Trees as memory representations
for simple visual patterns

JAMES P. CUNNINGHAM
Cornell University, Ithaca, New York 14853

The form of the memory representation for visual patterns consisting of a few line seg-
ments was investigated with three tasks. One task was the reproduction of remembered
patterns using pencil and paper. The order in which the subparts of those patterns were
reproduced was used to make suggestions about possible memory representations for the
patterns. The second task involved rating the "goodness" of subparts of the patterns.
These ratings were also interpreted as reflecting the underlying representations of the
patterns. In the third task, the time required to recognize that a small pattern was part
of a larger remembered pattern was used as further evidence about the memory represen-
tations of the patterns. The results from all three tasks are consistent with the theory that the
memory representations are tree structures. Methods for deriving and comparing tree struc-
tures are discussed.

This paper has two overall goals: (1) to give empirical
evidence that memory representations of simple visual
patterns can be modeled by a particular kind of net-
work, namely, a tree, and (2) to illustrate a methodology
that has recently been developed that can be used to
discover optimal tree representations in many research
domains.

Networks have often been proposed as models of the
way in which information is stored in memory (e.g.,
Anderson & Bower, 1973; Collins & Loftus, 1975;
Collins & Quillian, 1969; Norman & Rumelhart, 1975).
A graph, or network, consists of a set of objects called
nodes and a set of links between nodes. The pieces of
information to be stored in a memory network can be
represented by nodes or sets of nodes and the relation-
ships among the pieces can be represented by the way in
which the links interconnect the nodes.

Most of the empirical work associated with these
models has been in the general domain of language, but
there have also been several attempts to model visual
memory with networks. Both Palmer (1977) and Reed
(1974; Reed & Johnsen, 1975) present empirical evi-
dence that networks or network-like structures may be
appropriate for the representation of simple visual
patterns, but neither author proposed a specific repre-
sentation for a specific pattern. Anderson and Paulson
(1978) obtained data consistent with the idea that the
memory representation for faces is a network, but
presented only a schematic version of what that network

might be. Others (e.g., Palmer, 1975; Winston, 1975;
Baylor, Note 1) have presented more detailed theories
of how networks can be used to model information in
visual memory but have not been concerned with
providing empirical support for those theories. The
intent of this paper, on the other hand, is not only to
show that network representations of simple visual
patterns are plausible, but also to propose and empiri-
cally support specific network representations for
specific patterns.

Only a restricted class of network representations will
be considered in this paper: the class of free trees. A
"tree" is a particular kind of network in which there is
only one path (set of links) connecting each pair of
nodes. An example of how a tree can be used to repre-
sent some of the information in a simple visual pattern
is given in Figure 1. Figure 1a illustrates a pattern
consisting of six line segments, with each segment
labeled with a digit. (This is an example of the type of
pattern used in the experiments described below.)
Figure 1b illustrates a possible tree representation
for that pattern. The individual line segments of the
pattern are represented by the terminal nodes of the
tree (nodes with only one link attached), and the other
nodes of the tree have been given labels (LINE, ANGLE,
PATTERN) stating their significance in the organization
of the pattern. Note that only part of the information
in the pattern is represented by the given tree, since the
pattern could not be reconstructed from the information
in the tree. Additional information concerning the
lengths of lines and the sizes of angles would have to be
added in order for the representation to be complete.

There are several reasons why trees rather than
general networks will be considered as candidates for
modeling the memory representation of visual patterns,
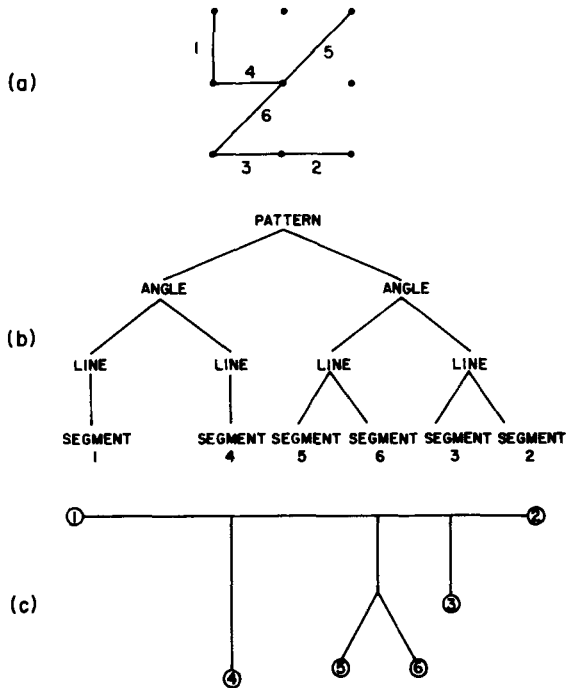with most of the reasons relating to the simplicity of

