Integrating Probabilistic Reasoning into a Symbolic Diagrammatic Reasoner

Joseph L. Bokor and Ronald W. Ferguson

College of Computing
Georgia Institute of Technology
801 Atlantic Avenue
Atlanta, GA 30332
{jlbokor, rwf} @cc.gatech.edu

Abstract

An important part of diagram understanding is the problem of glyph recognition. This problem is challenging because a glyph may be drawn in many different ways and with varying levels of precision. Diagrams in general are rarely perfect and often filled with poorly-drawn glyphs. In order to understand a wide range of diagrams, a diagrammatic reasoner should recognize glyphs that deviate from standard forms. This paper presents a new mechanism for glyph recognition which combines a probabilistic representation with an existing symbolic diagrammatic reasoner. This reasoner, GeoRep, recognizes diagram glyphs using a set of rules supported by a logic-based truth-maintenance system (LTMS). Here we extend GeoRep's LTMS to include nodes that encapsulate Bayesian networks. The result is a reasoner that can leverage the benefits of symbolic truth maintenance as well as that of probabilistic networks.

Introduction

Diagrams are important to a wide variety of tasks that include problem solving, communication, and collaboration (Glasgow, Narayanan, and Chandresekaran, 1995). These tasks are frequently aided by a diagram's ability to capture and convey many spatial relations at a glance.

In most cases, a crucial first step in diagram understanding is to combine the initial visual primitives into visual symbols or *glyphs*. Glyphs include such things as NAND-gates and AND-gates in digital logic diagrams.

Because glyphs are important for diagram understanding, the problem of correctly recognizing the glyphs is critical. Glyphs, however, are often created with limited precision, especially when drawn quickly, because it is assumed that others will still be able to identify the glyph being drawn. For example, Figure 1 contains a set of glyphs that are all easily interpreted as a NAND-gate, even though only (a) is drawn precisely.

Importantly, the type of imprecision in each glyph is not the type that is captured in a typical qualitative representation. A qualitative spatial reasoner, such as GeoRep, can recognize a glyph properly when the imprecision is captured by qualitative relationships. For example, in the NAND-gate, differently-sized arcs and segments do not confuse GeoRep as long as the basic qualitative relations, such as perpendicular corners, are preserved. However, in the imprecise glyphs these corners may in fact be intersections, may not be perpendicular, or may not be corners at all. GeoRep uses tolerance values to ensure that at least some of the spatial relations are captured for such imprecise glyphs. However, usually at least some of the expected qualitative relations lie outside those tolerances, and so may block glyph recognition. For this reason, these types of imprecision are not captured in a qualitative representation.

Therefore the problem is one of ambiguity at the level of the set of visual relations, rather than at the level of individual visual elements or relations.

The GeoRep reasoner (Ferguson & Forbus, 2001; Ferguson et al., 2003) uses a logic-based truth maintenance system (LTMS; Forbus and de Kleer, 1993; McAllester, 1990) and a rule engine to produce high-level spatial description of a diagram. While this system can interpret diagrams in domains with large glyph vocabularies, such as military Course-of-Action diagrams (Ferguson et al., 2000) the glyphs are required to be drawn with relatively high precision. Deviations from the standard glyph such as those shown in (b), (c), and (d) in Figure 1 are unrec-

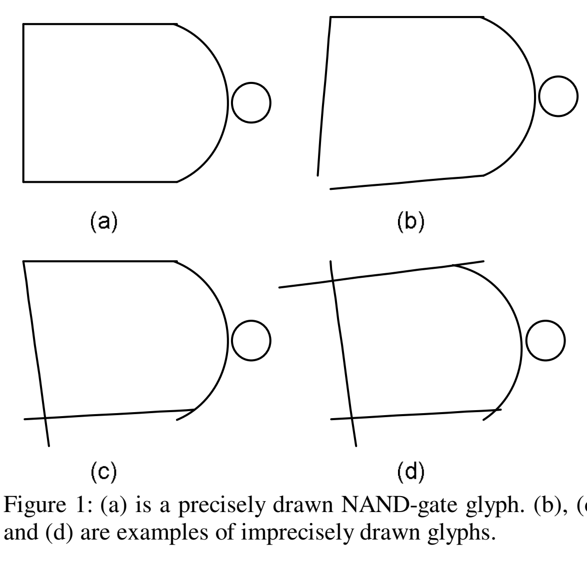


Figure 1: (a) is a precisely drawn NAND-gate glyph. (b), (c) and (d) are examples of imprecisely drawn glyphs.

ognizable by standard rules. In addition, it would be impractical to create rules for each possible variation of a glyph, since the set of possibilities is combinatorial.

