Perceptually Based Learning of Shape Descriptions

Olya Veselova & Randall Davis

Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: In the past few years our group has developed sketching systems for several design domains, including mechanical engineering and software. Each new domain required writing new shape recognizers. We are currently developing a generic sketching system where a new domain can be added simply by providing textual descriptions of the domain symbols in a shape description language [1]. Each symbol is described in terms of geometric primitives (lines, arcs, ovals, etc.) and constraints between them (connects, parallel, above, horizontal, shorter, etc.). Simply drawing the new shapes, however, would be even more natural than typing the descriptions. The system we are building will automatically produce a textual description of the symbol from a user’s drawing, preferably, using only one example.

Motivation: Interaction with the system should ideally be as easy as it would be with another person, who we can teach a set of new symbols by drawing them. Consider how people learn new symbols, as the one in Figure 2.

Previous Work: The Electronic Cocktail Napkin learns symbols composed of primitive shapes by recording the spatial relationships between the primitives [4]. The system, however, records all the constraints it finds; the user has to remove the unnecessary ones manually.

Several systems, like Rubine’s GRANDMA learn single stroke gestures, using features like size, aspect ratio, convex hull, stroke length, number of corners, etc. [2]. We would like to be able to deal with more complex symbols consisting of multiple strokes. The kind of features used by Rubine would not be sufficient to represent the structure of these symbols.

Machine learning techniques used typically require many training examples, which would be quite tedious for the user.

Our work is more similar to Winston’s learning and generalization of structural descriptions [6]. However, as we explain below, in our work constraints have explicit relevance ratings which the system uses to hypothesize “near misses,” only asking the user to accept or reject examples.
**Approach:** The input to the learning system is the user’s strokes segmented into simple geometric primitives by a toolkit developed earlier in our group [5]. It is relatively easy to find all the geometric constraints between these primitives in the particular example. The difficult task is to pick just the relevant subset of constraints that characterize the symbol. How do people do it?

To rank constraints we use knowledge of how people perceive and remember geometric shapes. One of the most useful psychological studies in this area was Erich Goldmeier’s work on perceived similarity of shapes [3]. The following example illustrates the setup of most of his experiments. Consider the shape in Figure 3a. Ask yourself which of 3b and 3c is more similar to 3a? The majority of subjects chose c. Note that the left side of b is exactly the same as a. Yet, even though in c all the lengths and angles are slightly changed, it is considered more similar because of preserved symmetry.

Figure 3: Which is more similar to a?

Goldmeier conducted many similar experiments on different drawings, analyzing which constraints people preferred to preserve. Those constraints constitute the perceived geometrical essence of the symbol. He found that people mostly attend to properties that he calls *singularities* - special cases in the space of geometric configurations. For example a vertical (or horizontal) line is a special case of possible line orientations, parallel lines are a special case for possible angles between two lines.

Singularities are very important for defining the perceived structure of the symbol. On the other hand, exact values of non-singular values most often do not matter. This allows us to describe shapes in a few qualitative terms (for example, for angles: *obtuse*, *right*, or *acute*, rather than a numerical value). Goldmeier’s experiments also explore how singularities compare in relevance among each other.

Several other studies of memory and perception provide knowledge about which constraints people find important and which they ignore. This knowledge is used to rank all the constraints found in the user’s drawn example, after which the program produces a textual description of the symbol that includes only the highly relevant constraints.

**Future Work:** Perceptual ranking of constraints, though largely implemented, needs further improvement. We are also working on a suitable user interface for correcting the description produced from a single example. The system itself is intended to guide this interaction, using the ranking of constraints to isolate the ambiguity to a few constraints with medium relevance. Only those constraints need to be verified with the user.

After completing the learning and verification components of the system, we plan to extend it to deal with more complicated curves and to handle symbols where primitives can be repeated any number of times (such as a resistor symbol in electrical circuits).

**Research Support:** This work is supported by MIT’s Project Oxygen.

**References:**


