1.

A Visual Routines Based Model of Graph Understanding

Yusuf Pisan
The Institute for the Learning Sciences
Northwestern University
y-pisan@nwu.edu

Abstract

We present a model of graph understanding and describe our implementation of the model in a computer program called SKETCHY. SKETCHY uses a combination of general graph knowledge and domain knowledge to describe graphs, answer questions, perform comparative analyses, and detect contradictions in problem solving assumptions. SKETCHY has generated reasonable graph summaries for 65 graphs from multiple domains. SKETCHY illustrates the robustness of our model of graph understanding.

Introduction

Understanding diagrams is an important part of human cognition, requiring integration of perceptual information and conceptual knowledge. Diagrams are used to solve problems, to give explanations, to summarize information and to represent spatial relations. Diagrams serve both as devices to aid in visualization of the situation and as short-term fast access memory devices for holding information (Larkin & Simon, 1987). Diagrams have been successfully integrated with computer programs to explain complex mechanical and dynamic systems (Forbus, Nielsen & Faltings, 1991; Kim, 1993). Diagram comprehension requires being able to identify objects, determine the relevant features for a particular problem and map the graphical features to the domain.

A graph is a specialized form of diagrammatic representation. Previous psychological research (Gattis & Holyoak, 1994; Pinker, 1990; Schiano & Tversky, 1992) shows that graphs form a symbolic system different than pictures with their own set of symbols and rules. Different graph formats emphasize different relationships between variables. For instance, pie graphs are used to show percentages, bar graphs and step graphs to show relative amounts, scatter plots to show trends in data and line graphs to show continuous changes. In this paper, we only consider line graphs.

We present a model of graph understanding and describe our implementation of the model in a computer program called SKETCHY. SKETCHY uses a combination of general graph knowledge and domain knowledge to describe graphs, answer questions, including comparative analyses, and detect contradictions in problem solving assumptions. SKETCHY has generated reasonable interpretations for all the graphs in a college level thermodynamics textbook (Whalley, 1992) as well as interpretations for a number of graphs from economics (Ekelund & Tollison, 1986).

Section 2 presents our model of graph understanding. Section 3 gives examples from SKETCHY, Section 4 discusses relevant work on graphs in psychology and vision and Section 5 describes possible extensions to the model and to the computer program SKETCHY.

A Model of Graph Understanding

Understanding graphs is a subset of the general problem of understanding diagrams. As such, graph understanding requires reasoning about spatial properties and relations and interpreting them in conceptual terms. Unlike general diagrams, graphs are composed of a small set of primitives (axes, lines, points, areas and labels), which simplifies object recognition. In a graph, points, lines and areas represent conceptual relationships in the domain. By characterizing the possible relationships among graph objects, we have constructed a model of graph understanding that is not tied to a specific domain.

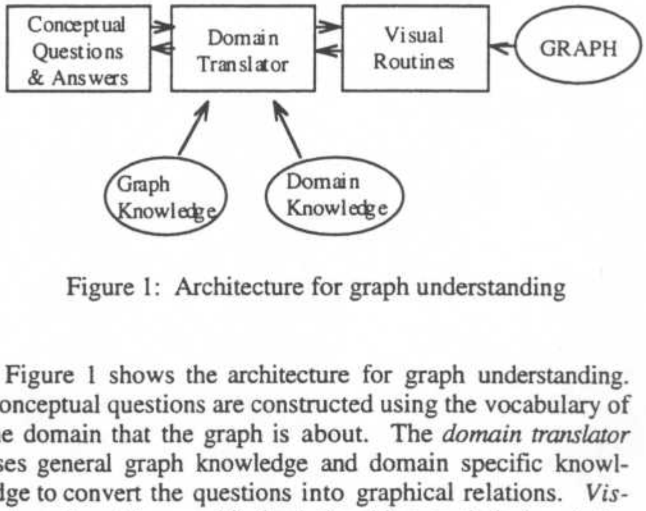


Figure 1: Architecture for graph understanding

Figure 1 shows the architecture for graph understanding. Conceptual questions are constructed using the vocabulary of the domain that the graph is about. The *domain translator* uses general graph knowledge and domain specific knowledge to convert the questions into graphical relations. *Visual routines* take graphical relationships as their input, inspect the graph to gather the necessary information and return the information to the domain translator. Depending on the results the domain translator might initiate other visual routines to answer the question. When all the necessary information is obtained from the graph, the domain translator converts the graphical relationships into the vocabulary of the domain and generates an answer to the question. This paper examines the information processing necessary for graph understanding. We ignore the problem of recognizing an image as a specific type of graph (Pinker, 1990) and how visual routines can be implemented (Ullman, 1984) as these problems have been addressed by other researchers.

