GeoRep: A Flexible Tool for Spatial Representation of Line Drawings

Ronald W. Ferguson
ferguson@ils.nwu.edu

Kenneth D. Forbus
forbus@ils.nwu.edu

Qualitative Reasoning Group
Department of Computer Science
Northwestern University
1890 Maple Ave
Evanston, IL 60201, USA

Abstract

A central problem in qualitative reasoning is understanding how people reason about space and shape with diagrams. We claim that progress in diagrammatic reasoning is being slowed by two problems: (1) Researchers tend to start from scratch, creating new spatial reasoners for each new problem area that they tackle, and (2) constraints from human visual processing are rarely considered. This paper describes 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, which suggests that our approach can overcome these problems. This paper outlines the architecture of GeoRep, explaining the domain-independent and domain-specific aspects of its processing and the motivations for the representations it produces. 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 – is described. Several limitations of GeoRep, as well as our plans for extending it, are discussed.

Introduction: How diagrams work

Diagrams are a ubiquitous part of everyday life. In everyday communications, through sketches, maps, and figures, humans use diagrams to convey information about a variety of domains. 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 the data exchanged between modules of a computer program, or the hierarchy within a corporation. In all such domains, both concrete and abstract, diagrams can be an extremely effective way of communicating information, even in comparison to the extraordinary power and flexibility of human language.

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 (Kosslyn, 1994; Tufte, 1990). One key difference between good and bad diagrams is how well a diagram utilizes the kinds of qualitative spatial relations that are easily perceived by the human visual system. These spatial relations, in the best diagrams, link cleanly with the conceptual relations the reader is meant to infer. For example, in a diagram of simple physical phenomena, 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, the wires may be drawn so that related wires are adjacent and parallel, so they can be visually grouped.

For this reason, if we want to understand what makes diagrams effective, we must understand how these perceived spatial relations can accommodate such reasoning tasks. We can model this process by decomposing it into two parts:

1. A low-level, domain-independent process that automatically detects a representative set of primitive spatial relations that people find visually salient. In terms of human vision, this process is intended to model low-level vision.
2. A medium-level, partially domain-specific process that embodies the particular visual skills needed for spatial reasoning in a domain, e.g., linking perceptual relationships to the domain's conceptual content.

This simple idea is the basis of GeoRep. GeoRep is an engine, a component for building diagrammatic reasoners. GeoRep takes as input a line drawing, given as a set of primitive visual elements. It produces as output a predicate calculus representation of the visual relations found in the drawing. Given the drawing, GeoRep first pulls out the primitive shapes in the figure and looks for a broad set of low-level visual relations. These relations are detected by a library of visual operations, and assumed to be domain-independent, covering a portion of the routines Ullman (1984) terms *universal visual routines*. Next, GeoRep uses these relations, in combination with domain-dependent rules, to generate higher-level relations specific to that diagram's domain. This two-level architecture provides GeoRep with a sophisticated understanding of how early visual relations are used for inferring conceptual relations.

From a qualitative physics perspective, GeoRep is a variant of the *metric diagram/place vocabulary* (MD/PV) model

(Forbus, 1980; Forbus, Nielsen, & Faltings, 1991). The MD/PV model 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, qualitative spatial reasoners must use two representation 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 linked to a particular place and task. The place vocabulary is constructed as needed by querying the metric diagram.

In the MD/PV framework, GeoRep's extensibility (and thus, its utility as a diagrammatic reasoning engine) is based on creating an intermediate place vocabulary which can be used to construct more complex place vocabularies. By generating an initial vocabulary of low-level visual relations from the metric diagram (in this case, the line drawing itself), then (assuming that these relations are consistent with human visual ability) it may be possible to use this initial vocabulary as the building blocks for a broad class of place vocabularies. Our conjecture is that this visually-grounded place vocabulary is both computationally useful and psychologically plausible.