Figure 1. (a) A sample pattern with digits labeling the line segments. (b) A possible tree representation of some of the information in the pattern. (c) The corresponding uninterpreted weighted free tree.

trees. One reason is that the use of trees in the representation and recognition of two- and three-dimensional patterns has received considerable attention in the automatic pattern recognition field (e.g., Pavlidis, 1977, Chapter 10), indicating their potential sufficiency as models. Another reason is that trees generally require less "memory capacity" than do networks to represent the same information (Friendly, 1977), although it might take more time to retrieve the information from a tree. A third reason is that it is usually better to resort to a complicated model only after a simpler version of the model has been shown to be inadequate. And a final reason is that there are convenient methodological techniques for the development and evaluation of tree models but not for general network models.

METHODOLOGICAL APPROACH

There are basically two approaches to obtaining empirical evidence that the memory representation of some particular pattern can be adequately modeled by a network. One approach is to make predictions concerning the outcome of some experimental task that would hold if the memory representation for the pattern was any member of some class of networks. If the predictions are verified, that is evidence that some network representation is appropriate, but the verification does not identify which particular network is the most appropriate. This approach has been used, for instance,

by Palmer (1977) and Reed (1974). A second approach to obtaining evidence is to propose a particular network representation and test predictions based upon that representation. But what is the source of the particular representation from which the predictions are made? The traditional model-building method is to invent it using past experience and intuition as guides. (An excellent example of this approach is the work of Anderson and Bower, 1973, on memory for sentences.) Another method is to hypothesize a class of representations, find the particular member of that class that most closely agrees with the observed data, and then evaluate that particular representation. This last approach is the one taken in this paper.

One way to get empirical evidence about the structure of a memory representation is to use an experimental task that is sensitive to the distance between subparts of the representation. For instance, in the domain of semantic memory, if a subject is given word A and the first semantic associate that occurs to the subject is word B, then words A and B can be considered to be "close together" in the subject's memory representation. Any model of the subject's semantic memory representation should be consistent with these distances. In the experiments discussed below, several tasks are presented that can be used in the domain of visual memory to estimate the distances between subparts of a memory representation.

As an illustration of what is meant by "distance" in a network representation, in the structure shown in Figure 1b, Segment 5 is much closer to Segment 6 than it is to Segment 4. To be more precise, the links of a weighted free tree can be considered as having different lengths (weights), and the "distance" between two nodes is the sum of the lengths of the links in the path connecting the two nodes. One conception of the reason that links differ in length is that they differ in the number of "elementary links" that they represent but that elementary links could all have identical lengths. Thus, in Figure 1b, the fact that the link connecting Segment 5 to "line" is shorter than the link connecting "angle" to "pattern" could indicate that the former contains fewer elementary links than the latter.

The general research strategy used is the following. The smallest unit of a simple visual pattern of concern will be the line segment. It is assumed that the relevant aspects of the memory representation of a simple visual pattern can be described by a weighted free tree, with some of the nodes of the tree representing the line segments of the pattern. Three different experimental tasks are used to estimate the distance in the memory structure among the nodes representing the line segments of a pattern: reproduction of the pattern from memory, goodness ratings of parts of the pattern, and the recognition of parts of the pattern. The data from each of these tasks is summarized by a matrix of distances, each entry in the matrix being the empirically

estimated distance between the representations of two line segments in the memory structure. Each matrix of distances is then analyzed with a procedure described by Cunningham (1978) in order to find the particular member of the class of weighted free trees that best represents the distances. The procedure finds the tree that has the organization and link lengths that produce internode distances corresponding most closely (in the least-squares sense) to the empirically estimated inter-segment distances. The tree resulting from this analysis is the proposed model of the memory representation of the pattern associated with that distance matrix.

