Human-Like Understanding of Two and Three 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

Abstract

We present a computational theory of mid-level vision within the Line Pair micro-domain, implemented in Mathematica. The theory assumes that categorization is the fundamental aspect of understanding, and that binding, symmetry, regularity detection, and proportionality detection are the mechanisms by which we categorize instances. We discuss additional aspects of "understanding" the domain, and suggest other domains to which the theory might transfer.

Introduction

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, at least 20 commonly used symbols can be created, and countless perceptually distinct variations exist within these classes. Figure 1 shows some common shapes from this domain, and elaborates on variations one might perceive for one of them. We claim that our theory solves the same computational problem, and thus captures the same type of information, as a human when exposed to an exemplar set within this simplified domain.

Our theory aims to not only account for how we differentiate between these two-line structures, but also to demonstrate how such an account helps a variety of processes typically associated with understanding. But what does it mean to understand a set of these structures? We view understanding as a family of processes, each of which widens the range of situations during which one could be said to 'understand' something. Six such processes are discussed here: categorization, outlier recognition, outlier correction, prototype generation, concept extension, and structuring justification. 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, and in fact the other 5 processes use the category representation system as their sole source of information without recourse to anything but minimal additions to the architecture.

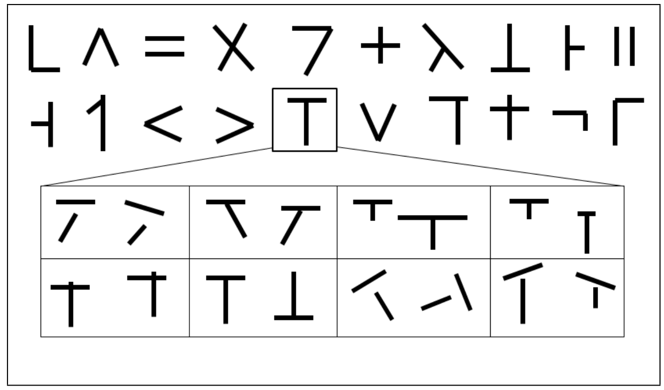

Figure 1: Some common shapes in the Line Pair domain. The expansion depicts a few easily differentiated concept classes within the range of "T"-like characters.

In our domain, "plausible" categories are those that are the most perceptually salient for humans. For example, looking at Figure 2, one typically groups the exemplars into "T" and "X" shapes. This isn't the only possible interpretation: we could group the instances by their size or relative angle, or any combination of factors. But since people find this one division so much more natural than the others, our theory will always choose it over the other interpretations. However, in a case where human opinions on the interpretation do differ, our theory will likewise be less confident in its decision, sometimes picking one over the other based on random factors.

Looking again at Figure 2, we can go further: the "X" shapes fall into two groups based on the relative sizes of the two lines, although they could also be categorized based on differences in their points of intersection. This raises two issues with regard to how our theory handles categorization: categorization is a recursive process of subdivision, and categorization can be ambiguous even when we strive for the standard human interpretation.

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

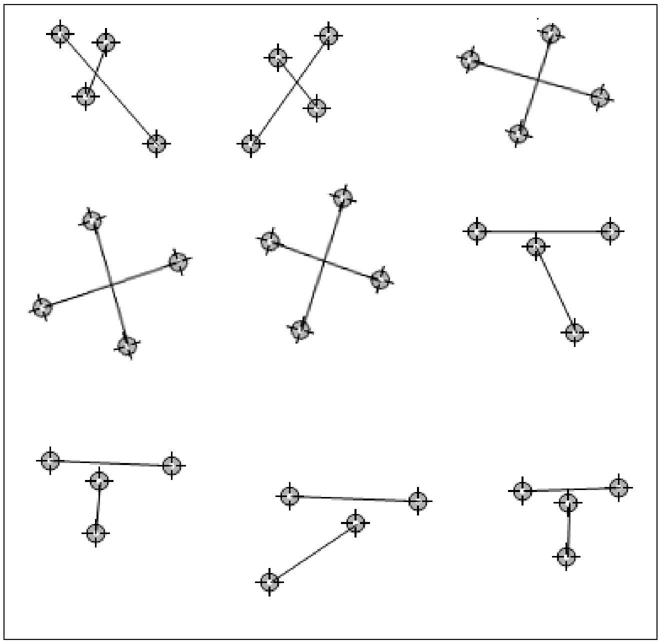

Figure 2: Exemplar set in the Line Pair domain, hand drawn in Mathematica as input to our program.

Content of the Theory

Three Perceptual Phenomena

We claim Reflection, Frames of Reference, and Scale Selection are perceptual phenomena that must be accounted for in any feature set sufficient for understanding in the Line Pair domain.

Reflection within the Line Pair domain consists of a type of symmetric relation, such that a figure is a reflection of another if one of their features differs only by its sign; Figure $3A$ depicts this type of relation for the orientation feature.

While all line pairs could be described solely in the frame of the viewer, proper generalization may require use of an object-centered frame of reference. For example, in Figure $3B$, we see a concept class where orientation with respect to the viewer is unimportant.

Finally, while some of the features measurable for two-line figures express absolute quantities, concepts may also rely on proportionally scaled relationships between elements. Figure $3C$ shows such a concept, where the stem of the "T" is always 1/3 of the way along the length of the hat.