In order to achieve more robust glyph recognition, a new method for recognizing glyphs is necessary. Various forms of probabilistic recognizers have been used successfully (Alvarado, Oltmans, & Davis, 2002; Cohen et al., 1997; Gross, 1996). This paper discusses the integration of a new probabilistic mechanism into the GeoRep architecture. The mechanism uses Bayesian networks encapsulated within a new probabilistic node in the LTMS. The use of Bayesian networks provides a new representation for sets of visual elements that can capture imprecision and uncertainty inherent in drawn diagrams. The encapsulation of these networks in the existing LTMS system allows the two systems to combine their strengths and both participate in the process of glyph recognition.

The next section describes the GeoRep diagrammatic reasoner. We then demonstrate how probabilistic reasoning is integrated, present some results, and compare these results to the original system.

The GeoRep Reasoner

GeoRep is a diagrammatic reasoner (Ferguson & Forbus, 2000). The input to GeoRep is the vector graphics representation of visual elements. Element types include line segments, circles, ellipses, arcs, spline curves, and positioned text. The most recent version of GeoRep (Ferguson et al., 2003) can also process visual elements dynamically as they are added to the diagram. GeoRep's output is a predicate calculus representation as given in Figure 2.

GeoRep creates this representation using a two-stage architecture (Figure 3). The first stage, the low-level relational describer (LLRD) represents a set of low-level visual relations. These relations are stored in a dependency network to allow quick revisions. Each network node in the network represents a set of alternative spatial relations between two visual elements. These sets, like many found in qualitative spatial vocabularies, are jointly-exhaustive and pairwise-disjoint (JEPD; Cohn, 1997).

The particular spatial relations captured in the LLRD's representation are designed to model those qualitative

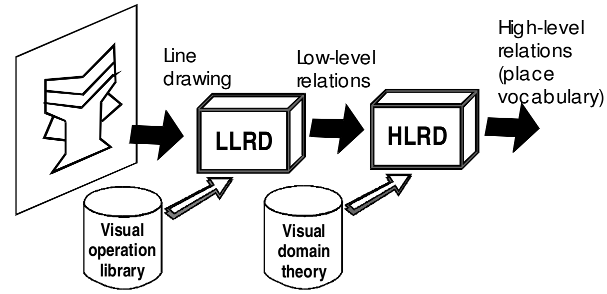


Figure 3: Simplified GeoRep architecture, containing stages for low-level and high-level visual reasoning.

spatial relations detected in early vision. For example, humans are sensitive to relative angles (such as perpendicular lines), indentations in figure boundaries (Hoffman & Richards, 1984), and to vertical and horizontal orientations in the assumed frame of reference (Rock, 1973).

The second stage, the High-Level Relational Describer (HLRD) uses these low-level relations and a rule-based *visual domain theory* to produce a description of the diagram. The output of the HLRD describes the diagram using domain-dependent high level relations.

The HLRD rules use a pattern-directed inference system that is supported by an LTMS. HLRD rules use the LLRD's low-level visual relations as well as domain knowledge. For example, to recognize the `(system-output wire-s11)` in the representation below, the system uses a rule that infers an output wire when the system detects a NAND-GATE and a proximate line, and that line connects to the circle of the NAND-GATE. Many of these rules thus combine spatial and domain-specific reasoning. In general, HLRD rules are constrained to run only on proximate visual elements to keep processing tractable.

