GeoRep: A Flexible Tool for Spatial Representation of Line Drawings

Ronald W. Ferguson and Kenneth D. Forbus

Qualitative Reasoning Group
Department of Computer Science
Northwestern University
1890 Maple Avenue
Evanston, IL 60201 USA
{ferguson, forbus}@ils.nwu.edu

Abstract

A key problem in diagrammatic reasoning is understanding how people reason about qualitative relationships in diagrams. We claim that progress in diagrammatic reasoning is slowed by two problems: (1) researchers tend to start from scratch, creating new spatial reasoners for each new problem area, and (2) constraints from human visual processing are rarely considered. To address these problems, we created GeoRep, a spatial reasoning engine that generates qualitative spatial descriptions from line drawings. GeoRep has been successfully used in several research projects, including cognitive simulation studies of human vision. In this paper, we outline GeoRep's architecture, explain the domain-independent and domain-specific aspects of its processing, and motivate the representations it produces. We then survey how GeoRep has been used in three different projects–a model of symmetry, a model of understanding juxtaposition diagrams of physical situations, and a system for reasoning about military courses of action.

Introduction: How Diagrams Work

Diagrams are ubiquitous. In daily communications, through sketches, maps, and figures, people use diagrams to convey information. Some diagrams depict intrinsically spatial domains, such as bus routes or furniture arrangements. Other diagrams use spatial concepts to compactly show more abstract relations, such as corporate hierarchies or data flow in a computer program. In all such domains, diagrams can be extremely effective.

It is also true, however, that there is a keen difference between effective and ineffective diagrams. Small visual differences may distinguish a diagram that elucidates from one that confuses (Tufte, 1990). A key difference between good and bad diagrams is how well they utilize the kinds of qualitative spatial relations most easily perceived by the human visual system. In the best diagrams, these spatial relations support the conceptual relations the reader is meant to infer. For example, in a thermodynamics diagram, an arrow may indicate the direction of heat flow, with thicker arrows to indicate greater flow, or tapering arrows to indicate heat dissipation. Or, in a circuit diagram, wires may be drawn so that related wires are adjacent and parallel, so they can be visually grouped.

For this reason, to understand how diagrams work, we must show how diagrams use visual characteristics to support particular qualitative inferences. In the system described here, we model this process as an interaction between two representation levels:

1. A low-level, domain-independent representation which involves a representative set of primitive spatial relations. This level models human low-level vision.
2. A high-level, domain-specific representation that models visual skills for a particular domain. This level links low-level visual relations to a domain's conceptual content.

These two representation levels form the basis of GeoRep. GeoRep is an engine for building diagrammatic reasoners. GeoRep takes as input a line drawing, given as a set of primitive visual elements. From this drawing, GeoRep creates a predicate calculus representation of the drawing's visual relations. To perform this task, GeoRep, given the drawing, examines the primitive shapes in the figure, looking for a broad set of low-level visual relations. These relations are detected by a library of visual operations (assumed to be domain-independent) which partially cover the set of *universal visual routines* (Ullman, 1984). Next, GeoRep uses these relations, in combination with domain-dependent rules, to generate the second, domain-specific representation. GeoRep's two-level

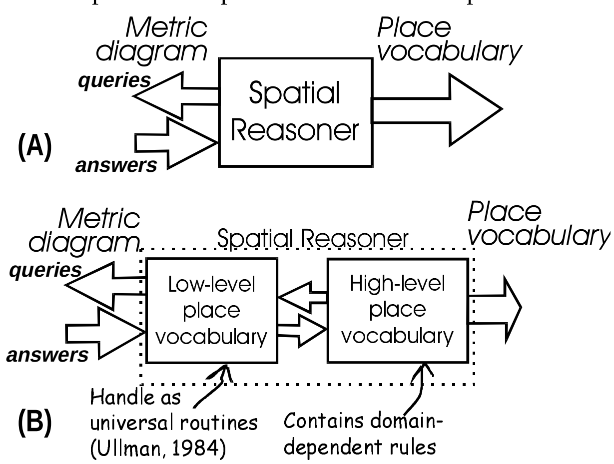


Figure 1: The Metric Diagram / Place Vocabulary framework (A) and how it is modified for GeoRep (B).

architecture provides a sophisticated model of how early visual relations are used for inferring conceptual relations.

