Human-Like Understanding of Two-Line Figures

Steven S. Hansen[1] and **David S. Touretzky**[1,2]
[1]Computer Science Department
[2]Center for the Neural Basis of Cognition
Carnegie Mellon University, Pittsburgh, Pennsylvania
shansen@andrew.cmu.edu, dst@cs.cmu.edu

Introduction

We describe a theory of perceptual understanding, implemented in Mathematica, that forms rich representations of very simple visual concepts. We suggest that this theory can represent any category distinction that a human is liable to make within its limited domain.

The Line Pair Domain

Concepts in the Line Pair domain are exemplar sets composed of pairs of line segments.[1] This domain is deceptively simple: although an exemplar can be fully specified by just four points, countless perceptually distinct variations exist. Figure 1 shows a small subset of these variations.

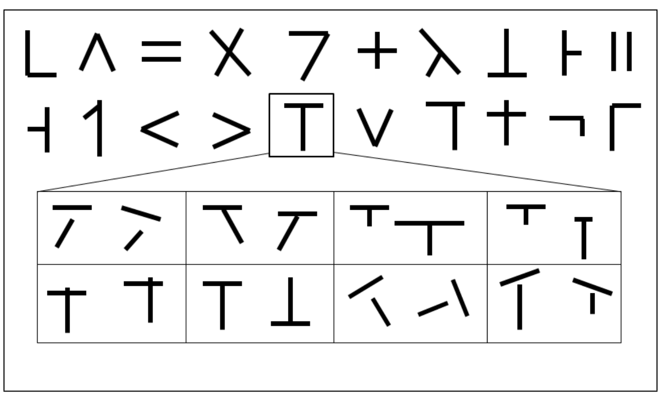

Figure 1: Some common shapes in the Line Pair domain. The expansion depicts a few easily differentiated concept classes within the range of "T"-like characters.

Our theory aims to account for understanding of two-line structures. We argue that the most fundamental aspect of understanding is that of categorization: if one can group instances of a domain into plausible categories one can be said to understand that domain. We futher claim that categorization within this domain is a recursive process of subdivision.

Despite its central role in our theory, categorization isn't the only way people demonstrate understanding of exemplar sets. In order to accommodate this, our theory supports a variety of other operations, such as outlier detection and concept extension. These draw heavily from the categorization produced, but widen the range of situations that the theory can successfully handle.

Content of the Theory

Three Perceptual Mechanism are necessary and sufficient to account for understanding in the Line Pair domain: binding, symmetry, and regularity detection.

The binding problem arises because line pairs are often asymmetric, the canonical example being the capital letter "T". Consider a "T" shape compared to a "T" that has been rotated 90 degrees. To appreciate the similarity in these shapes, we must know to compare the vertical line in the first shape to the horizontal line in the second. In our theory, bindings are determined via heuristic voting, since no one rule is adequate. A secondary binding problem arises in angular comparisons, where one needs to know the direction of rotation, and is handled in a similar manner.

Symmetry within the Line Pair domain consists of reflections, such that a figure has a symmetric relation with another if one of their features differs only by its sign. Figure $2A$ depicts this type of relation for the orientation feature.

Regularity is the measure of feature variation. While low variance features are often thought of as defining a concept, even features with large variation can be incorporated as part of a concept, in the sense that we learn that such features don't matter for this concept. Figure $2B$ shows a typical example, with Figure $2C$ showing that concepts may also rely on proportional relationships between elements.

Eight Distinguished Points can be regarded as the backbone of our Line Pair representation: the four endpoints, two midpoints, the projected intersection, and the point formed by projecting the closest endpoint of one line onto the other. While only the endpoints are needed to draw any line pair, they aren't sufficient to describe all possible line pair concepts. For example, in the case of the capital "T", the fact that the midpoint of the hat is bisected is one of the most

[1]We assume an unordered exemplar set to which subjects have full access. Modeling serial order effects, meaning exemplars are presented one at a time and the order of presentation affects the results subjects produce, would introduce additional constraints on our clustering algorithm, but is outside the scope of our investigation.