Four Distinguished Points

Features within our domain can be regarded as relations between four distinguished points on a line segment: the two endpoints, the midpoint, and the point formed by projecting the closest endpoint of one line onto the other (this becomes the intersection point if the lines intersect). Initially only the endpoints seem relevant, as they are all that is needed to draw any line pair, but 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 is bisected is one of the most salient features. Additionally, the canonical "X" and "+" rely on the relationship between midpoint and intersection point to be described in a natural manner.

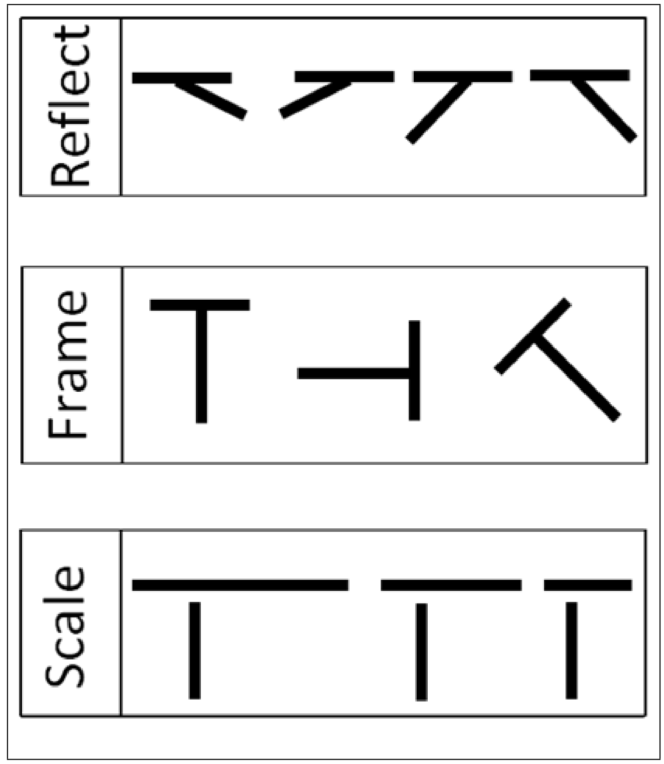

Figure 3: Three concept classes of "T"s which highlight different perceptual phenomena. A: Figures either upright or inverted. This aspect of the concept class involves reflective symmetry. B: Figures whose angle from stem to hat remains nearly constant, but whose orientation relative to the observer 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.

Three Types of Quantities

The quantities used by our theory are lengths, signed distances, and signed angles. It is one of the assumptions of our theory that nonlinear measurements aren't used by the human perceptual system for tasks of this nature.

The Binding Problem

When comparing two line pairs, several correspondence problems arise that require our representation to bind up to three variables to the pieces of a line pair. The bindings are determined using heuristic voting, since no one rule is adequate for all situations.

The primary binding problem comes from the fact that a line pair often has asymmetries between the two lines it contains, with the canonical example being the capital letter "T". Consider Figure $4A$, where a "T" shape is compared to a "T" that has been rotated 90 degrees. To appreciate

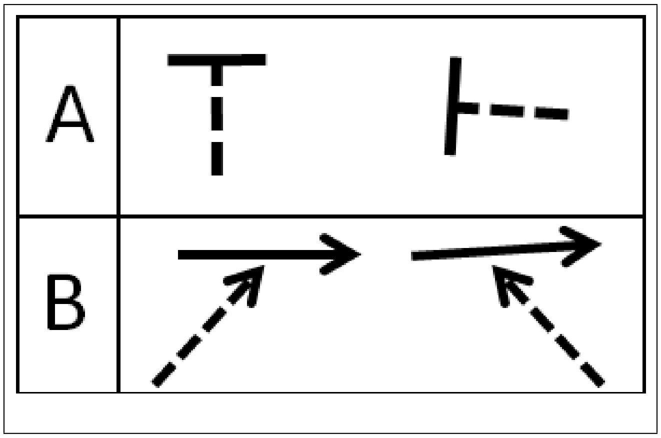

Figure 4: A: Two figures that need to be compared based on the roles each of their lines play. The stem role is shown via a dashed line, while the hat is solid. B: Two figures which need a description of angle between the stem and hat. Creating a label for the direction of each line disambiguates this description, and is shown by an arrowhead.

the similarity in these shapes, we must know to compare the vertical line in the first shape to the horizontal line in the second. But without a thoughtfully assigned binding, we might compare vertical line to vertical line and wind up thinking the two shapes are entirely dissimilar. We thus created roles for each line in our theory, with one segment designated the "stem" and the other the "hat". Following this intuition, the binding decision is made on the basis of heuristics, such as that the stem normally bisects the hat and not vice versa. Bindings are permitted to remain ambiguous when an exemplar is truly symmetric, such as an "X", in which case our algorithm must try both bindings to find which gives the better match.

The secondary binding problem arises in angular comparisons, where one needs to know the direction of rotation. For example, we might need to compare the relative angle of two stems with respect to their hats in order to differentiate the two line pairs in Figure 4B. Without providing the direction in which the stem points, we wouldn't be able to differentiate between cases where only the direction of rotation varies. Two additional bindings must thus be created. In the case of both the hat and the stem, one end point is labeled the start, and the other the finish, so that the direction can be seen as going from the start point to the finish point. Again following the intuitive bindings a capital "T" might have, we arrived at several plausible heuristics, such as that the stem should point *toward* the hat.