At the level of the high-level representation, GeoRep is a qualitative spatial reasoner. GeoRep's qualitative spatial reasoning uses a variant of the MD/PV framework (Forbus, 1980; Forbus, Nielsen, & Faltings, 1991). This framework is motivated by the *poverty conjecture*, which states that "there is no purely qualitative, general-purpose representation of spatial properties." (Forbus et al., 1991). For this reason, MD/PV reasoners use two representations levels: a *metric diagram*, which contains quantitative information (and often, some symbolic or qualitative representation), and the *place vocabulary*, which is a qualitative spatial representation fitted to the particular place and task (Figure 1-A). The place vocabulary is constructed as needed by querying the metric diagram.

Qualitative spatial reasoners using the MD/PV framework have been successful in many domains, including the analysis of mechanical systems (Forbus et al., 1991; Kim, 1993) and graphs (Pisan, 1995).

GeoRep elaborates on the MD/PV model by splitting the place vocabulary into higher and lower levels. The low-level place vocabulary represents low-level visual relations specific to early human vision, while the high-level place vocabulary is a task-specific spatial representation derivable from the low-level vocabulary (Figure 1-B). By generating the initial low-level place vocabulary directly from the metric diagram (which in this case is the line drawing itself), GeoRep can then use this initial vocabulary as the building blocks for a broad (if still finite) class of high-level place vocabularies. Thus GeoRep, while not a general diagrammatic reasoner, is general over the set of high-level place vocabularies derivable from this low-level place vocabulary, which is in turn bounded by cognitive constraints in human perception. Our conjecture is that this cognitively-grounded place vocabulary is both computationally useful and psychologically plausible.

GeoRep's consistency with human visual abilities–unlike previous systems using the MD/PV framework–provides robustness. Although some visual skills are domain-specific, the fact that people use visual reasoning in such a broad variety of tasks suggests that a sufficiently robust visual processing engine could provide similarly general services for diagrammatic reasoning. This generality addresses a key limitation of current research in diagrammatic reasoning, which is the tendency for every researcher to start from scratch, implementing a spatial reasoning system aimed at one class of problems.

Although most previous diagrammatic reasoning systems are motivated by human visual abilities (c.f. the systems described in Glasgow et.al. (1995)), their design has typically been driven more by the task than by the psychology of human vision. While this has lead to useful insights, we believe that an explicit concern with human vision can lead to better diagrammatic reasoners.

The next section describes GeoRep's architecture, explaining where its processing and representational choices have been influenced by perceptual psychology findings. We then demonstrate GeoRep's generality by showing its use in three systems: MAGI (Ferguson, 1994), a model of symmetry; JUXTA (Ferguson & Forbus, 1995), a model of juxtaposition diagram understanding; and COADD, a system for understanding military course of action diagrams. We close with a discussion of GeoRep's limitations and future development.

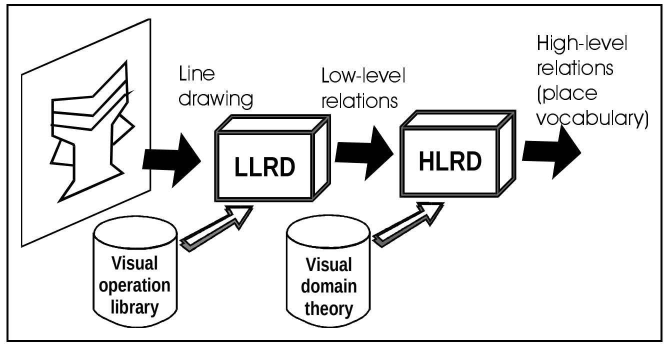


Figure 2: A simplified schematic of GeoRep's architecture.

Architecture

GeoRep's architecture is shown in Figure 2. GeoRep's input is a line drawing, given as a vector graphics file. This file uses the FIG graphics format. Using drawings rather than bitmaps avoids the problem of doing line detection, which is essential in machine vision, but not critical in diagrammatic reasoning. Using line drawings also makes diagram input simple: diagrams can be built using an off-the-shelf drawing program (Hendrich, 1999). Line drawings have been successful in several systems (e.g., Evans, 1968; Gross, 1996; Sutherland, 1963) and work well with existing spatial reasoning models.