Implementation

We have created a new mechanism for glyph recognition by extending the capabilities of the GeoRep reasoner and introducing Bayesian networks. Bayesian networks have been shown to be useful in similar recognition tasks (Al-

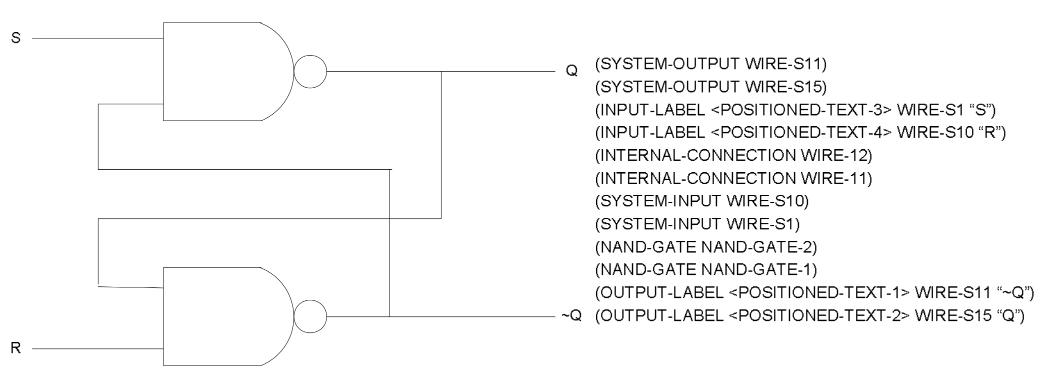


Figure 2: SR-Latch logic diagram and the representation produced by GeoRep.

varado, Oltmans, & Davis, 2002), where the networks were used to guide the interpretation of sketched visual elements for glyphs. Here we focus on the use of Bayesian networks in the context of bottom-up inferencing about glyphs given a set of partially-known qualitative visual relations.

In this system, the Bayesian networks are created dynamically as part of the rule firing process, which allows them to be created via a simple extension of the visual domain theory. The original GeoRep system's visual domain theories used a graphical rule language (Ferguson, 1994; Ferguson & Forbus, 1995). For example, Figure 4a shows the rule for recognizing a NAND-gate from this system. The new mechanism uses the same rule architecture to set up the Bayesian network structure (Figure 4b). This allows the system to use probabilistic reasoning while retaining the advantages of the symbolic visual domain theory.

The Mechanism for Recognition

We now look at how a rule within the visual domain theory creates integrated probabilistic nodes in GeoRep. For each glyph to be recognized, a rule such as the one in Figure 4b must be created. The rule's trigger captures the set of visual elements that make up the glyph. Alternatively, these visual elements can be seen as the set of proximate elements necessary to begin a glyph recognition attempt. In our implementation, the rule triggers include the existence of the visual elements as well as proximate relations between the elements.

In Figure 4, we can see how the rule in Figure 4b works on the glyph in Figure 4c. When this rule fires, it performs two actions. First, it creates a node in the LTMS dependency network representing the potential glyph interpretation for the visual element set. The node is created as an LTMS node with its existing implicational structure. This means that it can be retracted if any of the rule triggers becomes false (e.g., if a visual element is deleted from the diagram).

The rule's second action is to create a probabilistic node and embed it within the previously created LTMS node for the glyph interpretation. The probabilistic node captures the imprecision in a glyph using Bayesian networks. Al-

```
(grule ((attached ?L1 ?A1)
        (attached ?L3 ?A1)
        (parallel ?L1 ?L3)
        (polyline-member ?L1 ?polyline)
        (polyline-member ?L3 ?polyline)
        (polyline-member ?L2 ?polyline)
        (:different ?L1 ?L2)
        (corner ?L1 ?L2)
        (:different ?L3 ?L2)
        (corner ?L2 ?L3)
        (abuts ?C1 ?A1)
        (:gentemp ?gate-name "NAND-GATE-"))
     (:assume-with-implications
        (represents (composite ?polyline ?A1 ?C1)
                    (nand-gate ?gate-name))
         ((represents-specific-part ?L2
                          (input-plate ?gate-name))
          (represents-specific-part ?C1
                          (output-circle ?gate-name)))))
```

(a)

```
(grule ((line ?L1)
        (proximate ?L1 ?L2)
        (line ?L2)
        (proximate ?L2 ?L3)
        (line ?L3)
        (proximate ?A1 ?L2)
        (arc ?A1)
        (proximate ?arc ?circle)
        (circle ?C1))
(:eval (create-probabilistic-node
         (:tms-node <truth-maintenance-node-information>)
         (:network NAND-GATE)
         (:nodes (N1 :range (true false)
                  :parents nil
                  :elements (?L1 ?L2)
                  :visual-test (perpendicular ?L1 ?L2)
                  :p-table (t t t t t 1)
                           (t t t t f 0.9) ...) ...)
```