Combinatorial Feature Set

In order to adequately represent a line pair for purposes of our theory, we need a set of features that both uniquely identify the line pair and are plausibly perceivable by humans. We have identified a set of over 30 required features for this domain. The feature set is primarily the cross product of all distinguished points and all possible measurements, calculated in both reference frames. We assume that the human visual system computes these in parallel (Ullman 1984). Figure 5 demonstrates that the perceptually salient distinctions are immediately apparent in the feature values shown. Many of the features deal with distances between two distinctive points, one on the stem and one the hat. A noteworthy aspect of the feature set is its separation of sign from magnitude to capture symmetry. Reflective symmetry is expressed by magnitudes being similar while signs are opposite. So, in Figure 3A the magnitude of the relative angle and orientation features would have low variance, but the signs would have high variance. Another noteworthy aspect of the feature set is its use of redundancy to capture pairwise covariances between features. (It is an open question as to whether humans easily perceive covariances between larger subsets of features.). An example of this would be our 'difference of line lengths' feature, which can be used to see whether or not the hat and stem are consistently similar in length. Another invariant captured by our feature set is that of scale. The magnitude of any length or distance feature is accompanied by an additional feature encoding the ratio of the length or distance to the overall size of the figure (defined as the mean of the stem and hat lengths), thus allowing our theory to express relationships that are scale-invariant, such as Figure 3C.

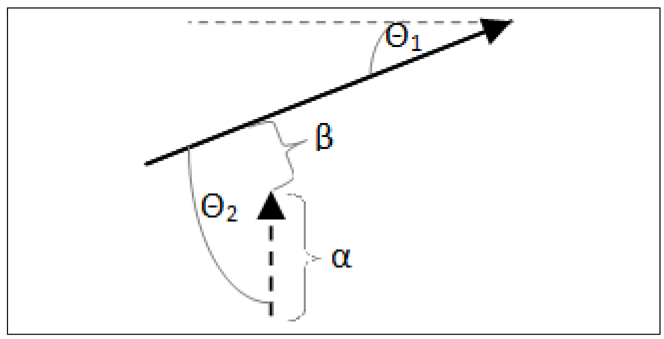

Figure 5: A small sample of the features calculated. α=stem length, β = minimum distance from stem to hat, θ_1=hat rotation, and θ_2 = stem angle relative to hat. Each feature corresponds to a dimension that most people would find salient.

Understanding as Hierarchical Concept Induction

As discussed above, we believe categorization is at the heart of understanding, and as such, it is at the heart of our theory. Categorization within our domain is akin to seeing the structure inherent in an image, so an understanding of the domain is best thought of as solving unsupervised concept learning problems, or equivalently as concept induction. This concept induction must also meet various specifications set by our above discussion of how humans perceive exemplar sets in our domain. Specifically, the concept induction process should be performed recursively until no perceivable divisions exist, and at each step the process should choose the most perceptually salient division as judged by an average human. It follows that the level at which a distinction is represented is indicative of the salience of that distinction.

So far, this notion of salience that underlies our theory has been underspecified. Let us operationalize the "salience" of a distinction between two line pairs to be the distance between the pairs in a feature space. Thus, the farther apart the two line pairs are in some yet to be defined feature space, the more salient the distinction between them. This definition was chosen for both its dovetailing with unsupervised learning theory and its correspondence to our intuitive notion of distinctions being more salient the more "different" the two objects are.

Categorization-based understanding within our domain can thus be thought of as a hierarchical process that uses operationalizations of human perceptual tendencies to guide it. In particular, our theory specifies definitions for the salience of perceptual distinctions, relevant features for comparison, and the threshold of automatic categorization in humans. With these specifications in place the process reduces to a relatively tractable unsupervised learning problem.

This process is illustrated by Figure 6, which shows the concept hierarchy produced by categorizing the exemplar set at the root of the tree. The two main branches of the tree indicate that the distinction between the "λ" shapes and the others was the first made, since these two instances were highly similar and distinct from the others along many dimensions. The two subdivisions of the non-"λ" node show that this concept class naturally clusters around two points in feature space. Specifically, these figures can be separated into "L", and "=" shapes. Notice that the node containing "λ" shapes has no children; this means that even though the elements of this concept class aren't identical, their differences are not judged to be significant.