The output of GeoRep is a description of the figure expressed in a domain-dependent high-level place vocabulary. Like previous approaches to spatial representation, the representation produced by GeoRep emphasizes compact, composable vocabularies that directly reflect visual structure. Entities are mapped onto geometric elements or object parts, with predicates to represent connections and arrangements (Biederman, 1987; Palmer, 1975). The composability of these arrangements are reflected in the composability of the vocabulary itself.

GeoRep's internal processing contains two stages: the low-level relational describer (LLRD) and the high-level relational describer (HLRD). The LLRD handles the domain-independent representation of the line drawing. It detects and represents a large set of useful visual relations. These relations are structural relations detected early in visual perception.

The HLRD, in turn, uses domain-specific rules that extend the LLRD's representation. These extensions include new visual relations (and how to compute them) and ways to recognize depicted items. The final output of the HLRD is one or more *representation levels*. A representation level is a set of propositions that corresponds to some specific task or type of analysis. For example, representation levels may include the LLRD's

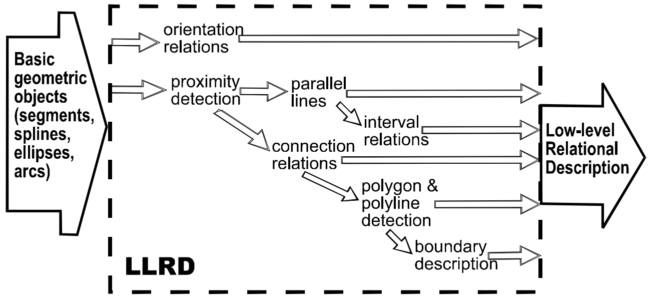


Figure 3: Processing of visual elements within the LLRD

basic visual representation, more complex visual relations, a representation of the depicted items, or potentially even limited reasoning within the diagram's problem domain (see Figure 5 for an example).

The Low-Level Relational Describer (LLRD)

GeoRep's first stage, the LLRD, creates GeoRep's low-level, domain-independent spatial representation. Starting with primitive visual elements, the LLRD detects and represents a broad set of early visual relations, using a library of *visual operations*. These operations correspond to Ullman's (1984) *universal routines*, which are routines that run in the absence of visual goals. These operations, while they do not attempt to model visual processes, model the visual relations such processes detect. Conceptually, these operations are applied in parallel over the visual field's proximate shapes. Due to dependencies between the LLRD's visual relations, it pipelines processing so that more complex visual relations are computed after simpler ones (Figure 3). For example, parallel lines and polygons, once detected, are fed to the interval relation and boundary description systems, respectively.

The LLRD recognizes five primitive shape types: line segments, circular arcs, circles and ellipses, splines (open and closed), and positioned text. The LLRD also subsumes some visual elements into polylines, polygons, and groups.

Computing proximity. Proximity is the LLRD's core attentional mechanism. Because it is impractical for the LLRD to detect all visual relations between all available element combinations, it only checks for relations between proximate elements. For example, in representing a human stick figure, the LLRD might relate the "foot" to the "leg", and the "leg" to the "torso", but wouldn't attempt to find relations between the "foot" and "hand".

To determine element proximity, GeoRep uses a calculation that is a function of element size, distance and shape type. Each visual element type has a prototypical *area of influence* based on the element's size. For example, a circle's area of influence is the area within twice its radius. Similarly, for a line segment the area of influence extends out from the segment for the segment length. Elements are considered proximate when their areas of influence overlap. Areas of influence are calculated as sets of circles and rectangles to make overlap checking efficient, and all proximity calculations are cached. Note that because all pairs of items must be considered, the time complexity of this stage is O(N^2) over the number of visual elements. This makes proximity detection the LLRD's most expensive stage. However, proximity detection also makes subsequent LLRD operators more efficient by limiting their application to either the set of visual elements or the set of proximate element pairs.

Though imperfect, this proximity measure has many advantages. It is easily constructed, relatively efficient, and captures the intuition that large elements (such as a large rectangles or long polylines) relate visually to many other elements. Similar approaches to rating nearness have been used successfully (e.g., Abella and Kender, 1993).