Domain Translator

Two kinds of knowledge are needed when translating a question from conceptual terms to graphical relations: general graph knowledge and domain specific knowledge. For example, to answer the question "When is SUPPLY equal to DEMAND?" the domain translator first needs to identify what objects are being referred to by SUPPLY and DEMAND. The graph labels serve as the necessary semantic information connecting the graph objects to the concepts in the domain. The domain translator initiates visual routines which inspect the graph to find the objects with labels SUPPLY and DEMAND. If no objects with those labels are found, domain knowledge is used to connect SUPPLY and DEMAND to the graph objects present.

Graph conventions make up an important part of general graph knowledge. When there are no scales on the axes, lines going up and to the right are interpreted as having a positive slope and signifying that variables on the axes are qualitatively proportional to each other. Steeper lines are interpreted as showing relations where the variable represented on the Y axis is increasing faster. Two regions with equal areas are interpreted as being equal in magnitude. Although domain specific knowledge can override graph conventions, graphs in most domains follow graph conventions closely. As a result, general graph knowledge can be applied to new domains to produce reasonable graph interpretations even when there is very little or no domain knowledge.

General graph knowledge also guides in identifying the important features of a graph. An image can be described in an infinite number of ways, so people use heuristics for summarizing graphical information, some general and some specific to task or domain. Some of the heuristics that we have observed people to use (and are implemented in SKETCHY) are:

- Only include information for objects with labels.
- Include coordinates of labeled points if the axes have scales.
- If a point is on a line or on the border of an area, include this information.
- Include information about any qualitative changes in line slopes and describe each qualitative region separately.
- If lines intersect, include this in the graph description.
- Mention changes due to modifications.

Each line in a graph represents a different relationship between the variables on the axes. For example, the supply line represents how the amount produced increases with increasing prices and the demand line represents how the amount demanded decreases with increasing prices. Intersection of two lines represents a point of equality between two relationships, often representing important values in the domain, and is always included in graph summaries. In the supply and demand example, the intersection point represents the equilibrium point for the market determining the current price. Qualitative changes in line slopes are included in the summary since a change in line direction represents a change in the type of relation between the variables. Points usually represent important domain specific values and are included in the graph summary.

Graphs provide a natural way of performing comparative analysis (Weld, 1990) by combining qualitative and quantitative information. Comparative analysis is the problem of predicting how a system will react to perturbations in its parameters. Purely qualitative techniques for comparative analysis, such as the methods used by Weld, are limited in their prediction capacity because the net effect of opposing influences cannot be determined. In graphs, lines carve up the two-dimensional space defining qualitative regions, which enable qualitative analysis while still maintaining access to numerical values. In Section 3, we present an example of how SKETCHY performs comparative analysis.

The domain translator uses the visual routine processor to extract information from the diagram. It begins by calling visual routines that identify entities in the graph. If the entities are not found, domain knowledge is used to suggest other graphical interpretations. Then it uses other visual routines to compute relationships between the objects based on the query. These relationships are then translated back into conceptual terms to produce an answer to the question.

Visual Routines

After the conceptual question is translated into graphical terms by the translator, visual routines are invoked to gather the necessary information from the graph. Ullman (1984) suggests how psychologically plausible elemental operations (such as bounded activation and boundary tracing) can be combined to construct visual routines. Visual routines are used to retrieve coordinates of objects, determine spatial orientations, find about interactions, and get information about size and changes in the graph.