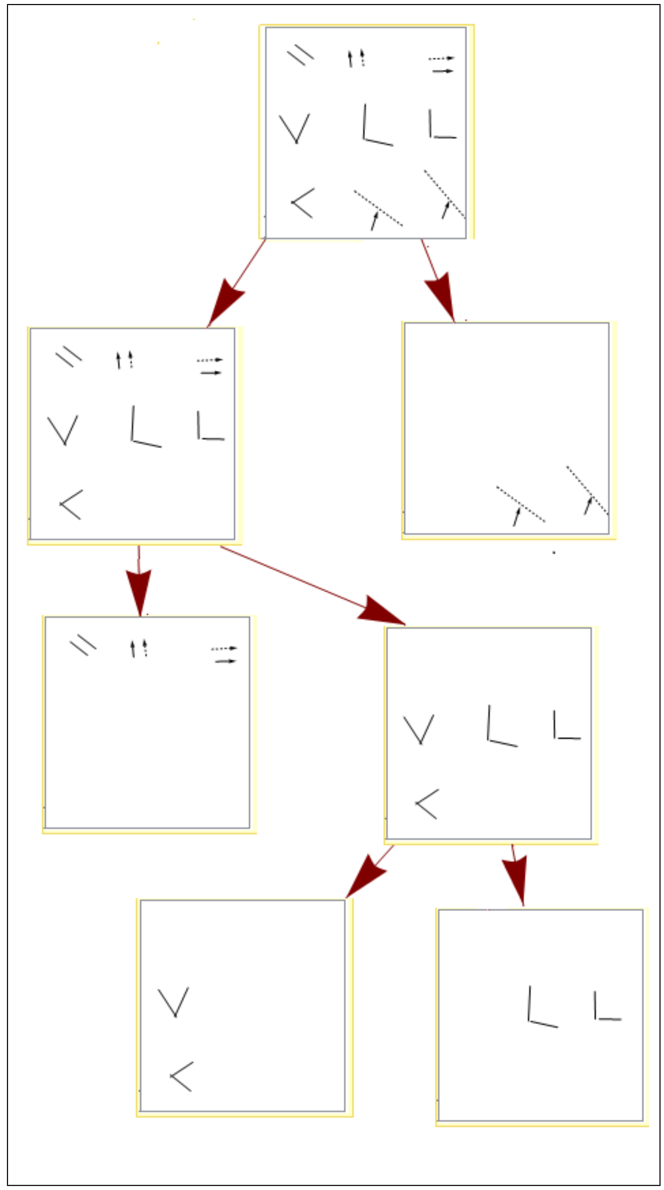

Figure 6: The hierarchical categorization of an exemplar set. The root of the tree is the original set, and the children represent sub-concept classes of the parent, with the depth representing their overall salience.

Line Triplet Heuristics: Context Sensitive and Semantically Plausible

The basic task of binding roles to individual lines in a Line Triplet sounds like a straight extension of the binding problem in the Line Pair domain. Namely, there is just an additional line, so the labels now depict the salient 'odd man out', the stem, and the two other lines, the primary and secondary leaves. All lines must still have an endpoint labeling, for the same reasons discussed in the Line Pair binding section.

However, this surface similarity belies a deep shift in the way binding must be thought when dealing with more than two lines. When dealing with a simple dichotomy, there is a sense in which there needn't be a 'correct' answer, so long as the results of such queries are always consistent, since the results are merely used for comparing 'stem' to 'stem' and not-stem to not-stem. But with three lines, consistency is obvious much harder to come by.

The binding heuristics used to solve this problem thus come in two varieties: Consistent heuristics are rather arbitrary and only promise that similar shapes are similarly labeled. E.g. lines point from left-to-right, the top-most line is the primary leaf, etc. These are the only type employed in the Line Pair domain. The other type is the semantic heuristic, each of which actually correspond to having a salient property (e.g. symmetry), and thus have meaning even when consistency isn't needed (even a solitary 'A' has a unique stem).

The bindings for Line Triplets also have an ordinal relationship with each order with regards to importance: the stem is more important to bind than the primary leaf, which in turn is more important to bind than the secondary leaf. The idea is that the best stem candidate shouldn't be labeled a leaf regardless of how much the primary leaf heuristics want him; it's always more important to have a good stem than a good primary leaf. This relationship means that we can (and should) decide on a stem before deciding on a primary leaf.

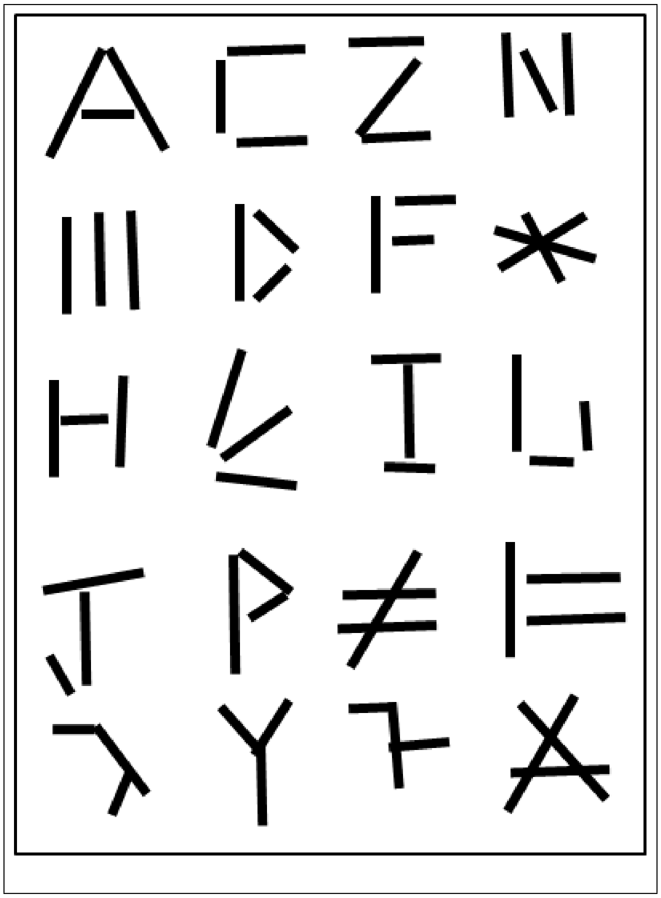

Figure 7:

In general, heuristics are needed for n-1 of the n labels. Thus the third label, the secondary leaf, is fixed by the heuristics for the stem and primary leaf, as the only remaining unlabeled line becomes the de facto secondary leaf. Likewise, the finish point is fixed via heuristics for the start point.

Now that some of the general properties of Line Triplet heuristics have been considered, the procedure for selecting the best bindings follows rather naturally. The best labeling for a specific triplet corresponds to the labeling that minimizes the weighted sum of the heuristics. The effective weight of each heuristic is determined by its a priori weight and the context weight. The a priori weight is simply our estimate of how powerful a heuristic is. The context weight is determined by the distribution of the heuristic's mismatches across the locally (bound using just the a priori weights) optimal labeling of each triplet in the neighborhood. The higher the mean of that distribution, the higher the average mismatch in the best case, and thus the lower the weight on that heuristic.