Running the visual operations. Once the LLRD determines which elements are proximate, it looks for other visual relations between proximate elements, using a visual operation library. Each visual operation detects specific visual relations that are part of early vision. All visual operations act on some combination of primitive visual elements, composite visual elements, and reference frames.

The rest of this section briefly surveys the set of visual operations the LLRD uses.

Orientation and the frame of reference. One fundamental characteristic of vision is the reference frame. Experiments have shown that figure orientation can have a critical effect on perception, including figure recognition (Rock, 1973). The LLRD detects many orientation-based relations, including *horizontally* and *vertically*-oriented elements, and *above* and *beside* relations between elements. The LLRD also looks for elements occupying the same horizontal or vertical extent.

Like humans, GeoRep can change its reference frame. GeoRep's default reference frame is gravitational, but can be changed based on clues in the scene, such as a preponderance of lines at one orientation, figural elongation, or symmetry. When the reference frame changes, LLRD relations using the old reference frame are retracted, and new relations asserted in GeoRep's knowledge base.

Parallel lines. The LLRD also detects parallel line segments, modeling the ease with which humans detect parallelisms. However, in practice, describing the parallel segments alone often doesn't adequately constrain the description of a drawing. To elaborate on parallel elements in a cognitively-plausible fashion, we extended the LLRD to categorize parallel segments using Allen's (1983) interval relations. Allen's interval relations were useful in describing parallel segments because they constrain the relative position of segment endpoints in a way invariant to the frame of reference. Admittedly, while in practice these interval relations have been extremely useful, empirical evidence for their role in vision is still marginal.

Connection relations. The LLRD also describes element connectivity, which is a central factor in perceptual organization. Connectivity is detected using standard computational geometry routines (Glassner, 1990) amended with strictness factors. The element types determine the type of connection relation. "Ended" elements, such as line segments and arcs, can connect or cross other elements. Specifically, the LLRD checks pairs of segments for corner, intersection, and mid-connections. Arcs may connect with segments as well, and their connections may be aligned or misaligned. Connections may also have a particular character. Corner angles, for example, are characterized as acute, obtuse, or perpendicular. The LLRD also detects and classifies connections between line segments and curved objects, such as circles and ellipses. Endpoint connections between a segment and an ellipse or circle are checked to see if the connection is radial or tangential. Other curved shapes, such as circles, ellipses, and arcs, are connected by abutment (i.e., when boundaries touch).

Building Composite Elements

A key insight of Gestalt theory is that the whole is seen differently than the sum of its parts. For GeoRep, which operates mainly in a bottom-up fashion, this means that it must recognize when individual elements can be subsumed into larger structures. When elements are collected into composite elements, we say they are *visually subsumed*, and the status of those elements changes. Once subsumed, a visual element is represented as part of its composite structure rather than individually.

GeoRep contains three mechanisms to perform visual subsumption. Elements may be subsumed as polylines and polygons, via grouping, and by constructing ad hoc composite elements, called *glyphs*. While these mechanisms lack the flexibility of human perception, they can simulate aspects of it, and can be extended when a particular visual domain requires it. These three mechanisms, although listed here with the LLRD, actually bridge lower and higher-level visual processing. The need to bridge these levels is due to the way subsumption is tied to perceptual organization. Perceptual organization itself often depends on either global element configuration or domain knowledge, which limits the effectiveness of bottom-up processing. Note that because GeoRep's visual rules can check if an element is a subsumed element, such rules may act only on unsubsumed elements, increasing reasoner efficiency.

Polylines and polygons. It has long been recognized that polylines and closed shapes are important in perception. The LLRD detects polylines and polygons using simple path-following algorithms. Despite the computational complexity of calculating closed shapes (c.f., Ullman, 1984), humans detect shape closure early in perception—perhaps pre-attentively (Treisman & Patterson, 1984).

Polygons, and their constituent corners and segments, have many characteristics derived from their boundaries. Their corners may be concave or convex, and groups of adjoining convex or concave corners constitute protrusions or indentations. Representing these characteristics is crucial to modeling human performance: inflexion points (indentations and protrusions) are critical in recognition tasks (Hoffman & Richards, 1984; Lowe, 1987), and recent studies have shown the importance of concavities in visual tasks such as symmetry judgment (Baylis & Driver, 1994; Ferguson, Aminoff, & Gentner, submitted).