Table 1: Examples of visual routines and how they are used

Visual Routine	Example of Use
examine label	Used to find the object being queried
coordinate-at-point	For calculating slope, getting the value of a point
right-of, left-of, above, below	Used for finding spatial relations of objects to each other. Necessary when axes do not have scales
inside, outside	Used for determining the relationship between an area and a point or line segment
steeper, flatter	For comparing slopes of lines qualitatively
bigger, smaller	Comparing sizes
vertical, horizontal	Special cases for line slope being zero or infinity
change-in-slope	For dividing lines into regions
touches, intersects	Possible relationship between objects
on-line, on-border, forms-border	Specifying a limit point either for an area or a line

Table 1 shows the visual routines used to interpret graphs. The visual routines in Table 1 are given in terms of object pairs, but they can also be used to find objects that satisfy a specific relationship.

Examples from SKETCHY

SKETCHY is a fully implemented computer program based on our model of graph interpretation. Given a graph produced by a simple interface, SKETCHY can provide natural summarizations, answer questions, perform comparative analyses, and detect contradictions in problem solving assumptions. SKETCHY has been fully tested on 65 graphs from two domains (economics and engineering thermodynamics), which suggests that the model is robust. This section illustrates SKETCHY's operation on representative examples, to better show how the model works.

Graph Summarization

Figure 2 shows a graph from a thermodynamics textbook. Understanding this graph is essential for solving many thermodynamics problems since all substances exhibit the same qualitative behavior shown. The graph shows three regions (liquid, liquid/vapour, and vapour regions) corresponding to the phase(s) a substance can be in. The temperature lines, which are contours of equal temperature, effectively add a third dimension to the graph. SKETCHY produces the graph description given in Figure 3 using general graph knowledge and graph labels, but without in-depth domain knowledge about temperature, pressure, volume or the phases a substance can be in.

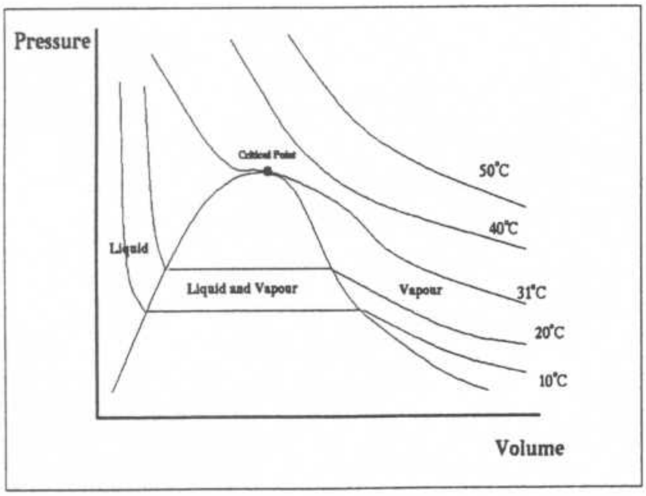


Figure 2: Compression of carbon dioxide

SKETCHY's summary captures important features of the graph, but it contains more information than a person might give in explaining the graph to someone else. Part of becoming an expert in the domain is learning how to concisely state the relevant features of a graph for the current task. Including task specific control information would make SKETCHY's summary more concise.

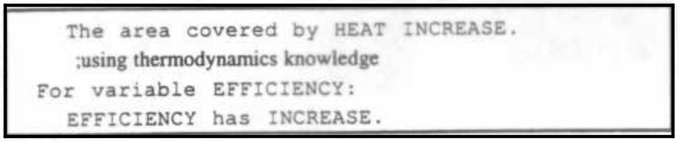
```
  For line 31-C:
    VOLUME and PRESSURE are inversely propor-
tional.
  For line 20-C:
    The slope of 20-C has discontinuities
    ;associating discontinuities with regions
    Inside region LIQUID:
      VOLUME INCREASE and PRESSURE DECREASE.
    Inside region LIQUID-AND-VAPOUR:
      VOLUME INCREASE and PRESSURE CONSTANT.
    Inside region VAPOUR:
      VOLUME INCREASE and PRESSURE DECREASE.
  CRITICAL-POINT is on lines (31-C)
  CRITICAL-POINT is on regions (LIQUID LIQUID-
AND-VAPOUR VAPOUR)
  For TEMPERATURE contour:
    As TEMPERATURE increases
      the slopes of TEMPERATURE lines become
more LINEAR.
    ;basis for Boyle's Law
    For a constant PRESSURE:
      As VOLUME increases TEMPERATURE INCREASE.
      VOLUME and TEMPERATURE are directly pro-
portional.
    For a constant VOLUME:
      As PRESSURE increases TEMPERATURE IN-
CREASE.
      PRESSURE and TEMPERATURE are directly pro-
portional.
```