(b)

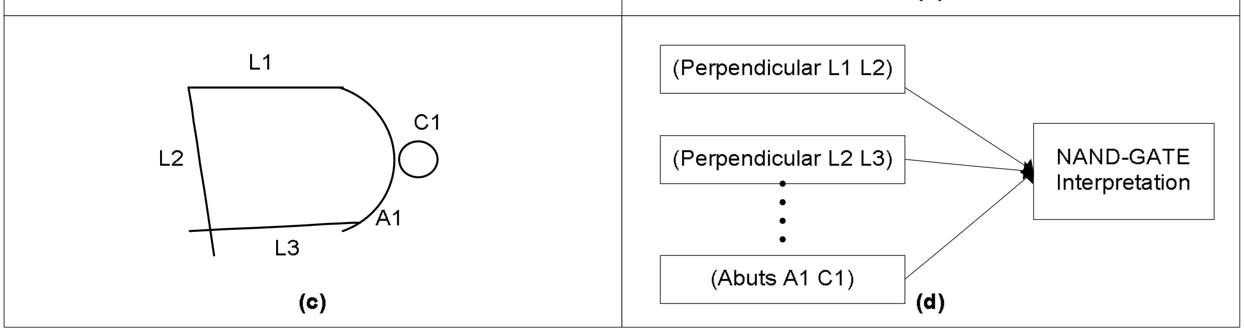


Figure 4: (b) is a rule in the current system, it demonstrates the information that must be encoded in order to create the Bayesian networks. (a) is a rule from the original GeoRep system to recognize a NAND-gate. (c) is a sample nand-gate that the new system can recognize, but that the old rules would fail on. (d) is a fragment of the Bayesian network created by the rule.

though Bayesian networks are typically used to represent the causal dependencies between events, here we use the networks to represent the contribution that sets of visual relations have on the final glyph interpretation. For instance, in the NAND-gate figure (Figure 4c) the existence of *perpendicular* relations between lines and an *abuts* relation between the circle and the arc contribute strongly to the probability that the glyph should be recognized as a NAND-gate. Similarly, if these relations do not exist, a NAND-gate is less likely.

The network created by the system is, as in Figure 4d, a single-layer Bayesian classifier. To determine the probability of the interpretation in the created network, the mechanism must first decide which nodes in the Bayesian network are evidence for the interpretation.

This is done in two ways. First, the relation that the Bayesian network node represents is searched for in GeoRep's low-level and high-level representations. If the relation is found, then that node is considered evidence. If the relation is not found, then the mechanism uses the visual test stored in the node to observe if the relation should be part of the representation.

For example, in the network in Figure 4d, if the mechanism could not find the relation `(Perpendicular L1 L2)`, it then runs the visual test `(Perpendicular L1 L2)` and applies the result as evidence in the Bayesian network. Additionally, the new relation is added to GeoRep's representations.

Communication between representation levels

The last important part of the mechanism is how the created probabilistic node interacts with the GeoRep reasoner (Figure 5). There are two main ways in which communication occurs. The first is by affecting the truth labeling for the glyph interpretation node in the LTMS. As described above, when a rule for a glyph fires, it creates an LTMS node for that interpretation. This node is initially labeled *Unknown*. The glyph recognition process in the probabilistic node generates a probability for the interpretation based on the existing evidence. When this probability reaches a threshold value, the labeling of the LTMS node is changed to *True*. Similarly, if the probability given the current evidence is below the threshold, the LTMS node will be labeled *False*.

The second method of communication is via the evidence collection process. Here, a probabilistic node accesses GeoRep's representations to collect evidence for the Bayesian network. It may also change the existing diagram representation if it determines that a visual test is required to determine if a relation is evidence for an interpretation.

This represents a more sophisticated method of top-down influences than in earlier versions of GeoRep. Previously, rules could make callbacks to the LLRD visual routines, but would only do so by satisfying a set of rule conditions. The visual test results were never added into the diagram representation. This new implementation allows the anticipated glyph structure represented in the probabilistic networks to guide further visual processing.