The LLRD represents indentations and protrusions as groups of concave or convex points. Protrusions are also represented relative to the current reference frame, indicating the protrusions' relative vertical placement.

Grouping. Grouping requires some measure of similarity between grouped elements. This required similarity metric makes grouping too broad an effect to model with the LLRD. However, the LLRD can model limited grouping effects by using domain-specific similarity metrics in the HLRD.

Grouping in GeoRep thus depends on a set of domain-specific *grouping rules*. These rules determine which element pairs are similar enough to be grouped. For example, triangle groups may be collected with a grouping rule that pairs triangles of similar size. While there are limits to this approach—a new rule is needed for each new group type, and the rules are not generative—this mechanism has proven adequate for our current visual domains, and easily accommodates the construction of new grouping-based place vocabularies. We are currently looking at a tractable grouping sub-case using factors, such as similar size, orientation, and shape, that have been shown to allow items to be grouped pre-attentively (Julesz & Bergen, 1983; Treisman & Gelade, 1980).

Glyphs. Along with other basic shapes, GeoRep includes *glyphs*, which are arbitrary collections of visual elements that constitute a symbol or other divisible visual form. Glyphs implement visual symbols, such as depictions of NAND gates or military units. Glyphs are treated as a single element with location and extent alone.

The High-Level Relational Describer (HLRD)

GeoRep's reasoning does not end with the LLRD's low-level place vocabulary. Built upon this low-level vocabulary is a high-level vocabulary specific to a visual reasoning task. For example, depicting connectivity in a wiring diagram or the meshing of gears may involve spatial relations that are not domain-general, but are still better expressed in a diagram than through text.

This high-level place vocabulary is created by GeoRep's second stage: the HLRD. The HLRD's input is the LLRD's description. The HLRD contains a rule engine utilizing a logic-based truth-maintenance system (LTMS; Forbus & de Kleer, 1993). The complexity of this stage thus depends on the domain. The rule engine loads rules from a visual domain theory, and creates a description using those rules.

HLRD rules are similar to those for other rule-based systems, but are set apart by the rules' visual vocabulary, which form a convenient abstraction layer for discussing domain-dependent visual symbols (e.g., the symbology of maps) and spatial relations. HLRD rules contain special forms for delimiting the application of rules to proximate objects and for calling the LLRD's visual operation library.

While HLRD rules are domain-specific, there are some rules used across domains. For example, one rule set handles *representational links* between visual elements and what they represent. In thermodynamics, for instance, a trapezoid may represent a fluid container. While the specific mappings from geometry to conceptual entity are domain-specific (trapezoids may have different meanings in other domains), the properties of representational links are more general. These rules dictate that each visual element represent only one thing (excluding partonomic relations). Multiple element interpretations are then resolved via various heuristics (e.g., when conflicting interpretations exist, choose the interpretation that accounts for the most visual elements and retract the other interpretation in the LTMS).

Because the HLRD uses the LTMS, the HLRD can explain *why* it believes that particular visual elements represent particular things: e.g., why a polygon represents a coffee cup. Another advantage of explicit representational links is that they can be used to extend the place vocabulary. For example, given a drawing of two coffee cups, GeoRep can determine which cup contains more liquid by returning to the polygons representing the cups and comparing them to see if one cup is taller or wider.

Once the HLRD has generated a high-level description, it can either be retrieved from the HLRD directly, or filtered by relation type to simulate different diagrammatic representation levels. For example, one representation level might list only individual glyph properties, while another level might relate patterns of glyphs.

HLRD's ability to handle arbitrary place vocabularies is limited by the LLRD's capabilities. However, the advantage is that when HLRD rules use only the LLRD's representation or visual operations, it is cognitively plausible that the resulting description will contain relations that are visible to people. The LLRD's representation is valuable because it provides an easy-to-use and extensible vocabulary. But it is also valuable because, used correctly, it should tell us not just the relations a drawing depicts, but why a person would notice those relations.

