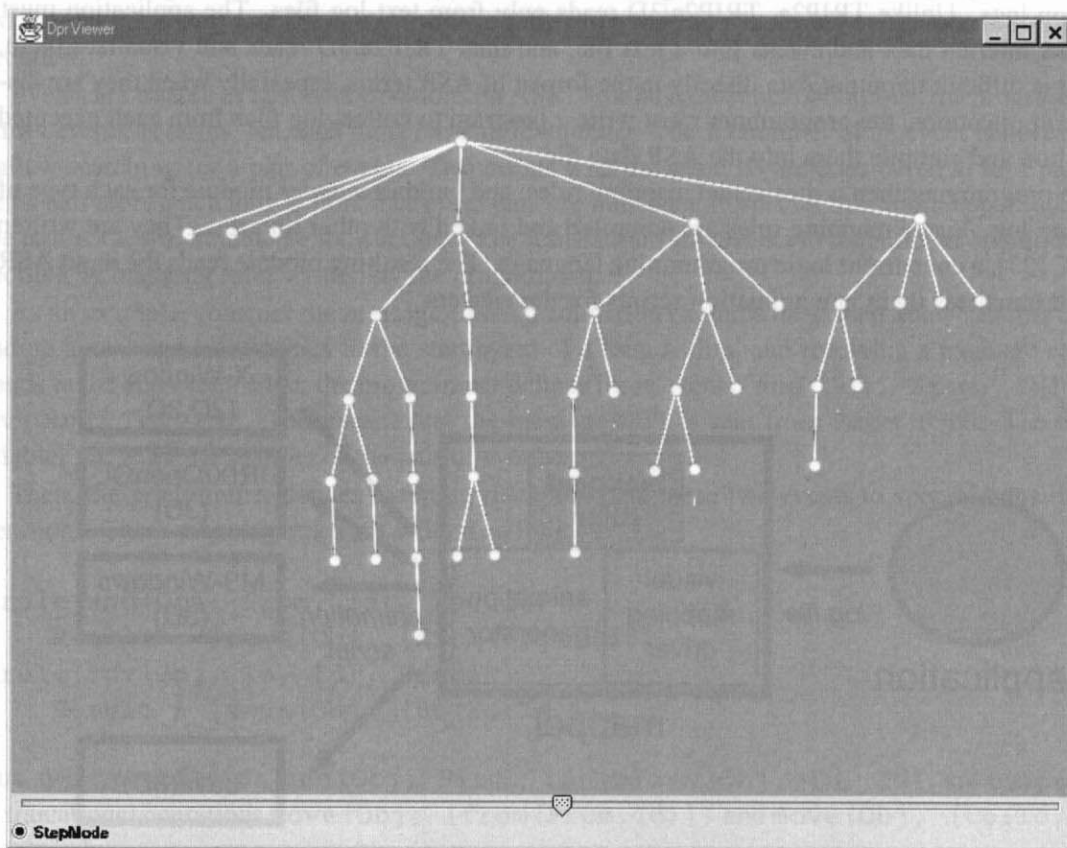

Figure 5.23: Screenshot of Java-TRIP2a/3D.

5.5.5 An Example

The N-Queen Problem

Figure 5.24 shows a search tree for the N-Queen (N=6) problem. Each branch represents a choice of position where a queen can be placed. In this search, queens are placed from the first to the last row in order. Only one queen can be placed in each row, so there are only six choices. However, the program cuts branches when some of the choices conflict with the queens already in position. If there is no choice, the node cannot extend a branch.

This animation shows two types of action. The first type of action is the extension of branches. When the N-Queen program has found the position choices for the next row, the animation shows the emergence of new child nodes. For example, when the program has found that there are three choices of position, the animation shows the creation of three new children (branches).

Another type of action is the sending of messages to the root node. A message is sent when the program has succeeded in placing six queens on the board. In this case, messages are sent from the node at a depth of six. A message is also sent when the program has failed to find a choice of position in the next row. In this case, the program cannot extend any more branches, and the depth of such a node is less than six.