The rationale behind this is that we see how happy each heuristic is with the context-free choice of binding for each of the triplets in the neighborhood. Those heuristics that were generally very happy (low mean) are weighted higher. This weighting scheme makes any triplet with ambiguous context-free bindings decide based on the heuristics most satisfied in the surrounding triplets, such as in Figure 8.

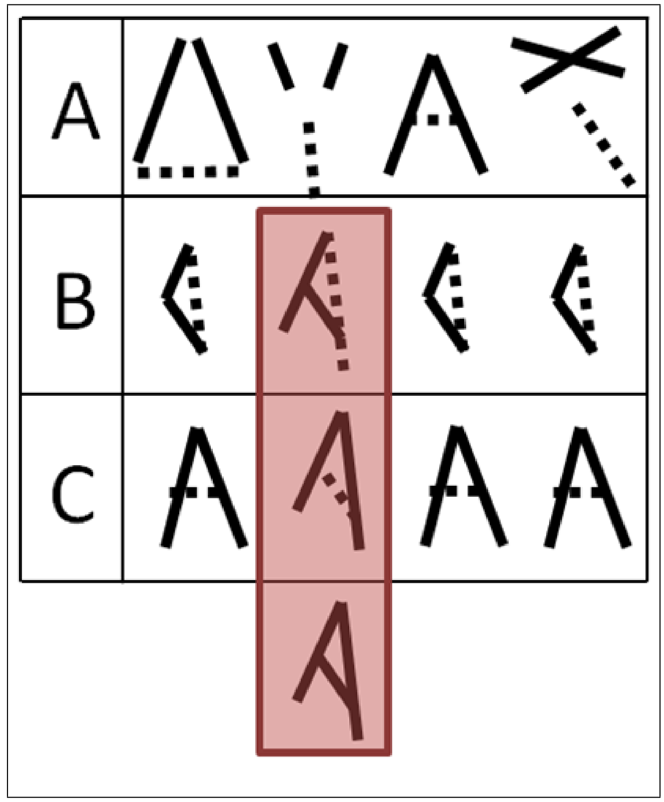

Figure 8:

Heuristic List

- Stem Heuristics
  - Topological Symmetry - Checks to see how closely the proposed stem topological feature with the primary leaf matches the topological feature of the proposed stem and secondary leaf. Here 'matches' means that in addition to the distance measured being similar, the two topological features should be symmeterical in some way. Normally this means the same distinguished points are involved. However, the stem can use different endpoints in the two features, as per a sort of reflective symmetry (think of the stem in a 'Y' versus an 'H').
  - Minimum Distance Symmetry - Dropping the symmetry constraint of the topological symmetry heuristic gives a heuristic that judges similarity based solely on how large the gaps between lines are (good for finding a stem in asymmetric cases, such as an 'h' like shape).
  - Rotational Symmetry - Checks how closely the proposed stem's angle with the primary leaf matches the angle between the proposed stem and the secondary leaf. The angle is the minimum angle of the leaf relative to the stem, and can thus vary from 0 to 90 degrees. The minimum is used to see the symmetry in identical and reflected angles as being equivalent for the purposes of

binding a stem (think of finding a stem for an isosceles triangle versus a π like shape).
  - Length Uniqueness - Checks to see how much the proposed stem's length differs from the lengths of the leaves. Ideally, the stem's length is the 'odd man out'. This check is made by taking the difference between the lengths of the proposed leaves, as the stem that minimizes this distances would be the best candidate.
- Primary Leaf Heuristics
  - Length - The longer leaf is judged to be primary.
  - Left-of-Stem - The leaf most to the left of the stem is judged to be more primary.
  - Top-of-Stem - The leaf most above the stem is judged to be primary.
- Endpoint Directionality Heuristics
  - Leaves point toward stem - Minimize the distance between the endpoints of the leaves and the stem.
  - Stem points toward leaves - Minimize the distance between the endpoint of the stem and the closest leaf.
  - Points to the right - The most rightward endpoint is the better finish point.
  - Points Upward - The most vertical endpoint is the better finish point.

Topological Feature Space: A Mechanism for Gross Similarity Distinctions

While the Line Pair domain could be dealt with using a single feature space to represent every instance of the domain, the increased complexity of the Line Triplet domain not only led to an increase in the size of the total feature space as expected, but also to the partitioning of the space into two separate feature sets based on how fine a grain of distinction they made.

The insight that led to this split was that it made little sense to use a rich feature set to compare an 'F'-like Triplet to an asterisk-like Triplet - the two shapes aren't similar enough to be meaningfully compared within the traditional feature space. The question then arises as to how one can determine when two Triplets are similar enough to use this feature space.

The intuitive answer, which has thus far held up empirically, is that we compare shapes only when their topologies are sufficiently similar. By topology, we mean the standard mathematical notion of what remains constant under continuous, non-tearing, transformations. We define the topological structure of a Line Triplet under the assumption that the two closest important points on each pair of lines in the Triplet are seen as 'glued together' and thus preserved under topological transformations.