Results

Using this mechanism, we were able to improve the glyph

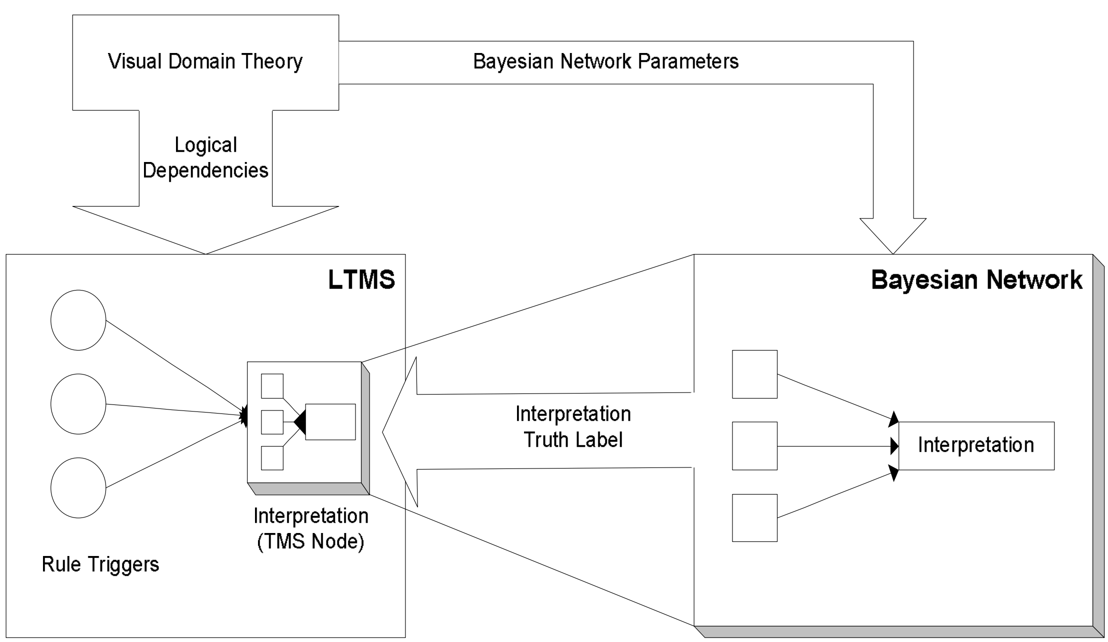


Figure 5: The interactions and communications between the truth maintenance system and the Bayesian networks.

recognition capabilities of the GeoRep reasoner. Figure 6 shows an example of the representation for the SR-latch diagram generated by the reasoner using the new probabilistic mechanism.

The original reasoner could not correctly interpret this diagram, because the imprecision in the drawn glyphs provided inadequate visual relations to fire visual domain theory rules. However, the new mechanism easily recognizes the NAND-gate glyphs and the rest of the circuit. The importance of recognizing the glyphs goes beyond simply the recognition alone. Recognizing the NAND-gates and adding them to GeoRep's high level representation allows further rules to be processed. For instance, the rule for system-input depends on the relationship between a segment and a recognized NAND-gate. The probabilistic mechanism recognizes the glyphs and then other rules from the visual-domain theory may be triggered.

The two systems working together in this integrated fashion, allows for better overall diagram understanding.

Related Work

There is a wide variety of work on diagram and sketch recognition systems (Davis, 2002; Alvarado, Oltmans, & Davis, 2002; Alvarado & Davis, 2001; Cohen et al., 1997; Ferguson & Forbus, 2002; Gross, 1996; Landay, 1995; Stahovich, 1998). These systems vary in the amount of glyph recognition they perform and the contextual knowledge they integrate into the glyph recognition process. They also vary in how well recognizers in one area can be used in another.