In the animation, the messages are represented as flying circles, which fly from the source node to the root node. The color of nodes in the search tree represents the machine on which the search is executed. For example, two machines are used for the search shown in Figure 5.24. The left and the right halves of the tree are executed on different machines. The light green nodes on the left part of the tree are executed on one machine (A), and the other brown nodes are executed on another (B).

The execution is started on machine A. The first six choices are divided into two parts, and half are sent to machine B. The execution on machine A proceeds faster than that on machine B (Figure 5.24). In addition, the message from the last node arrives immediately at the root node. On the other hand, many messages have been sent from machine B, but have yet to arrive at the root node (Figure 5.24), because the left part of the tree is executed on a different machine from the root node. This is represented by many flying circles in the animation.

Figures 5.25 and Figure 5.26 show a program that solves the N-Queen problem[1]. The program runs on several machines in parallel. The log data output from the program are shown in Figure 5.27. The data are converted to ASR by the special converter. The resulting ASR are shown in Figure 5.28. The visual mapping rule for displaying a search tree is presented in Figure 5.29.

5.5.6 Summary

We have described the problem of parallel program animation in our previous model and the solution. In our previous model, it was difficult to overlap unit animations. This problem is alleviated in our new model by allowing programmers to specify the start and end events of unit animations. We have also described the TRIP2a/3D system based on the new model and an example animation created with TRIP2a/3D.

5.6 Conclusions

We have proposed the extended bi-directional translation model for animations, and have implemented a prototype system based on this model. Animations can be created by providing visual and

---

[1]This program was written in Schematic[122].
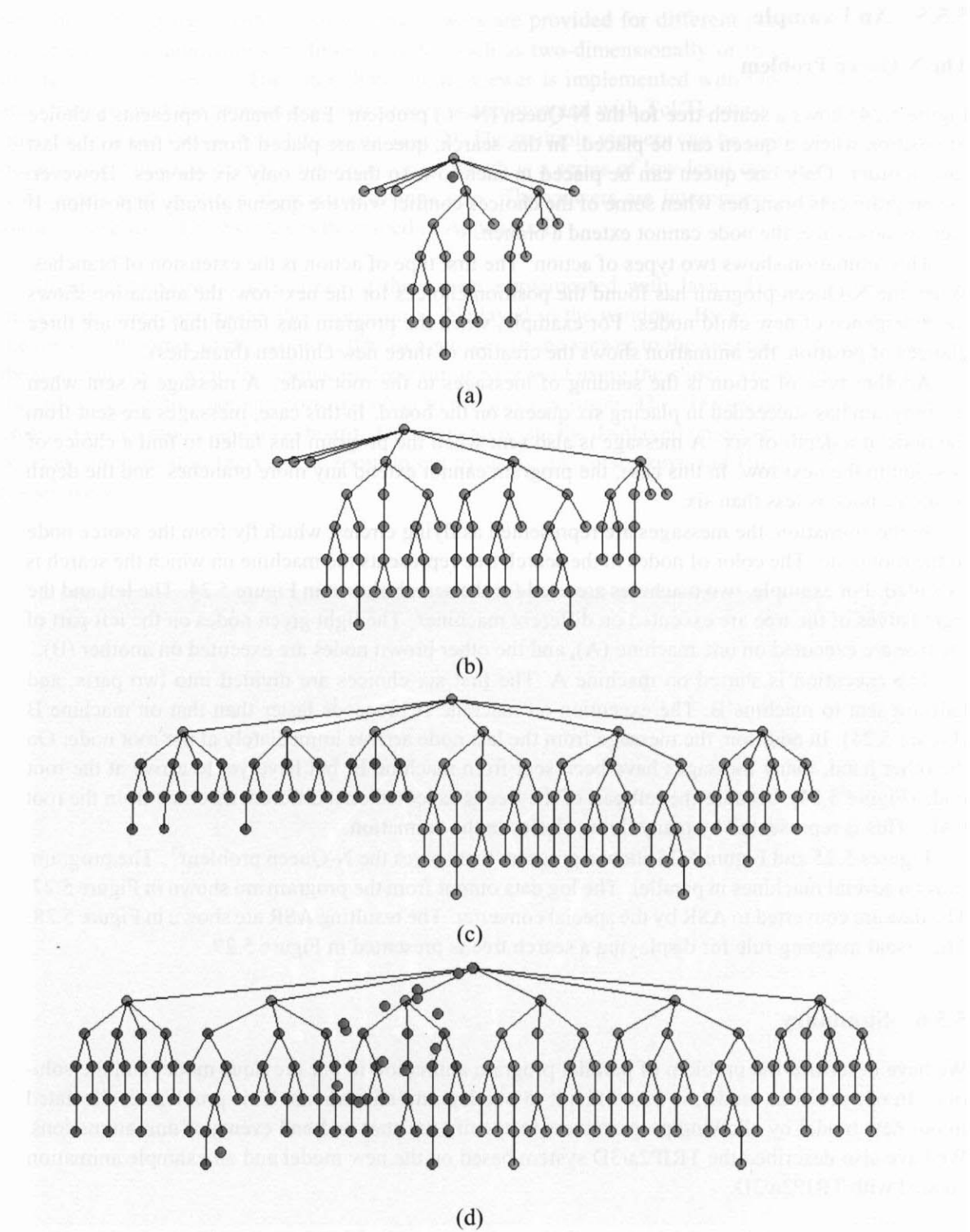
(a)