Given this definition of topology within the Line Triplet domain, we can define a topological feature set which represents this structure and a means of comparing two Triplets of arbitrary topology. The topological feature set consists of the three pairs of important points that comprise the 'gluing spots' in a Triplet (one per pair of lines), along with the distance between each pair of points. Notice that this means

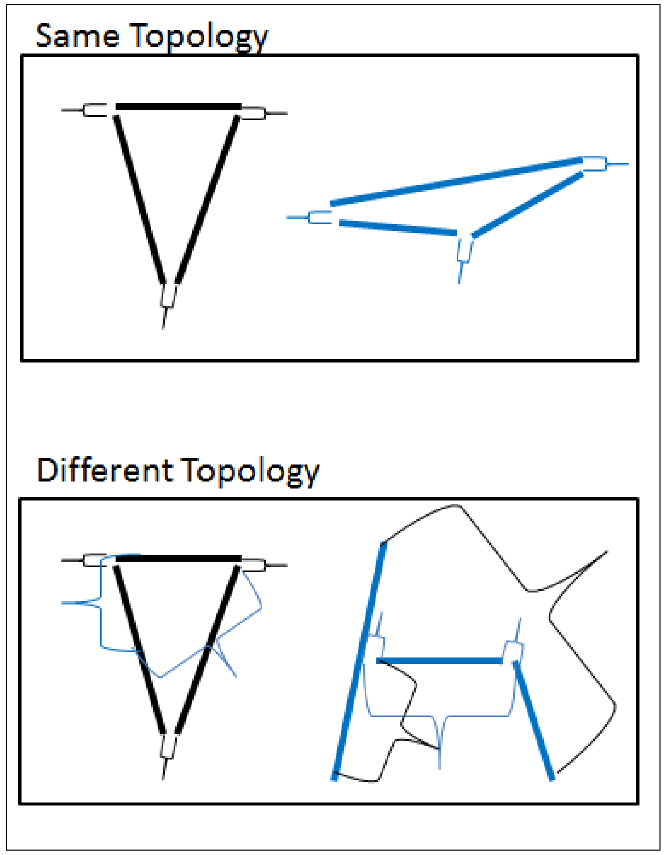

Figure 9:

two Triplets might not have the same topological feature set when they have different topologies. Thus, in order to compare two Triplets, we compute the union of the two feature sets for each Triplet and then compute their Euclidean distance, as shown in Figure 9.

This new feature set can now be used to compare two Triplets of different topologies, with the rich feature set used only when all of the Triplets being compared are nearly identical in topological space. However, this process is refined slightly by noticing that humans don't find topologically equivalent shapes to always be similar, such as the two shapes Figure 10 . Thus the topological feature set is supplemented by a set of rotational features representing the minimum angle between each pair of lines in the Triplet.

Fine Grained Feature Space: An Extension of the Two Line Case

Now that a mechanism is in place for comparing Triplets with radically different topologies, we can return to the original problem of creating a fine-grained feature set that that can capture the more subtle differences between Triplets. This majority of this feature space is concerned with capturing the features that relate one line in the Triplet to another. These are a direct extension of the Line Pair feature set, as the three possible pairings of lines within a Triplet can be represented by three instances of the Line Pair fea-

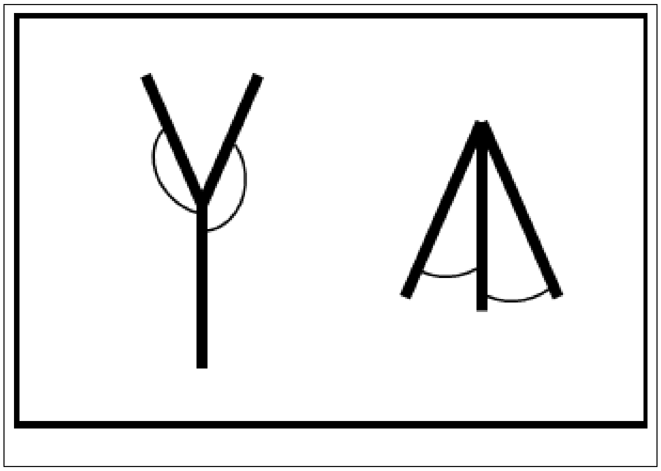

Figure 10:

ture set. As noted previously on the Line Pair feature set section, each Line Pair feature set is built combinatorially from the following Cartesian product of distances: (important point on line A) X (important point on line B) X (x or y dimension) X (frame of the page or of line A) X (absolute or proportional to the average line length) X (signor magnitude).

While this feature space captures most salient features, there are some three-way relationships the humans can readily pick up on. The most obvious of these is the illusory line relationship, in which the three lines of the Triplet 'line up' in some way. As the top section of figure 11 illustrates, this 'lining up' corresponds to three important points, one on each line, forming a straight line. A feature can encode this phenomenon by recording the total error that's keeping the three important points in question from forming a perfectly straight line. Since potentially any combination of mid and end-points could lend itself to such an illusory line, we must add features corresponding to following Cartesian product of errors: (important point on line A) X (important point on line B) X (important point on line C).

It should be noted that an illusory line isn't the only type of three-way relationship perceivable. The bottom section of Figure 11 illustrates depicts the different instance of the more general relationship, the shared contour. A great variety of such relationships can be perceived and many can be captured through the existing feature set via values among many pair-wise features. However, no explicit representation exists for these more complex contours within the Line Triplet feature set.

This omission is not an oversight, but rather a theoretical claim: non-linear contours (and other advanced three-way features) have no place within our domain, as they rely on knowledge outside of our domain. Specifically, most contours rely on implicitly interpreting a Triplet as having surfaces, and in some cases, depth values. The surprising dependence of the Line-Triplet domain on various facets of cognition is explored more generally in the discussion section.