The adequacy of a particular tree as a model of the relevant aspects of the memory representation of a particular pattern is evaluated in several ways. One way is to compare the predicted intersegment distances with the empirically estimated intersegment distances by computing the correlation between them. A high correlation indicates a good fit of the tree model to the observed data. Another way of evaluating a tree is to see if its structure is intuitively interpretable (has face validity). An example of the sort of tree that could be obtained from the analysis of a distance matrix is given in Figure 1c. This tree corresponds exactly to the tree in Figure 1b, but no interpretation is imposed. (Note, for instance, that the node labeled "Pattern" in Figure 1b does not even exist in Figure 1c, because it has only two links connected to it; thus its location in no way affects the distances between other nodes.) If a free tree obtained from this analysis can be sensibly interpreted in a way analogous to the way in which Figure 1c can be interpreted as representing Figure 1b, then that is evidence that the tree is a good model of the memory representation of the associated pattern. Another, and perhaps the most powerful, way used to evaluate the adequacy of a particular tree is to see if the same tree representation can model memory for a particular pattern in a variety of experimental situations. If the structure of the memory representation of a particular pattern can be tapped with several different experimental tasks, and a different tree is needed to model behavior in each task, then either there is no common representation underlying performance on all the tasks or the tree is a poor model of that representation. Two different approaches to this use of converging evidence are employed. In one approach, results from one experimental task are compared with the tree representation derived from the results of a different task. In the second approach, a separate tree representation is derived from the results of each task, and the trees themselves are compared with each other.

Three experiments are presented, and the results of each suggest that memory for simple visual patterns can be adequately modeled by weighted free trees. The first experiment uses the order in which the segments of a pattern are reproduced from memory as a way of inferring memory structures for a relatively large number of patterns. In the second experiment, subjects repro-

duce a small number of patterns from memory and also rate the "goodness" of various parts of the patterns. In the analysis of this experiment, it is necessary to introduce several ways of comparing trees with one another, since the similarity of trees derived from the two different tasks needs to be evaluated. In the third experiment, the time required to recognize whether or not a set of line segments is part of a particular remembered pattern is used as information about the structure of the memory representation for that pattern. The results of this task are compared with the results of a reproduction task and a goodness rating task performed by the same subjects.

EXPERIMENT 1: REPRODUCTION

In this experiment subjects reproduced simple visual patterns with pencil and paper. The order in which the subparts of the patterns were reproduced can be used to make suggestions about possible memory representations for the patterns. The general idea is that subparts which have representations close to each other in the representation of the overall pattern also tend to be reproduced close to each other in time. So if two subparts of a pattern tend to be drawn one immediately after the other, this is evidence that the two subparts are represented close to each other in the memory structure for the pattern. The idea that output order reflects the way in which memory is organized is not new. For instance, Henley (1969) used it to study the semantic memory organization of animal names, and Reitman (1976) suggested its use in the study of memory for the location of pieces in the game of Go.

Method

Subjects. The subjects were eight undergraduates at Cornell University who were enrolled in an introductory psychology course and volunteered to participate in the experiment. All subjects were right-handed.

Stimuli. The stimuli were simple patterns of intersecting straight lines. Each pattern was made up of a 3 by 3 matrix of dots and 6 (of 16 possible) line segments connecting adjacent dots. Each of the 19 particular patterns used is shown in Figure 2. Each pattern had exactly two diagonal segments, tended to fill the area defined by the dot matrix, and had no disconnected parts. Similar sorts of stimuli were used in experiments by Derks (1972) and Palmer (1977).

Slides of the individual patterns were projected onto a screen from a distance of approximately 4 m. The projected patterns were approximately .5 m in height, and subjects sat approximately 2.5 m away from the screen. The patterns appeared as black lines and dots against a white background.

Procedure. Each subject was given a 57-page booklet. All of the booklet pages were identical and had three sections. At the top of each page was a 3 by 3 matrix of dots with all of the adjacent dots connected and with each line segment numbered with a digit from 1 to 16. In the middle of each page was a dashed-line version of the dot matrix with all 16 connecting line segments, and at the bottom was a row of blanks where line segment numbers could be written. The first blank already contained a segment number.