(b)

(c)

(d)

Figure 5.24: Screenshot of an animation that shows a search tree for the N-Queen problem.

```
;;;
;;; N-Queen Probelm
;;;
(module schematic-user)
(include "schematic2.sch")
(include "counter.sch")
(include "logger.sch")

(definep (n-queen-sub c j N M soremade l)
  (cond ((> j N)
         (begin
           (now (output-log-with-time l "~s ~s ~s () šend~%"
              (list *pe* j soremade)))
           (add! c 1)
           (now (output-log-with-time l "~s ~s ~s () šent~%"
              (list *pe* j soremade)))
           1))
        ((= j 1)
         (let ((candidates (safe-place-list 1 1 N soremade)))
           (if (null? candidates)
               (begin
                 (now (output-log-with-time l "~s ~s ~s ~s šend~%"
                    (list *pe* j soremade candidates)))
                 (add! c 1)
                 (now (output-log-with-time l "~s ~s ~s ~s šent~%"
                    (list *pe* j soremade candidates))))
               (now (output-log-with-time l "~s ~s ~s ~s~%"
                  (list *pe* j soremade candidates))))
           (psum (map (plambda (i)
                        (let ((processor
                               (quotient
                                (- N i)
                                (inexact->exact (ceiling (/ N M))))))
                          (future (n-queen-sub c 2 N M (cons i sore-
made) l)
                            :on processor)))
                 candidates)))))
        ((or (= j 2) (= j 3))
         (let ((candidates (safe-place-list 1 j N soremade)))
           (if (null? candidates)
               (begin
                 (now (output-log-with-time l "~s ~s ~s ~s šend~%"
                    (list *pe* j soremade candidates)))
                 (add! c 1)
                 (now (output-log-with-time l "~s ~s ~s ~s šent~%"
                    (list *pe* j soremade candidates))))
               (now (output-log-with-time l "~s ~s ~s ~s~%"
                  (list *pe* j soremade candidates))))
           (psum (map (plambda (i)
                        (future (n-queen-sub c (+ j 1) N M (cons i sore-
made) l)))
                 candidates)))))
        (else
         (let ((candidates (safe-place-list 1 j N soremade)))
           (if (null? candidates)
               (begin
                 (now (output-log-with-time l "~s ~s ~s ~s šend~%"
                    (list *pe* j soremade candidates)))
                 (add! c 1)
                 (now (output-log-with-time l "~s ~s ~s ~s šent~%"
                    (list *pe* j soremade candidates))))
               (now (output-log-with-time l "~s ~s ~s ~s~%"
                  (list *pe* j soremade candidates))))
           (sum (map (plambda (i)
                       (n-queen-sub c (+ j 1) N M (cons i soremade) l))
                 candidates))))))))
```