In this way, unlike previous systems that used the MD/PV model, GeoRep has been designed to be a general-purpose system, with careful attention to consistency with human visual abilities. 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 less by the psychology of human vision, and more by the domain task. While a focus on the domain task can lead to useful insights, we believe an explicit concern with human vision can lead to better diagrammatic reasoners. This approach may also reduce the tendency for each researcher to start from scratch, implementing a spatial reasoning system aimed at a particular class of problems. While visual skills clearly contain domain-specific components, the fact that people use visual reasoning in a broad variety of tasks suggests that a sufficiently robust visual processing engine could provide similarly general services for diagrammatic reasoning.

We start by describing the architecture of GeoRep and how it works, pointing out where our processing and representational choices have been influenced by the findings of perceptual psychology. Next, we demonstrate GeoRep's generality by showing how it has been used in three systems: MAGI (Ferguson, 1994), a model of symmetry; JUXTA (Ferguson & Forbus, 1995), a model of juxtaposition diagram understanding, and a system for understanding military Course of Action diagrams. We close with a discussion of GeoRep's limitations and plans for future work.

Architecture

A simplified schematic of GeoRep's architecture is given in Figure 1. GeoRep's input is a line drawing, given as a vector graphics file. By using drawings rather than bitmaps, we avoid the problem of performing line detection, a process that is critical in machine vision, but adds little to a model of diagrammatic reasoning. Using line drawings also makes diagram input simple: diagrams can be built using an off-the-shelf drawing program. Line drawings have been successful in a number of systems (e.g., Evans, 1968; Gross, 1996; Pisan, 1995; Sutherland, 1963) and work well with many existing spatial reasoning models.

The output of the system is a spatial representation, expressed in propositional statements. As with many previous approaches to spatial representation, the representation produced by GeoRep emphasizes compact, composable vocabularies that directly reflect visual structure. Entities tend to be mapped directly onto geometric elements or object parts, with predicates to represent connections and arrangements (for similar approaches, see (Biederman, 1987; Palmer, 1975, 1989)). The composability of these connections and arrangements are reflected in the composability of the vocabulary itself.

GeoRep's architecture consists of 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 autonomously detects and represents a large set of useful visual relations. These relations are assumed to be 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 basic visual representation, more complex visual relations, a representation of the depicted items, or potentially even some reasoning within the diagram's problem domain.

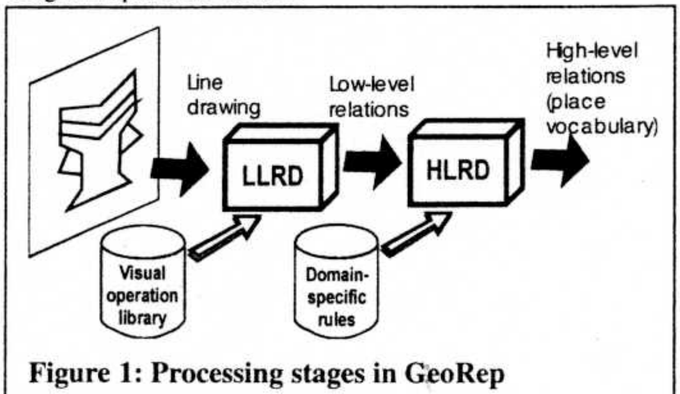

Figure 1: Processing stages in GeoRep

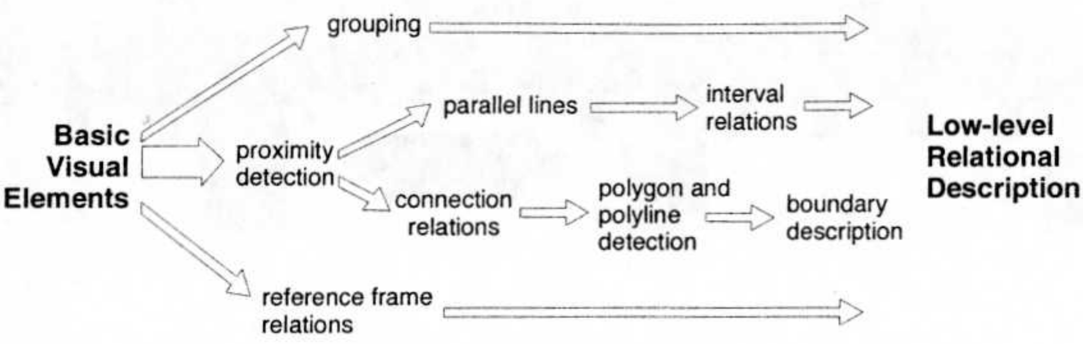

Figure 2: Processing of visual elements within the LLRD

The Low-Level Relational Describer (LLRD)