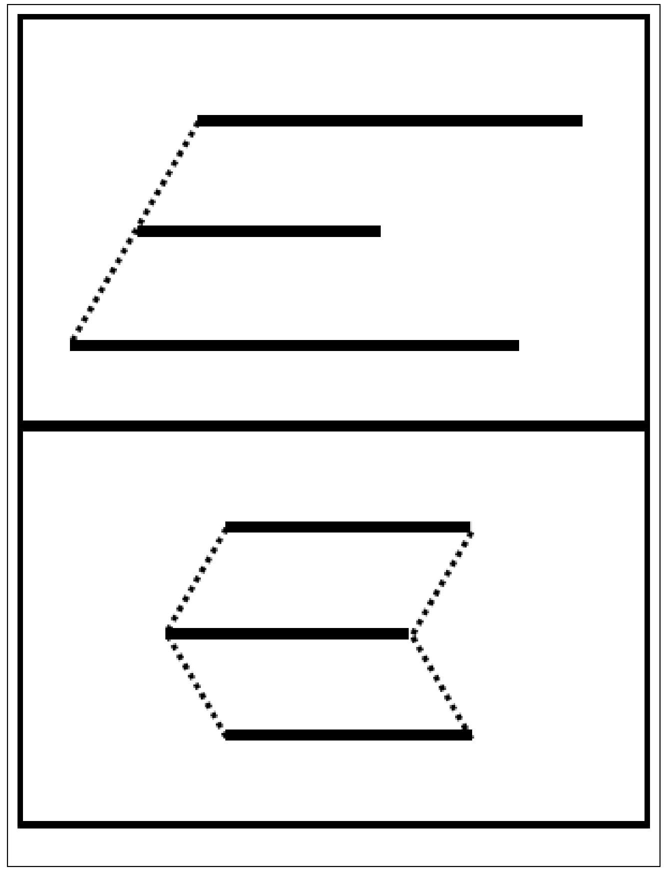

Figure 11:

Demonstrating "Understanding"

Once the hierarchy of concept classes has been formed, there are many ways to utilize this structured information. The five discussed here are outlier recognition, outlier correction, prototype generation, concept extension, and structuring justification. These examples are very closely tied to the concept representation system, using it as the sole source of information without recourse to anything but minimal additions to the architecture.

Outlier Recognition

One hallmark of understanding is the ability to recognize that which seems to be unnatural or out of place. Examples of this can be found in a huge range of domains, from pop-out effects in visual search (as well as search in the other modalities) to identifying incorrect board positions amongst chess experts. In our domain, 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. Using these two definitional aspects of outliers, we can detect outliers within our theory's framework by scoring every concept class based on its distance to the other concept classes and the number of elements it contains. Figure 12 demonstrates a case where one shape can clearly be seen as an outlier based on the types of support we described.

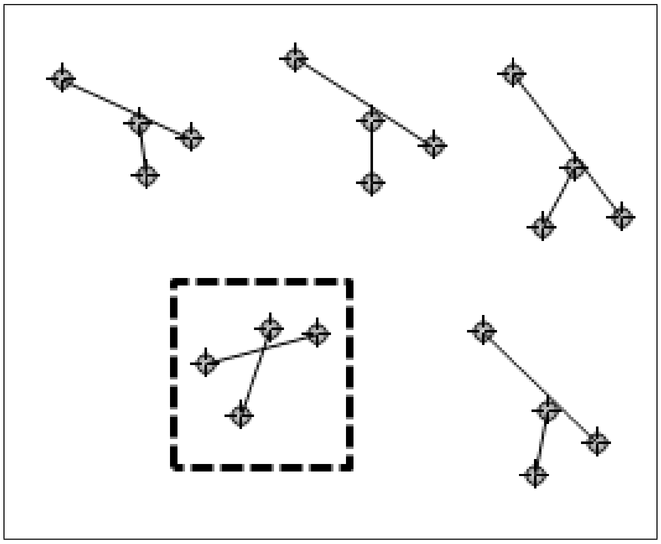

Figure 12: An exemplar set where the figure in the dashed box has been detected as an outlier.

Outlier Correction

Correction of an outlier is achieved by adjusting the feature values until the exemplars can be incorporated into a more acceptable concept class. Specifically, each outlier element is matched to the closest sibling concept class in feature space. If adding the element to the concept class doesn't disrupt the coherence of the concept, then it should be added. If not, we can find the minimal feature deformation necessary for the element to be accepted into the concept class. This feature deformation can be achieved by repeatedly selecting the feature contributing the most to the distance and adjusting it to better conform to the concept class.

Prototype Generation

Another dimension of understanding is the ability to form and use multiple representations of a concept, as the more plausible representations one has, the greater the confidence that one truly understands the domain. The most compact representation examined is to store the prototypical example of a category as being descriptive of the concept class. This variety of representation is discussed at length in the literature on perceiving natural kinds, though its use in mid-level vision isn't as widely researched.

A prototype based representation is advantageous in many situations, such as when new classifications must be made quickly, or when we simply no longer care about the specific set of line pairs. Such a representation can be made by constructing a figure based on the concept class's distributions for each feature. Since our feature set over-specifies the exemplar, at each step in the construction process we must evaluate how much each of the applicable features should contribute to the prototype. This contribution is made through a linear combination of the applicable features' predictions, weighted by their variances. Once constructed, this unique pair of lines is prototypical of its concept class. Note that while this means assuming a unimodal distribution for each feature, the hierarchical nature of our theory's categorization means that at some level of sub-classification, this assumption is likely to hold.