Figure 5.25: The target parallel N-queen-solving program (1).

```
(definep (safe-place? i j y soremade)
  (if (null? soremade) #t
      (let ((x (car soremade)))
        (cond ((= i x) #f)
              ((= (- i j) (- x y)) #f)
              ((= (+ i j) (+ x y)) #f)
              (else
                (safe-place? i j (- y 1) (cdr soremade)))))))

(definep (safe-place-list i j N soremade)
  (if (> i N) '()
      (if (safe-place? i j (- j 1) soremade)
          (cons i (safe-place-list (+ i 1) j N soremade))
          (safe-place-list (+ i 1) j N soremade))))

(definep (psum list)
  (if (null? list) 0 (+ (touch (car list)) (psum (cdr list)))))

(definep (sum list)
  (if (null? list) 0 (+ (car list) (sum (cdr list)))))

(definep (n-queen N)
  (let* ((c  (counter 0))
         (t0 (current-time))
         (n  (n-queen-sub c 1 N *npe* '() (logger)))
         (t1 (current-time))
         (m  (get-value c)))
    (format #t "~s ~s : ~s msec~%" n m (- t1 t0))))

(definep (schematic-main args)
  (n-queen (string->number (cadr args)))
  (newline))

;;;
;;; counter.sc---Simple Counter class
;;;

(define-generic (add! self x))
(define-generic (get-value self))

(define-class counter ()
  value)

(define-method! counter (add! self x)
  (become value :value (+ x value)))

(define-method counter (get-value self)
  value)

;;;
;;; an object for logging
;;;

(define-class logger ())

(define-generic (output-log-with-time self log-format args))
(define-generic (output-log self log-format args))

(define-method logger (output-log-with-time self log-format args)
  (format #t "~s  " (current-time))
  (apply format `(#t ,log-format ,@args)))

(define-method logger (output-log self log-format args)
  (apply format `(#t ,log-format ,@args)))
```

Figure 5.26: The target parallel N-queen-solving program (2).

```
30659177   0 1  ()  (1 2 3 4 5 6)
30659179   0 2  (4)  (1 2 6)
30659179   0 3  (1 4)  (3 5)
30659179   0 4  (3 1 4)  (5 6)
30659180   0 5  (5 3 1 4)  (2)
30659180   0 6  (2 5 3 1 4)  ()
30659180   0 5  (6 3 1 4)  (2)
30659180   0 6  (2 6 3 1 4)  ()
30659181   1 2  (1)  (3 4 5 6)
30659181   1 2  (2)  (4 5 6)
30659182   1 2  (3)  (1 5 6)
30659183   1 3  (3 1)  (5 6)
30659184   0 2  (5)  (1 2 3)
30659184   0 3  (1 5)  (4 6)
30659185   0 4  (4 1 5)  (6)
30659185   0 5  (6 4 1 5)  (3)
30659185   0 6  (3 6 4 1 5)  ()
  ...omitted...
```

Figure 5.27: An example of log data.

```
node(t0,red).
node(t1,red).
node(t2,red).
node(t3,red).
node(t4,red).
node(t5,red).
node(t6,red).
trees([tree(t0,[t1,t2,t3,t4,t5,t6])
]).
end_of_state.
node(t0,red).
node(t1,red).
node(t2,red).
node(t3,red).
node(t4,red).
node(t10,red).
node(t16,red).
node(t40,red).
node(t5,red).
node(t6,red).
trees([tree(t0,[t1,t2,t3,t4,t5,t6])
,tree(t4,[t10,t16,t40])
]).
end_of_state.
node(t0,red).
node(t1,red).
node(t2,red).
node(t3,red).
node(t4,red).
node(t10,red).
node(t118,red).
node(t190,red).
node(t16,red).
node(t40,red).
node(t5,red).
node(t6,red).
trees([tree(t0,[t1,t2,t3,t4,t5,t6])
,tree(t4,[t10,t16,t40])
,tree(t10,[t118,t190])
]).
end_of_state.
...omitted...
```