GeoRep's first stage, the LLRD, creates GeoRep's domain-independent spatial representation. Starting with the drawing's primitive visual elements, the LLRD detects and represents a broad set of early visual relations, using a set of *visual operations*. Conceptually, it applies these operations in parallel over the set of all proximate shapes in the visual field. These operations correspond to the sorts of relations that Ullman (1984) proposed are calculated by *universal routines*, which are visual routines that run in the absence of visual goals. To take advantage of logical dependencies between visual relation types, the LLRD pipelines the processing of visual operations so that more complex visual relations are computed after simpler ones (as in Figure 2). 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 strings. These primitive shapes, common across many different vector graphics formats, are sufficiently expressive to cover a broad range of line-drawn diagrams. The LLRD can also subsume these visual elements into polylines, polygons, and groups.

Computing proximity

Simple proximity is the core attentional mechanism within the LLRD. Because it is neither practical nor possible for the LLRD to detect all visual relations between all available combinations of visual elements, 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 the "hand". To determine which elements are proximate, GeoRep uses a proximity 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 object's size. For example, a circle has an area of influence covering everything within twice its radius. Similarly, for a line segment, the area of influence is the area extending out from the line segment for the length of the line segment. Elements are considered proximate when their areas of influence overlap. Areas of influence are always calculated as sets of circles and rectangles, in order to make overlap checking efficient.

Although later we show how this measure can be improved, this type of proximity measure has many advantages. It is easily constructed, efficient to use, and captures the intuition that large elements (such as a large rectangle or a long horizontal line) relate visually to many objects in the figure.

Running the visual operations

After computing the sets of proximate visual elements, the LLRD runs its library of visual operations on the figure. Each visual operation detects a specific visual relation that is part of early vision. All visual operations in the LLRD act on some combination of primitive visual elements, composite visual elements, and reference frames. Most also take a numeric strictness factor. If the visual operation takes more than one argument (i.e., if it does not detect some characteristic of a single visual element, such as the orientation of a line segment), then it is applied only to sets of elements that are proximate. These operations are relatively efficient, utilizing well-known algorithms from computational geometry (e.g., Glassner, 1990).

The LLRD notes which relations were detected in the figure, and generates a representation of those relations at the request of the HLRD. In the remainder of this section, we briefly survey the set of visual routines the LLRD currently detects.

Orientation and the frame of reference. One of the fundamental spatial qualities of vision is the frame of reference – the sense of what "up" is. Although there are many aspects of human vision that are orientation-invariant, experiments have shown that object orientation often has an effect on visual structure, and can affect the recognition of objects, especially abstract figures (Rock, 1973). The reference frame also affects the relations the LLRD produces. The LLRD generates several relations that depend on the reference frame, including *horizontal* and *vertical*, *above* and *beside*. Another pair of relations describe elements that occupy the same extent vertically or horizontally within the given reference frame.

Note that for humans, the frame of reference is usually gravitational and retinocentric, but need not always be. Humans use reference frames in flexible ways. Our visual system may change the frame of reference based on clues given in the scene, such as a preponderance of lines at the same orientation, elongation in a figure, or symmetry. For this reason, GeoRep initially assumes a gravitational reference frame, but can change its frame of reference when needed. When the frame of reference changes, relations based on the old reference frame are retracted, and the appropriate new relations asserted.

Parallel lines. The LLRD detects a drawing's parallel line segments, which models the ease with which humans detect parallel lines. However, in working with parallel lines in GeoRep's domains, we found that describing the parallel segments alone often doesn't adequately constrain the description of a drawing. Due to this problem, we extended the LLRD so that it also categorizes each pair of parallel line segments using Allen's (1983) interval relations. We found that these interval relations were useful in describing parallel segments because they could consistently describe the relative position of segment endpoints in a way that was invariant with respect to the reference frame. Admittedly, although we found interval relations in this context to be extremely useful, we know of no psychological evidence yet concerning whether such interval relations are detected in early perception. Note also that the symmetry of parallel line segments halves the set of valid interval relations available.