Applications of GeoRep

To date, GeoRep has been used in three different projects: symmetry detection of abstract figures, diagrams of simple physical phenomena, and military Course-of-Action (COA) diagrams. We briefly survey each of them here, and provide references for those who wish to explore, for each system, GeoRep's role in greater depth.

Symmetry detection. GeoRep is used in the MAGI symmetry-detection model (Ferguson, 1994; Ferguson, in preparation). MAGI, which maps similar relations in a representation to determine its symmetry, uses GeoRep to detect symmetry in drawings, including functional drawings such as logic circuits. It has also been used to simulate experimental results. In (Ferguson, Aminoff, & Gentner, 1996; Ferguson et al., submitted), subjects in two experiments judged the symmetry of randomly-generated polygons after brief presentation times (50 ms). The experiments found that qualitative visual structure, such as

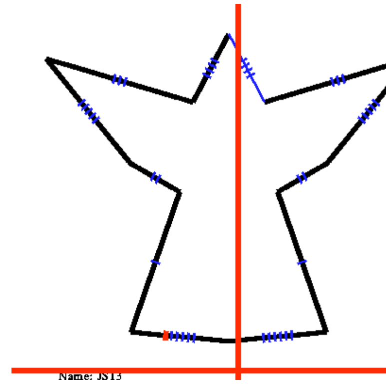


Figure 4: Sample figure from asymmetry study, with axis and correspondences are drawn in by MAGI.

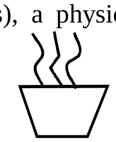
boundary concavities, had a significant effect on whether a figure was judged symmetric.

To simulate the experimental results, GeoRep was given the polygon set, using the same segment data used for the experimental stimuli. For each figure, GeoRep generated a low-level relational description. This was then passed to the MAGI model, which determined the qualitative symmetry of the figure (Figure 4). The simulation was successful, resulting in the same general pattern of symmetry judgments found in the human subjects. MAGI, like human subjects, detected asymmetries more easily when the asymmetry involved differences in qualitative visual structure, such as mismatches in vertex concavity or in the number of vertices.

Juxtaposition-based diagrams of simple physical phenomena. GeoRep is used as part of a system called JUXTA (Ferguson & Forbus, 1995; Ferguson & Forbus, 1998), which critiques simplified diagrams of physical phenomena.

For each diagram, GeoRep generates three different levels of description: a visual level (using the LLRD, and some additional rules), a physical level (interpreting the

Original Drawing



Visual Level

(POLYGON poly1)
(NUMBER-OF-SIDES poly1 4)
(UPRIGHT-TRAPEZOID poly1)
(SPLINE spline1)
(SPLINE spline2)
(SPLINE spline3)
(SPINE-GROUP group1
(GROUP spline1 spline2 spline3))

Physical Level

(CONTAINER cup1)
(LIQUID liquid1)
(CONTAINS (CONTAINER cup1)
(LIQUID liquid1))
(STEAM-HEAT steam1)
(RISING-FROM steam1 liquid1)

Process Level

(FLOW HEAT liquid1 atmosphere steam1)

Figure 5: A subset of the representations produced by GeoRep for JUXTA, with the original figure.

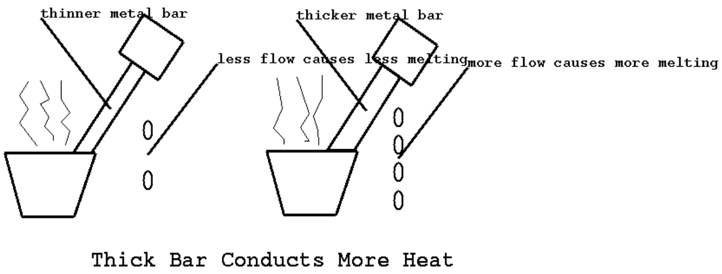


Figure 7: JUXTA's labeling of the aligned differences detected in a diagram, as related to the caption

visual elements as domain objects using a set of structural templates), and a process level (giving the physical processes inferred from the diagrams). A representative sample of each level is given in Figure 5.

Using MAGI to detect the repeated parts of the scene, JUXTA detects the physical and process differences between those parts, and attempts to relate those differences to the caption. The resulting system can critique the diagram based on how the diagram meets the expectations set in the caption. Based on the caption, for example, JUXTA can label the figure's critical differences (Figure 7).