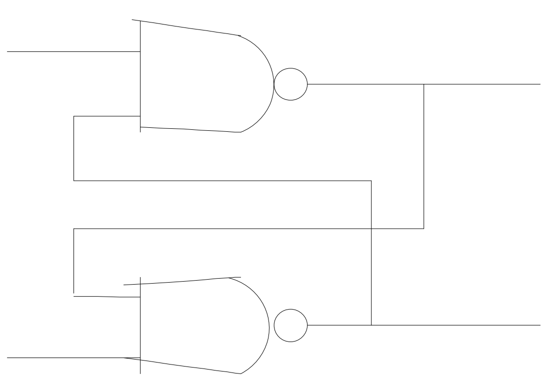
Several previous systems have shown the general effectiveness of probability-based methods of glyph recognition. These systems have used techniques such as partial template matches (Gross, 1996), and hidden-Markov models and neural networks (Cohen et al., 1997) to flexibly recognize glyphs given noisy data. In general, however, probabilistic recognizers built using these techniques can be difficult to generalize, often requiring extensive training for new glyphs.

The use of contextual information and additional modalities can also aid recognition (Oviatt, 1999). Extra information beyond what exists solely in the diagram is used to aid the process of sketch or diagram understanding. A strong demonstration of the power of this technique may be found in Quickset (Cohen, 1997), which combines information from a speech-understanding module with the glyph recognition system to improve recognition. Similarly, NuSketch Battlespace (Forbus, Ferguson, & Usher, 2000; Ferguson & Forbus, 2002) uses speech understanding to avoid intensive glyph recognition, while retaining the ability to make domain inferences based on spatial characteristics of the drawn glyphs. The use of a probabilistic framework in the system described here may allow for easier integration of other modalities.

There are also a number of systems that, like the one described here, use a two-level approach to the problem of glyph recognition, where a low-level domain-independent module recognizes primitive shapes and elements, and a high-level domain-dependent system uses the low-level representation to perform glyph recognition (Ferguson, 1994; Ferguson & Forbus, 2000). This approach can be made considerably less brittle by the incorporation of flexible recognizers at the low level, allowing for sketch input (Davis, 2002).

A significant limitation of the current system is that it does not yet work from sketch data. Although this system does not yet have the power and flexibility of systems that work from sketch data, it does have a number of potential advantages long-term. First, many existing systems do not work with static diagrams, but use timing information from the sketch as part of glyph recognition. In contrast, this system should be able to recognize diagrams independent of the process of creation. Second, unlike some of these systems, this technique focuses on ambiguity that exists at the level of the set of available visual relations, rather than at the level of visual elements.

There has also been work implementing probabilistic reasoning in truth maintenance systems. These systems are called belief maintenance systems and typically extend the TMS labeling from {*True*, *False*, *Unknown*} to labeling via intervals such as in probabilistic logic (Ramoni & Riva, 1994; Nilsson, 1986). The system presented here does not attempt to integrate probabilistic reasoning directly into the TMS framework by, but rather to allow two different knowledge representations to interleave, each relying on the other's output for a more robust diagram interpretation.



(NAND-GATE NAND-GATE-2)
(NAND-GATE NAND-GATE-1)

(SYSTEM-OUTPUT WIRE-S11)
(SYSTEM-OUTPUT WIRE-S15)
(INTERNAL-CONNECTION WIRE-12)
(INTERNAL-CONNECTION WIRE-11)
(SYSTEM-INPUT WIRE-S10)
(SYSTEM-INPUT WIRE-S1)

Figure 6: Results of running the new mechanism on an imprecise SR-latch diagram. The predicates in bold indicate the NAND-gate glyphs that were recognized via the new mechanism. The remaining predicates are a result of the mechanism detecting these glyphs. The earlier GeoRep system would have been unable to detect any of these.

Beyond the truth maintenance framework, there is important work on merging symbolic and probabilistic representations. Koller and Pfeffer's (1998) probabilistic frame-based systems demonstrate how representations of different types can be combined to produce a more robust final result, in this case, the ability to compensate for the lack of structural representation in Bayesian networks and the limited ability to represent uncertainty in frame-based systems.

Conclusions

Robust recognition of diagram glyphs will lead to better diagram understanding. We have presented a new mechanism that can recognize glyphs that vary widely in the level of precision in which they are drawn. By recognizing these glyphs the complete process of diagram understanding can continue. The mechanisms together provide this benefit.

Future work will look to incorporating this system into a more dynamic sketching environment. We also plan to look at some of the interesting properties of visual relations (such as jointly exhaustive and pairwise disjoint relation sets) to see if these properties can be better leveraged in probabilistic networks. We are also developing a more compact formalism for the conditional probability tables, and are investigating the reuse of probability tables.