Figure 3: SKETCHY's description of carbon dioxide compression graph

Comparative Analysis

Graphs are an ideal representation for comparative analysis since they combine qualitative and quantitative information. SKETCHY demonstrates comparative analysis can be done via visual processes on a graph. Analyzing engineering cycles is an important task in thermodynamics. The basic cycle for a steam power plant is the Rankine cycle, shown in Figure 5. A common modification to the Rankine cycle is superheating the steam in the boiler to increase the efficiency of the cycle. The net work of the cycle before modification is represented by area 1-2-3-4-1 and after modification by 1-2-3'-4'-1. The area under 1-2-3-3' represents the total heat put into the system.

```
  ;using graph interpretation rules
  For point 3:
    The ENTROPY of 3 INCREASE.
    The TEMPERATURE of 3 INCREASE.
  For point 4:
    The ENTROPY of 4 INCREASE.
    The TEMPERATURE of 4 CONSTANT.
  For region WORK:
    The area covered by WORK INCREASE.
  For region HEAT:
```



```
   The area covered by HEAT INCREASE.
   ;using thermodynamics knowledge
 For variable EFFICIENCY:
   EFFICIENCY has INCREASE.
```

Figure 4: SKETCHY's explanation

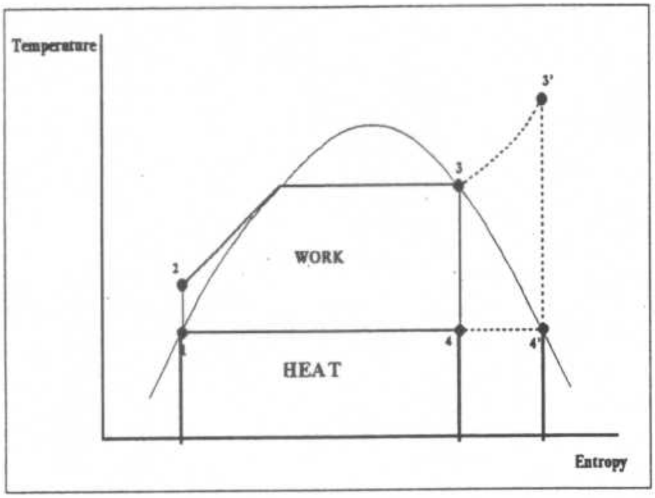


Figure 5: Effect of superheating on Rankine cycle

Qualitative methods alone are sufficient to reach the conclusion that WORK and HEAT have increased as a result of modification. Efficiency, defined as the amount of work divided by the amount of heat, is represented indirectly through work and heat as areas in the graph. Determining whether efficiency has increased or not cannot be resolved qualitatively. SKETCHY uses visual routines to calculate the changes in areas and determines that the efficiency of the cycle is increased.

Using SKETCHY in Problem Solving

We have connected SKETCHY to CyclePad (Forbus & Whalley, 1994) an intelligent learning environment for engineering thermodynamics. An important problem in such learning environments is detecting contradictory student assumptions and explaining them in an easily grasped fashion. SKETCHY uses student-supplied assumptions and numerical values computed by CyclePad to automatically draw temperature-entropy diagrams. Students can express design changes using these diagrams. Modifications to CyclePad's parameters that lead to visually detectable contradictions are found by SKETCHY's thermodynamics domain rules and it warns the student about them (c.f. Figure 7).