To perform this analysis, the distinction between levels of interpretation is crucial. Visual differences can be relevant or irrelevant depending on the caption's interpretation. Because GeoRep can represent multiple abstraction levels, JUXTA can distinguish between visual differences that could confuse the reader and differences that, while noticeable, would not be confusing.

Course-of-Action Diagrams. In DARPA's High-Performance Knowledge Bases (HPKB) initiative, GeoRep is being used for spatial reasoning about Course-of-Action (COA) diagrams (Ferguson, Rasch, Turmel, & Forbus, 2000). These diagrams, drawn by the military for tasks such as troop movement planning, use a well-defined set of line-drawn symbols to indicate important areas, unit locations and types, tasks, movement paths, and obstacles. Most work performed with COA diagrams is done by hand, using grease pencils on clear acetate. Diagrams are frequently redrawn to remove irrelevant details or change

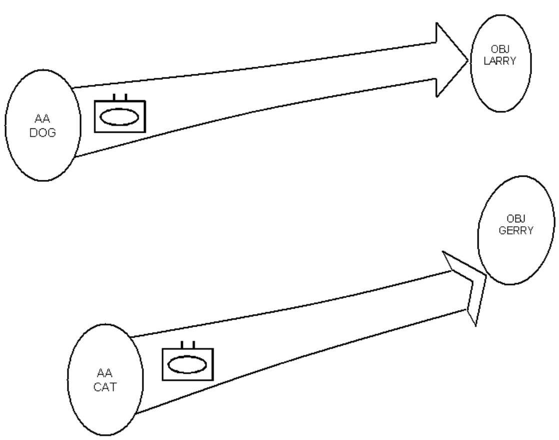


Figure 6: Example from a Course-of-Action diagram

the level of description.

The COA diagram describer (COADD), built using GeoRep, takes a line drawing of a COA diagram (as in Figure 6), and produces a description of the units, areas, and tasks given in the figure. Recognition of symbols in the COA diagram is handled by an HRLD rule set. It is worth noting that the initial prototype, which handles enough of the COA symbols to do simple but recognizable COA diagrams, was completed in less than 10 person-days, and involved only minimal changes to the LLRD (mostly to improve recognition of arrows). COADD's diagrams are the largest handled by GeoRep, containing as many as 197 visual elements.

Because most COA diagrams are constructed interactively, we are investigating extending GeoRep to handle interactive freehand sketches as input, instead of line drawings. A key technique is the use of glyphs in GeoRep to limit the low-level processing. A completed COA geographical reasoner utilizing GeoRep was recently used in a COA critiquer in HPKB (Ferguson et al., 2000)

Limitations and Areas for Future Work

GeoRep has evolved considerably as various projects have made demands on it. While GeoRep has shown itself to be a flexible and useful tool in our own research, it has significant limitations. These limitations must be addressed to make the model truly general.

First, GeoRep needs a cognitively accurate model of proximity. While GeoRep's proximity metric is sophisticated enough to incorporate the relative shape and size of elements considered proximate, the human attentional mechanism is much more complex, often balancing one proximity against another. For example, shapes A and B might be seen as proximate only if there is not some shape C that lies between them. We are investigating techniques for incorporating this model of proximity into the LLRD.

GeoRep's processing is mainly bottom-up, with only limited top-down influences on shape perception. Top-down influences occur when the HLRD calls LLRD operations to verify visual relations that are not checked by default. By using this limited top-down mechanism, GeoRep enforces the use of LLRD relations. In other words, GeoRep enforces the cognitive constraint that inferences be sanctioned by easily-perceived qualitative visual relations. We are currently examining ways to extend top-down influences while maintaining these vision-driven cognitive constraints.

GeoRep's intended use as part of an interactive sketching system highlights two other areas for improvement. GeoRep currently processes drawings in batch mode. For sketching, drawings will be processed incrementally. GeoRep currently expects each visual element to be accurately classified when read in. Although the strictness of LLRD operations can be varied, GeoRep does not have mechanisms to resolve ambiguous figures. Nor does it handle multiple variant feature interpretations of a single figure. For sketching, where a single pen stroke might be a spline, line segment, or arc depending on the context, GeoRep will have to be more flexible about choosing between alternate interpretations. These modifications will also allow GeoRep to be used with less-reliable data formats, such as vector data derived from scanned bitmaps of pre-existing diagrams.