References

Alvarado, C. and Davis R., (2001). Resolving ambiguities to create a natural sketch based interface. In *Proceedings of IJCAI-2001*.

Alvarado, C., Oltmans, M., and Davis, R. (2002). A framework for multi-domain sketch recognition. *AAAI Spring Symposium on Sketch Understanding,* pp. 1-8, March 25-27 2002.

Cohen, P., Johnston, M., McGee, D., Oviatt, S., Pittman, J., Smith, I., Chen, L., and Clow, J., (1997). QuickSet: Multimodal interaction for distributed applications. In *Proceedings of the Fifth Annual International Multimodal Conference*, Seattle, 31-40.

Cohn, A. G. (1997). Qualitative spatial representation and reasoning techniques. In Brewka, G., Habel, C., and Nebel, B. (Eds.), *Proceedings of KI-97* (pp. 1-30). Freiburg, Germany: Springer-Verlag.

Davis, R. (2002). *Position Statement and Overview: Sketch Recognition at MIT.* Paper presented at the AAAI Spring Symposium on Sketch Understanding. Palo Alto, CA.

Ferguson, R. W. (1994). MAGI: Analogy-based encoding using symmetry and regularity. In A. Ram & K. Eiselt (Eds.), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 283-288). Atlanta, GA: Lawrence Erlbaum Associates.

Ferguson, R.W., Bokor, J.L., Mappus, R.L., & Feldman, A. (2003). Maintaining spatial relations in an incremental diagrammatic reasoner. *Conference on Spatial Information Theory*, Ittingen, Switzerland, September 2003.

Ferguson, R. W., & Forbus, K. D. (1995). Understanding illustrations of physical laws by integrating differences in visual and textual representations. *Paper presented at the Fall Symposium on Computational Models for Integrating Language and Vision.* Cambridge, Massachusetts.

Ferguson, R. W., and Forbus, K.D. (2000). GeoRep: A flexible tool for spatial representation of line drawings, *Proceedings of the 18th National Conference on Artificial Intelligence* (pp. 510-516). Austin, Texas: AAAI Press.

Ferguson, R. W., and Forbus, K.D. (2002). A Cognitive approach to sketch understanding. *Proceedings of IAAI-2002*.

Ferguson, R. W., Rasch, R. A. J., Turmel, W., and Forbus, K. D. (2000). Qualitative spatial interpretation of Course-of-Action diagrams, In *Proceedings of the 14th International Workshop on Qualitative Reasoning*. Morelia, Mexico.

Forbus, K. D., and de Kleer, J. (1993). *Building Problem Solvers.* Cambridge, MA: The MIT Press.

Forbus, K. D., Ferguson, R. W., and Usher, J. M., (2000). Towards a computational model of sketching. In *Proceedings of the International Conference on Intelligent User Interfaces*, Sante Fe, NM.

Forbus, K. D. and Usher, J. (2002). Sketching for knowledge capture: A progress report. *IUI-2002*.

Gross, M. D. (1996). The Electronic Cocktail Napkin - computer support for working with diagrams. *Design Studies* 17(1): 53-70.

Glasgow, J., Narayanan, N. H., and Chandresekaran, B. (1995). *Diagrammatic Reasoning: Cognitive and Computational Perspectives.* Menlo Park, CA: The AAAI Press/The MIT Press.

Hoffman, D. D., and Richards, W. A. (1984). Parts of recognition. *Cognition, 18*(1-3), 65-96.

Koller, D. and Pfeffer, A. (1994). Probabilistic frame-based systems. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence* (pp. 580-587).

McAllester, D. (1990). Truth maintenance. *Proceedings of the Eighth National Conference on Artificial Intelligence* (pp. 1109-1116). Boston, MA: AAAI Press.

Nilsson, N. (1986). Probabilistic logic. *Artificial Intelligence*, 28:71-87.

Ramoni, M., and Riva, A. (1994). Belief maintenance in Bayesian networks. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence,* pp. 498--505.

Rock, I. (1973). *Orientation and Form.* New York, NY: Academic Press.

Stahovich, T. F., Davis, R., & Shrobe, H. (1998). Generating multiple new designs from a sketch. *Artificial Intelligence*, 104(1-2), 211-264.