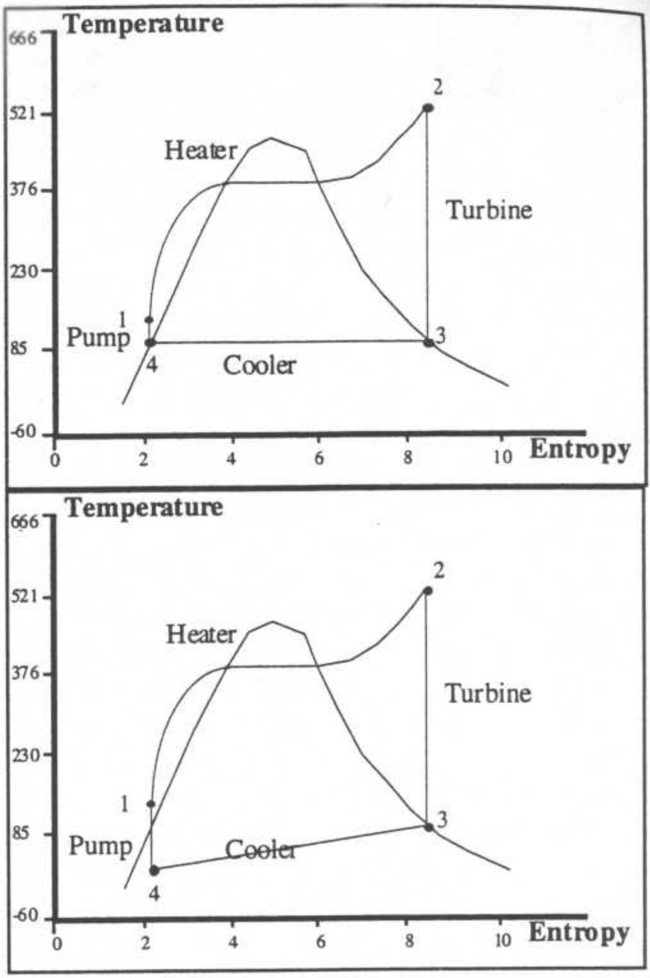
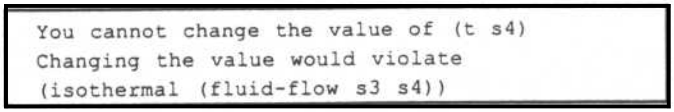


Figure 6: The graph before and after user modification



```
You cannot change the value of (t s4)
Changing the value would violate
(isothermal (fluid-flow s3 s4))
```

Figure 7: SKETCHY's report of detecting the contradiction

Related Work

One inspiration for SKETCHY is the Metric Diagram/Place Vocabulary model of spatial reasoning (Forbus, 1980; Forbus, Nielsen & Faltings, 1991). SKETCHY's Visual Routine Processor is its Metric Diagram.

Ullman (1984) introduced the concept of visual routines as a goal-oriented visual processing facility. Visual routines express domain-specific visual skills. Mahoney (1992) extends Ullman's work by defining image chunks, formed using topological information, that can be used for higher level goals. In SKETCHY we ignore the problem of recognizing and identifying graph objects and concentrate on interpreting their interactions. A natural extension to SKETCHY would be implementing image chunks, which would enable SKETCHY to analyze scanned images. This extension would not fundamentally alter our model of graph understanding.

POLYA (McDougal & Hammond, 1993) uses visual operators to specify which objects in the diagram to inspect in the course of solving geometry proofs. POLYA's operators are very specific to the geometry domain (such as LOOK-AT-LEFT-BASE-ANGLE). SONJA (Chapman, 1991) on the other hand uses very general action oriented visual operators for playing a video game. SKETCHY's operators are specific for examining line graphs.

Pinker (1990) describes psychological factors contributing to difficulty in reading graphs. Pinker suggests a similar architecture to SKETCHY, but his main emphasis is on recognition of different graph types through general graph schemas and the difficulties in understanding different graphs, rather than providing a concrete computational model for graph interpretation. Currently SKETCHY does not have any internal model for processing capacity or selective attention, both of which would be useful in increasing its psychological plausibility.

Gattis and Holyoak (1994) look at the impact of goals and conceptual understanding on graph interpretation. Gattis and Holyoak's most significant finding is that the variable being queried should be assigned to the vertical axis, so that steeper lines can map to faster changes in the queried variable. We view this result as further evidence that graph semantics and graph interpretation is separate from the domain the graph is about.