Figure 5.28: The ASR data for N-queen animation.

```
:- module vmr.

rule(default, Result) :- true | Result = [].

rule(node(X), Result) :- true |
        Result = [sphere(X, 5, [material([ambient(0.2, 0.5, 0.4)])])].
rule(node(X, red), Result) :- true |
        Result = [sphere(X, 5, [material([ambient(0.8, 0.5, 0.4)])])].
rule(node(X, blue), Result) :- true |
        Result = [sphere(X, 5, [material([ambient(0.2, 0.2, 0.8)])])].
rule(node(X, green), Result) :- true |
        Result = [sphere(X, 5, [material([ambient(0.2, 0.8, 0.2)])])].
rule(send(Mess, Id, From), Result) :-
        Result = [move(Mess, [from(From, Id)])].
rule(sent(Mess, Id, To), Result) :-
        Result = [move(Mess, [to(t0, Id)])].

rule(trees(X), Result) :- true |
        treesmap(X, Res1),
        layout_leaf(X, Res2),
        generic:new(merge, {Res1, Res2}, Result).

%%%---------------------------------------------

treesmap([], Result) :- true | Result = [].
treesmap([tree(X, L)|T], Result) :- true |
        L = [H|_],
        treesmap(T, RestResult),
        connections(X, L, Connections),
        TmpResult = [x_parallel(L, []),
                     x_average(X, L, []),
                     z_relative([X, H], 0, []),
                     y_relative([X, H], -30, [])],
        generic:new(merge, {Connections, RestResult, TmpResult}, Result).


connections(X, [], Connections) :- true | Connections = [].
connections(X, [H|T], Connections) :- true |
        Connections = [line(l(X, H), 1, []), con-
nect(l(X, H), X, H, [])|Rest],
        connections(X, T, Rest).


layout_leaf(List, Result) :- true |
        List = [tree(X, L)|Rest],
        makeLeafList([X], List, Leafs),
        trip:flat(Leafs, FlatLeafs),
        makeLeafConstraints(FlatLeafs, Relatives),
        %Result = [place(X, 0, 0, 0, []), x_relative(FlatLeafs, 15, [])].
        Result = [place(X, 0, 0, 0, [])|Relatives].

makeLeafConstraints([_], R) :- true | R = [].
makeLeafConstraints([X, Y|T], R) :- true |
        R = [x_relative([X, Y], 15, [])|R1],
        makeLeafConstraints([Y|T], R1).


makeLeafList([], Trees, Result) :- true | Result = [].
makeLeafList([Node|Rest], Trees, Result) :- true |
        searchTree(Node, Trees, Tree),
        subMakeLeafList(Node, Tree, Trees, R1),
        makeLeafList(Rest, Trees, R2),
        generic:new(merge, {R1, R2}, Result).

subMakeLeafList(Node, nil, Trees, R) :- true | R = [Node].
otherwise.
subMakeLeafList(Node, Tree, Trees, R) :- true |
        Tree = tree(X, L),
        makeLeafList(L, Trees, R).

searchTree(Node, [], Tree) :- true | Tree = nil.
searchTree(Node, [tree(Node, L)|Rest], Tree) :- true | Tree = tree(Node, L).
otherwise.
searchTree(Node, [T|Rest], Tree) :- true | searchTree(Node, Rest, Tree).
```

Figure 5.29: The visual mapping rule set for the N-queen animation.

transitional mapping rules, and by annotating an application program so that the data and the operation of the application are passed to the animation system. We have applied this model to several algorithms, such as sorting algorithms, the tower of Hanoi, and Kruskal's algorithm for finding a minimum spanning tree. Our experience has shown that animations can be created quite effortlessly.

In future studies, we plan to provide a way to specify animations by demonstration using the techniques of TRIP3/IMAGE[88, 87] described in Appendix A.