Connectivity, polygons, and boundaries. Connectivity between line segments is a key relation type detected by the LLRD. Line segments may connect as corners, as intersections, or as mid-connections. Arcs may connect with line segments as well, and such connections may either be aligned or misaligned. The LLRD also detects and classifies connections between line segments and curved objects, such as circles and ellipses.

The LLRD also performs simple path-following to detect polylines and polygons. It has long been recognized that closed shapes are important in perception. Despite the computational complexity of calculating closed objects (c.f., Ullman, 1984), humans appear to detect shape closure very early in perception, with evidence that they detect closure pre-attentively (Treisman & Patterson, 1984).

For polygons, the LLRD also analyzes the shape boundary. The boundary of a polygon is represented at several levels. Individual corners are marked as concave or convex. Groups of adjoining convex or concave corners are represented as protrusions or indentations. Inflexion points (indentations and protrusions) have been found to be very useful in visual tasks (Hoffman & Richards, 1984; Lowe, 1987), and recent studies has shown the importance of concavities in visual tasks such as symmetry judgment (Baylis & Driver, 1994; Ferguson, Aminoff, & Gentner, Submitted).

Grouping. Grouping is a very broad aspect of perception, because it must utilize some notion of similarity between grouped items. This makes grouping too broad an area for the LLRD to handle, and the current implementation is somewhat *ad hoc*. However, we are currently implementing a particular grouping mechanism that may get around this limitation by concentrating on grouping via pre-attentive factors.

Currently, grouping in GeoRep (which is only partially in the LLRD) depends on a set of *grouping rules* for each domain, which determine which pairs of items can potentially be grouped. The problem with this method, of course, is that a new rule must be written for each new group type. The grouping rules are not generative: a grouping rule for similar triangles can't automatically handle a more specific type of group (say, similar equilateral triangles) or a less specific type (groups of dissimilar triangles). However, this form of grouping has been adequate for the current set of diagrams, and allows for the quick construction of place vocabularies that use groups as part of their representation (e.g., a group of spline curves which represent steam rising from a hot liquid).

We are currently implementing a grouping mechanism that depends on pre-attentive grouping factors, such as similar size, orientation, and shape, to detect groups. These factors have been shown to allow items to be grouped pre-attentively (Julesz & Bergen, 1983; Treisman & Gelade, 1980). We believe that this mechanism will allow for a more flexible and cognitively valid grouping mechanism than is currently available in GeoRep.

The High-Level Relational Describer (HLRD)

Diagrammatic reasoning does not end with the generic visual relations that the LLRD produces. When we understand diagrams, we often depend on visual relations that are domain-dependent. For example, understanding the connectivity of a wiring diagram or the meshing of gears may involve detecting spatial relations that are not domain-general, but still are sufficiently spatial that they are best expressed in a diagram, rather than through text.

As previously noted, we call such sets of visual relations *place vocabularies* (Forbus, Nielsen, and Faltings, 1991).

The construction of place vocabularies is handled by the second stage of GeoRep: the HLRD. The HLRD contains a rule engine, which takes the structured representation created by the LLRD and applies domain-dependent rules to produce a place vocabulary for that domain. The rule engine used by the HLRD is Forbus and de Kleer's LTRE (1993), which is a rule engine built on top of a logic-based truth maintenance system (LTMS).

Writing rules in the HLRD is very similar to writing rules for any other rule-based system. What sets apart the HLRD's rule engine from other rule engines is the visual vocabulary that the rules use, which form a convenient abstraction layer for discussing domain-dependent visual symbols (e.g., the symbology of maps) and spatial relations. In addition, the HLRD rules contain special forms for delimiting the application of rules to proximate objects (just as the LLRD does), and for calls to the library of visual operations used by the LLRD. Part of the skill of diagrammatic reasoning in a domain comes from the use of a set of visual routines that map conceptual domain problems into visual terms, thus replacing inference with perception.