Lohse (1993) describes a computer program called UCIE which uses graph schemas to predict response times to answer questions about the graph. UCIE's graph schemas for information retrieval are similar to SKETCHY's general graph knowledge. UCIE's short-term and long-term memory models could be incorporated into SKETCHY to get similar response time predictions.

SKETCHY's graph descriptions are mainly produced by domain independent graph rules. Tabachneck, Leonardo and Simon (1994) demonstrate how novices have difficulty integrating visual and verbal information. Novices fail to provide answers that could be obtained by simple perception whereas experts see the answer immediately. When domain rules are not used, SKETCHY suffers from a similar problem. SKETCHY cannot answer any questions about variables besides the ones explicitly mentioned on the graph even when the answer is visually available. Part of becoming an expert in a domain is creating the necessary domain rules, so that inferences about objects not labeled in the graph can be made.

Discussion

We have presented a model for interpreting graphs and illustrated its capabilities via examples solved by SKETCHY, a computer implementation of the model. SKETCHY has generated reasonable interpretations for 65 graphs from thermodynamics and economics showing that our model is broadly applicable.

Extending SKETCHY to other graph types such as bar graphs and pie charts appears straightforward. The major difficulty appears to be increasing the library of visual routines to recognize and compare these compound graphical elements. Extending our model to general diagrams would require developing functional representations for objects that will be in the diagrams. Currently we are incorporating SKETCHY into a new cognitive simulation of student problem solving in engineering thermodynamics.

Acknowledgements

Special thanks to my advisor Ken Forbus for his comments and encouragement. This research was funded by the Computer Science Division of the Office of Naval Research.

References

Chapman, D. (1991). *Vision, Instruction, and Action*. Cambridge, MA: MIT Press.

Ekelund, R. B. & Tollison, R. D. (1986) *Economics*. Boston: Little, Brown.

Gattis, M. & Holyoak, K. J. (1994). How graphs mediate analog and symbolic representation. In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 346-350). Hillsdale, NJ: Lawrence Erlbaum Associates.

Forbus, K. (1980) Spatial and qualitative aspects of reasoning about motion. In *Proceedings of the First Annual AAAI Conference* (pp. 170-173). Los Altos, CA: William Kaufmann Inc.

Forbus, K., Nielsen P. & Faltings, B. (1991) Qualitative spatial reasoning: the clock project. *Artificial Intelligence*, 51, 417-471.

Forbus, K. & Whalley, P. B. (1994). Using qualitative physics to build articulate software for thermodynamics education. In *Proceedings of the Twelfth Annual AAAI Conference* (pp. 1175-1182). Menlo Park, CA: AAAI Press/MIT Press

Kim, H. (1993) Qualitative reasoning about fluids and mechanics. Doctoral Dissertation. Urbana-Champaign: University of Illinois at Urbana-Champaign.

Kosslyn, S. M. (1989). Understanding charts and graphs. *Applied cognitive psychology*, 3, 185-226.

Larkin, J. H. & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65-99.

Lohse, G. L. (1993). A cognitive model for understanding graphical perception. *Human-Computer Interaction*, 8 (4), 353-388.

Mahoney, J. V. (1992). Image chunks and their applications. Technical Report EDL-92-3, Xerox Parc, Palo Alto, CA.

McDougal, T. F & Hammond, K. J. (1993). Representing and using procedural knowledge to build geometry proofs. In *Proceedings of the Eleventh Annual AAAI Conference*, (pp. 60-65). Menlo Park, CA: AAAI Press/MIT Press.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73-126). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schiano, D. J. & Tversky, B. (1992). Structure and strategy in encoding simplified graphs. *Memory & Cognition*, 20, 12-20.

Tabachneck, H., Leonardo A. M. & Simon, H. A. (1994). How does an expert use a graph? A model of visual and verbal inferencing in economics. In *Proceedings of the Sixteenth Annual conference of the Cognitive Science Society* (pp. 842-847). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ullman, S. (1984). Visual Routines. *Cognition*, 18, 97-159.

Weld, D. S. (1990). *Theories of comparative analysis*. Cambridge, MA: The MIT Press.

Whalley, P. B. (1992). *Basic Engineering Thermodynamics*. Oxford, NY: Oxford University Press.