Concept Extension

An additional way of representing a learned category is to create a generator capable of producing an infinite number of new exemplars. Additionally, if one has such a generator for each concept in a hierarchy, one could generate similar exemplar sets to the one that led to the hierarchy's formation, or add to the set in such a way that extended the concepts it showcased.

Now all we must do to extend an exemplar set is to generate instances drawn from the stored distributions such that each concept class contributes new instances in proportion to its importance to the global structure. While many methods for deciding this importance are possible, the one we have found to yield the best results is to simply draw instances from the second level of sub classified concepts in proportion to the number of instances already in each class. The reasoning behind choosing only the second level concept classes to draw from is that drawing from deeper in the tree tends to yield extensions that too closely follow the exact structure of the initial exemplar set, rather than merely capturing its theme. Of course, one could always change what level to draw from depending on how broadly one construes the notion of pattern.

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. This aspect of understanding can be seen in various protocol analysis experiments, where the subject is asked to describe the process and reasoning used to come up with his conclusion.

If, during each categorization step, we eliminate all of the features with too little variance to be used in breaking a concept into sub classes, then we have enough information to explain our classification decisions in terms of similarities and differences in prototypical feature values. In other words, our theory can give a symbolic justification for each concept class distinction made by using the information gained in this new dimensionality reduction step.

This justification can be formed for any sibling concept classes A and B by defining two sets, AB_S and AB_D, where AB_S contains the features shared by A and B, and AB_D contains the features which distinguish A from B. Additionally, both sets must only contain well defined features (low enough variance to not require further sub-classification) so as to only describe the clearest differences and similarities between the categories.

We can populate the two justification sets by making a series of comparisons between the dimensions reduced in the superclass C (which we will call F_C), and those reduced in A and B (which will be referred to as F_A and F_B). The important insight that makes these comparisons work is that eliminating a dimension for clustering is equivalent to saying that that dimension is a common property of the to-be-clustered concept class, since its elimination was due to its lack of contribution to the overall conceptual distance between members. Likewise, if a dimension was selected for use in clustering, it must be a feature that the current concept class differed on considerably. So, $F_C \subset AB_S$, since $A \cup B \subset C$ and all elements in C have similar values for F_C by definition. We can then complete the set of features in AB_S by adding those features in $F_A \cap F_B$ that have similar distributions. Those that don't will make up AB_D.

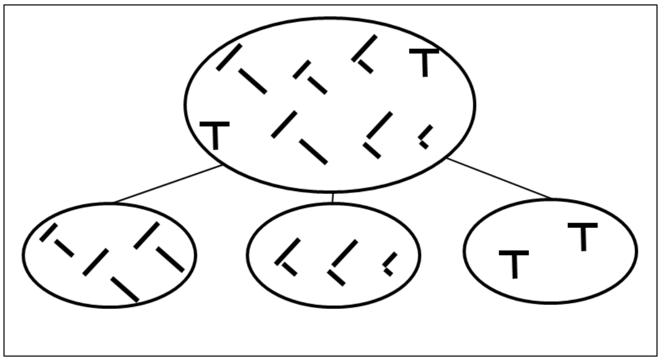

Figure 13: The top concept class has been split into three child classes. Since we eliminated the low variance features before making this split, we can form a justification for this particular split.

Take, for example, the task of constructing the justification for the bottom left and bottom center concept classes in Figure 13. We can say that these concepts were formed because, while they had the same hat lengths, relative orientation, and orientation of the stem relative to the horizontal, they were defined by different values of stem length and placement of the stem along the hat. The first part of this justification simply stated the features found to be in AB_S and the second part stated the features found in AB_D.

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 1997b).

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. A careful investigation of what people look at when classifying such simple images might lead to insights into how the human feature set differs from the one used here, and how those differences differentially affect performance in various domains.

Our implementation's judgments on the salience of various distinctions should correlate with human subjects' self-reported salience values. In other words, if our implementation labels some distinction as very salient, one should be able to confirm that most human subjects rate that distinction as very salient. So, even though our theory relies on human phenomenology as an evaluation criterion, our claims are falsifiable.

Additionally, the hierarchal representation used here seems phenomenologically plausible in the broad sense that rather than simply making a single classification, humans seem able to see the differences within each category. However, some runs of our theory's implementation go several levels deep in their sub-categorization. An investigation into how deeply people typically process exemplar sets in our, or a similar, domain might lead to important insights into the validity of keeping such a deep concept hierarchy.

Related micro-domains for theorizing

Our domain is but one of many in which one can form theories of visual object understanding. It is an open question as to whether other domains can be handled by the type of analysis we've described, or if the Line Pair domain is uniquely tractable. If our present theory could generalize to these other domains, the converging evidence might help determine which parts of our theory are applicable within real world domains.

A related micro-domain is single continuous lines with a limited number of curvature changes, which has previously been investigated by Feldman (Feldman 1997a).This would include characters such as the letter "S" and the numeral "6". We hypothesize that an exemplar set in this domain would be perceived by human observers in a way similar to our own domain. That is to say, this new domain lends itself to being understood based on a hierarchical categorization. So, while the feature set and the variety of symmetries would be quite different from that discussed in the current theory, at least some of the underlying theoretical framework would remain constant between domains. Moreover, any invariants between the two domains might provide insights into the general nature of the regularities humans are capable of detecting.

References

Feldman, J. 1997a. Curvilinearity, covariance, and regularity in perceptual groups. *Vision Research* 37(20):2835–2848.

Feldman, J. 1997b. 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.

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