To build a new diagrammatic reasoner, one need only write a set of HLRD rules describing the diagrammatic conventions (such as the symbology) of the domain. These rules may trigger directly on visual relations already detected by the LLRD, but a rule may also perform limited top-down reasoning by directly calling visual operations based on the visual context. For example, an arrow-detector might check a potential arrow's "wings" to see if the two wings are of approximately equal length. The HLRD can also use domain knowledge given as a set of asserted facts, or may take information asserted about the domain by other reasoners (for example, information from a symmetry-detection engine about corresponding parts — see below).

While there is a domain-specific component to visual routines, there are also some visual routines that are widely used, and thus built into the HLRD. One rule set often used by the HLRD handles *representational links* between visual elements and what they represent in the diagram's domain. For instance, in a diagram understanding system, a trapezoid may represent a coffee cup. While the specific mappings from geometry to conceptual entity are somewhat domain-specific, there are important general properties of representational links that these rules express. For instance, each visual element should represent something, and each visual element can represent only one thing (excluding partonomic relations). These rules can often help resolve ambiguous interpretations, using simple heuristics (e.g., choose the interpretation which accounts for the greatest number of visual elements in the diagram).

The HLRD's reasoning engine can be used to provide insight into the diagrammatic reasoning process. Because these rules use a truth-maintenance system, the HLRD can explain why it believes that a particular set of visual elements represent a particular thing. For example, to explain why a particular polygon represents a coffee cup, it might explain that in this domain (as in the second example domain described in the next section) all trapezoids with a particular set of characteristics are assumed to be cups. Alternatively, the context of surrounding objects, such as "steam" rising from the polygon, might also be used to infer a cup.

Another advantage of explicit knowledge about representational links is that it can also be used to extend the place vocabulary. For example, given a drawing of two coffee cups, HLRD rules can figure out which cup has perceptibly greater volume by going back to the polygons that represent the coffee cups and checking to see if one of the cups is taller or wider than the other.

Once the HLRD has been run, and has generated a place vocabulary for the given diagram, this representation can be retrieved from the HLRD directly. Alternatively, it can be filtered by relation type to simulate different representation levels in the diagram. For example, one representation level might contain a representation of primitive visual relations such as intersections and interval relations, while a higher-level representation might include such things as the set of symbols in the diagram, or what those symbols represent.

The HLRD cannot generate arbitrary place vocabularies because it is limited by the capabilities of the LLRD. However, it has the advantage that, as long as the rules use only the LLRD's representation or its visual operation library, the resulting representation is likely to be cognitively plausible as something that would be noticed by the average human. The LLRD's representation is valuable because it provides an easy-to-use and extensible vocabulary. But it is also valuable because, if used correctly, it should tell us not just the relations a line drawing depicts, but the reasons why a human 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. The first project involved dozens of figures in a domain closely linked with human perception. The latter two projects, each tested on only a handful of representative diagrams, are ongoing efforts to apply GeoRep to practical diagrammatic tasks. We outline each project in turn.

Symmetry detection. GeoRep has been used as part of the MAGI symmetry-detection model (Ferguson, 1994; Ferguson, in preparation), which detects symmetry in

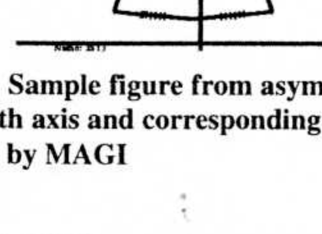

Figure 3: Sample figure from asymmetry study, with axis and corresponding parts as drawn in by MAGI

simple drawings, including functional drawings such as logic circuits. Here we describe how GeoRep has been used with MAGI to simulate the results of a set of psychological experiments. 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 concavities in the polygon boundary, had a significant effect on whether a polygon was perceived as symmetric. Subjects were more accurate at judging the symmetry of objects when asymmetries were qualitative, involving mismatches such as an unequal number of vertices on the left and right sides of the polygon, or corners that were concave on the left-hand side but convex on the right. The greater accuracy for qualitative factors was significant even accounting for a number of metric measures of asymmetry, such as difference in area between sides.