Acknowledgements

This research was supported by the Cognitive Science and Computer Science programs of the Office of Naval Research, by the Defense Advanced Research Projects Agency, under the High Performance Knowledge Bases program, and by the National Science Foundation, under the Learning and Intelligent Systems program. Useful feedback and/or assistance was provided by Laura Allender, Jim Donlan, John Everett, George Lee, Yusuf Pisan, Rob Rasch, Bill Turmel, Jeff Usher and three anonymous reviewers.

References

Abella, A., & Kender, J. R. (1993). Qualitatively describing objects using spatial prepositions, *Proc. Eleventh National Conference on Artificial Intelligence.*

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Comm. of the ACM*, 26, 832-843.

Baylis, G. C., & Driver, J. (1994). Parallel computation of symmetry but not repetition within single visual shapes. *Visual Cognition*, 1, 377-400.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94(2), 115-147.

Evans, T. G. (1968). A program for the solution of a class of geometric-analogy intelligence-test questions. In M. Minsky (Ed.), *Semantic Information Processing* (pp. 271-353). Cambridge, MA: MIT Press.

Ferguson, R. W. (1994). MAGI: Analogy-based encoding using symmetry and regularity. *Proc. 16th Ann. Conference of the Cognitive Science Society* (pp. 283-8). Atlanta.

Ferguson, R. W. (in preparation). A structure-mapping model of symmetry detection.

Ferguson, R. W., Aminoff, A., & Gentner, D. (1996). Modeling qualitative differences in symmetry judgments, *Proc. 18th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum Associates.

Ferguson, R. W., Aminoff, A., & Gentner, D. (submitted). Early detection of qualitative symmetry.

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

Ferguson, R. W., & Forbus, K. D. (1998). Telling juxtapositions: Using repetition and alignable difference in diagram understanding. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in Analogy Research* (pp. 109-117). Sofia: New Bulgarian University.

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

Forbus, K. D. (1980). Spatial and qualitative aspects of reasoning about motion, *Proc. First National Conference on Artificial Intelligence*. Palo Alto, California.

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

Forbus, K. D., Nielsen, P., & Faltings, B. (1991). Qualitative spatial reasoning: The CLOCK project. *Artificial Intelligence*, 51(1-3).

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

Glassner, A. S. (Ed.). (1990). *Graphics Gems*. Chestnut Hill, MA: AP Professional.

Gross, M. D. (1996). The Electronic Cocktail Napkin: A computational environment for working with design diagrams. *Design Studies*, 17(1), 53-69.

Hendrich, N. (1999). JavaFIG: The Java diagram editor, [Web page]. Computer Science Department, University of Hamburg, Germany. Available: http://tech1.informatik.uni-hamburg.de/applets/javafig/ [1999, December 28].

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

Julesz, B., & Bergen, J. R. (1983). Textons, the fundamental elements in preattentive vision and perception of textures. *The Bell System Technical Journal*, 1619-1645.

Kim, H. (1993). Qualitative reasoning about fluids and mechanics (Technical Report #47). Evanston, IL: Institute for the Learning Sciences, Northwestern University.

Lowe, D. G. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, 31, 355-395.

Palmer, S. E. (1975). Visual perception and world knowledge. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in Cognition* (pp. 279-307). San Francisco: W. H. Freeman and Company.

Pisan, Y. (1995). A visual routines based model of graph understanding, *Proc. 17th Annual Conference of the Cognitive Science Society*. Pittsburgh: Erlbaum.

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

Sutherland, I. E. (1963). Sketchpad, a Man - Machine Graphical Communication System. Unpublished PH.D., Massachusetts Institute of Technology, Cambridge, MA.

Treisman, A., & Patterson, R. (1984). Emergent features, attention, and object perception. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 12-31.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.

Tufte, E. R. (1990). *Envisioning Information*. Cheshire, Connecticut: Graphics Press.

Ullman, S. (1984). Visual routines. *Visual Cognition*. 18(1-3), 97-159.