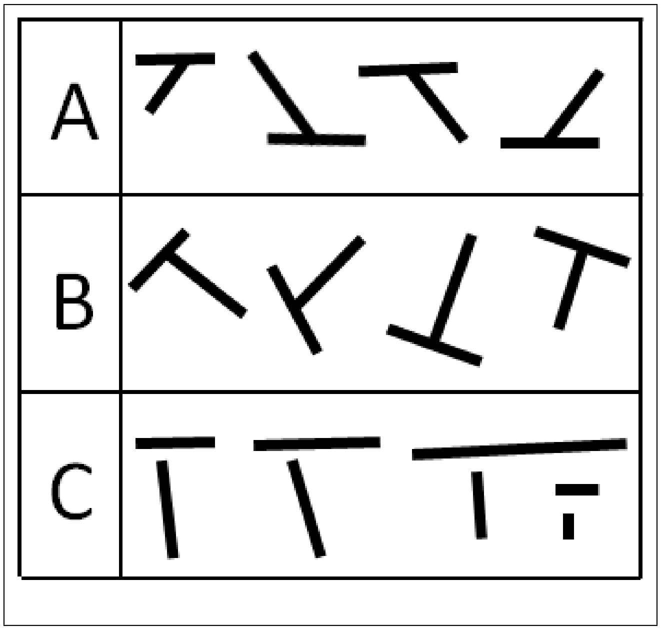

Figure 2: Three classes of "T"-like concepts highlight different perceptual phenomena. A: Figures either upright or inverted. This aspect of the concept class involves reflective symmetry. B: Figures whose angle from stem to hat remains nearly constant, but whose orientation relative to the observer is irrelevant, and whose stem length varies considerably. All such regularities, or lack thereof, must be captured in the concept's representation. C: Figures whose distance from stem to hat remains constant, but whose stem is a fixed proportional distance along the length of the hat.

salient features. Additionally, the canonical "X" and "+" rely on the relationship between midpoint and intersection point to be described in a natural manner.

Two Frames of Reference are used, since while all line pairs could be described solely in the frame of the viewer, proper generalization may require use of an object-centered frame. This point can best be illustrated by a class of "T" shapes whose stems are a certain distance below the hat. If one of the exemplars of this class were rotated, we would say that the stem is still the same distance "below" the hat, which implies use of an object-centered frame.

Three Types of Quantities used by our theory are lengths, signed distances, and signed angles. While these appear to be the only options for this domain, other domains may require additional measurement types, such as areas, or degrees of curvature.

A Combinatorial Feature Set primarily composed of the distances between all pairs of distinguished points is used to represent all perceivable aspects of a Line Pair. Seperate features are calculated for both reference frames, and for both types of regularity (constant and proportional), with separate encoding of sign and magnitude to capture reflective symmetry. The number of potential distance features is 295; there are also 3 possible angle features.

The Aspects of Understanding

Hierarchical Concept Induction forms the heart of the theory. Such categorization is akin to seeing the structure in an image. Given any exemplar set, our implementation partitions it into concept classes, recursively subcategorizing these classes if the data supports this. This is best thought of as solving an unsupervised concept learning problem, using operationalizations of human perceptual tendencies as heuristics.

Outlier Recognition and Correction allows one to recognize and eliminate that which seems to be unnatural or out of place, and can thus be thought of as one hallmark of understanding. We assume that concept classes are well-supported, so outliers are defined as those concept classes with very few members (possibly only one). The property of being an outlier may be further supported by a large distance to the nearest well-represented concept. Once outliers are identified, our program can correct them (merge them into existing concepts) by finding the minimal feature deformations necessary to be accepted into the concept class.

Prototype Generation is discussed at length in the literature on perceiving natural kinds, though its use in mid-level vision isn't as widely researched. Within our theory, this means constructing and storing the prototypical example of each category.

Concept Extension is an additional way of representing a learned category via a generator capable of producing an infinite number of new exemplars. Our program uses weighted combinations of feature distribution information to generate new instances that adhere to the "spirit" of the concept.

Structuring Justification The ability to explain one's choice of categorization, which consists of being able to explicitly identify the reasoning behind each concept class's formation, is important for full understanding. Our theory can give a symbolic justification for each concept class distinction made.

Discussion

Our choice of this micro-domain was inspired in part by the work of Hofstadter and his students on understanding visual analogies, most notably the Tabletop and Letter Spirit programs (Hofstadter 1996). We also gained insight from Feldman's work on perceptual grouping (Feldman 1997).

While we claim our feature set gives human-like performance for the Line Pair domain, other feature sets might yield similar results, and possibly generalize better to more complex domains. Curved lines with a limited number of concavities and non-selfintersecting polygons are two domains we have considered.

References

Feldman, J. 1997. Regularity-based perceptual grouping. *Computational Intelligence* 13(4):582–623.

Hofstadter, D. 1996. *Fluid Concepts And Creative Anaogies: Computer Models Of The Fundamental Mechanisms Of Thought*. Basic Books.