To simulate the experimental results, GeoRep was given the set of polygon figures (80 figures from experiment 1 and 160 figures from experiment 2), using the same line segment data used to display figures for human subjects. GeoRep generated the low-level relational description for each figure. This representation was then passed to the MAGI model, which judged the symmetry of the figure (Figure 3). MAGI, using the visual representation produced by GeoRep, succeeded in producing the same general pattern of symmetry judgments found in humans subjects. Like them, MAGI more accurately judged figural symmetry when asymmetries were qualitative.

Juxtaposition-based diagrams of simple physical phenomena. As part of the first author's dissertation, GeoRep is being used as part of a system called JUXTA (Ferguson, In preparation), which critiques simplified diagrams of physical phenomena.

For each diagram, GeoRep generates three different representation levels: a visual level, a physical level, and a process level. The visual level, designed to represent the abstract visual relations in the diagram, is built using the LLRD and some additional HLRD rules. The physical level, which represents simple physical entities and physical relations (e.g., immersion), uses simple structural description rules in the HLRD. The process level's rules find the physical processes inferable in the diagrams. A representative sample of each level is given in Figure 5.

Using MAGI to detect the repeated parts of the scene,

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.

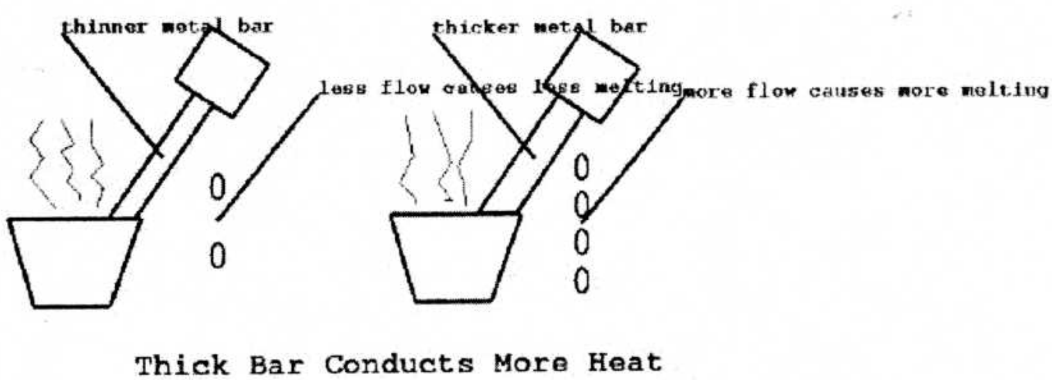

Figure 4: JUXTA's labeling of the aligned differences detected in a diagram, as related to the caption

JUXTA detects the physical and process differences between the two parts of the scene, and attempts to relate those differences to the caption (given to JUXTA as a qualitative relation). The result is a system that is able to critique the diagram based on how well the diagram's figure meets the expectations set in the caption. For example, based on the caption, JUXTA is able to label the critical differences found in the figure (Figure 4).

To perform this analysis, the distinction between representation levels is crucial. Visual differences are noted as relevant or irrelevant depending on the interpretation of the caption and how they affect the understanding the physical objects or processes in the drawing. Because GeoRep can represent multiple levels of description, JUXTA can distinguish visual differences that could affect someone's understanding of the diagram, and visual differences that, while noticeable, would not be confusing to the reader.

Course of Action Diagrams. As part of DARPA's High-Performance Knowledge Bases (HPKB) initiative, GeoRep is being used for spatial reasoning about Course of Action (COA) diagrams. These diagrams, used extensively 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, boundaries of available movement, and obstacles. Most work performed with COA diagrams currently is done by hand, using grease pencils on clear sheets of acetate. Diagrams are often redrawn several times to remove irrelevant details or change the level of description.

The COA reasoner we are building 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 HLRD 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 the recognition of arrows). Because most COA diagrams are constructed interactively, we are currently investigating how to extend GeoRep to handle interactive freehand sketches as input, instead of complete line drawings. Work continues on this reasoner, whose output will be used by several other technology developers within the HPKB initiative.

Limitations and areas for future work

GeoRep is being used in several projects within our group, and a reasoner based on it (the COA diagram interpreter), is being released for broader use by researchers within DARPA's HPKB initiative. If GeoRep shows itself to be useful and stable, we plan to release GeoRep more widely.

GeoRep has evolved considerably as various projects have made demands on it. We expect that it will continue to evolve as it is incorporated into future projects. While GeoRep has shown itself to be a flexible and useful tool in our own research, it has significant limitations. These limitations will need to be addressed to make the model truly general.

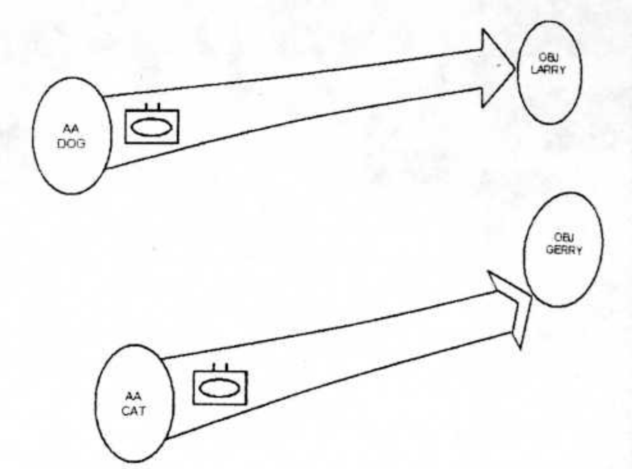

Figure 6: Example from a Course of Action diagram

First, GeoRep needs a more cognitively accurate model of proximity. While GeoRep's notion of proximity is fairly sophisticated, since it incorporates the relative shape and size of elements considered proximate, the human visual system's notion of proximity 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 currently investigated techniques for incorporating this model of proximity into the LLRD.

Because we plan to use GeoRep as part of an interactive sketching system, there are two other areas where it needs to improve. GeoRep currently processes drawings in batch mode. For sketching, it will need the ability to process drawings incrementally, allowing shapes to be added, modified or removed.

Second, GeoRep will need to handle ambiguous shapes based on context. Currently, GeoRep handles ambiguity in line drawings by using strictness factors in the LLRD's visual operations. Shape primitives, such as circles or arcs, are assumed to be unambiguous, and thus correct when read in. For sketching, where a single pen stroke might be a spline, line segment, or arc depending on the context, GeoRep will have to allow for the possibility of multiple competing interpretations, not only for visual relations but also for the shape classifications themselves.

Acknowledgements

This research was supported by the High-Performance Knowledge Bases initiative of the Defense Advanced Research Projects Agency, the Air Force Office of

Scientific Research, the Cognitive Science and Computer Science programs of the Office of Naval Research, and by the National Science Foundation, under the Learning and Intelligent Systems program. For thoughtful feedback and discussions of this research, we also wish to thank other members of the Qualitative Reasoning Group (including Tom Mostek, Robert Rasch and Bill Turmel) as well as two anonymous reviewers.

References

Allen, J. F. (1983). Maintaining Knowledge about Temporal Intervals. *Communications 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. 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. (in preparation). MAGI: A model of symmetry and repetition detection. .

Ferguson, R. W. (In preparation). *Understanding Diagrams through Analogical Encoding and Qualitative Reasoning.* Unpublished Ph.D., Northwestern University, Evanston, IL.

Ferguson, R. W., Aminoff, A., & Gentner, D. (1996). Modeling qualitative differences in symmetry judgments, *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* . Hillsdale, NJ: Lawrence 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.* Paper presented at the Fall Symposium on Computational Models for Integrating Language and Vision., Cambridge, Massachusetts.

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

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: The AAAI Press/The 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.

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.

Kosslyn, S. M. (1994). *Elements of Graph Design.* New York: W. H. Freeman and Company.

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: Notes on a model of sensory-cognitive interaction. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in Cognition* (pp. 279-307). San Francisco: W. H. Freeman and Company.

Palmer, S. E. (1989). Reference frames in the perception of shape and orientation. In B. E. Shepp & S. Ballesteros (Eds.), *Object Perception: Structure and Process* (pp. 121-163). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pisan, Y. (1995). Visual Routines Based Model of Graph Understanding, *Proceedings of the Cognitive Science Society* : 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. In S. Pinker (Ed.), *Visual Cognition* (pp. 97-159). Cambridge, MA: MIT Press.