On each trial, a pattern was displayed for 10 sec, and then a blank white field was displayed. As soon as the blank white
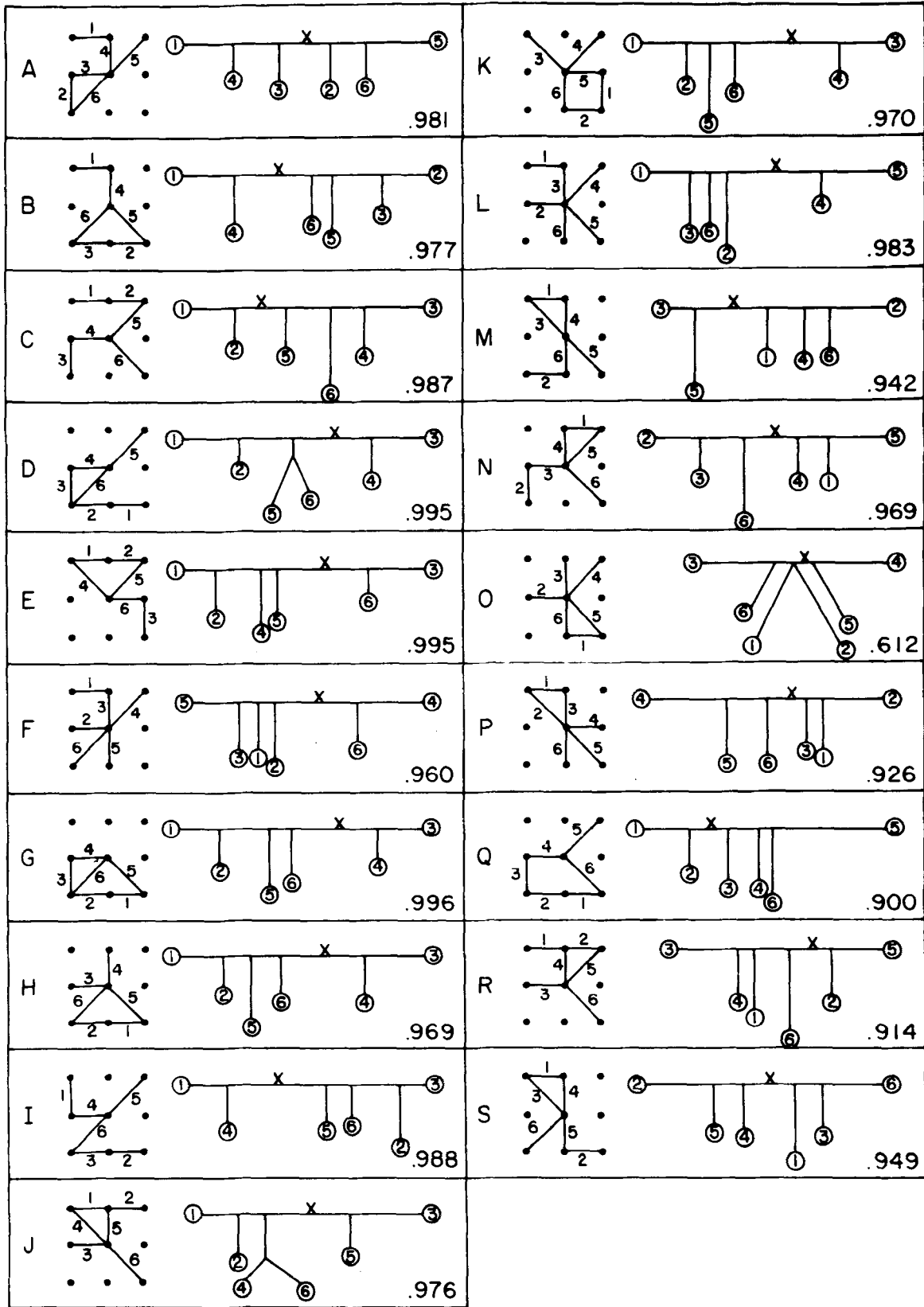
Figure 2. The patterns used in Experiment 1 (with each line segment given an arbitrary numerical label), the weighted free tree representation for each pattern (numbered nodes represent the corresponding line segments), and the correlations between the internode distances in the tree and the corresponding data values. For each tree, the longest link between non-terminal nodes is marked by an X.

field appeared, the subjects drew the pattern that they had just seen by filling in the appropriate parts of the dashed-line pattern in the middle of the current page in the booklet. They began drawing with the segment identified by the number in the first blank at the bottom of the page. They recorded the order in which the line segments were drawn by writing the sequence of line segment numbers in the rest of the blanks at the bottom of the page. After 30 sec of drawing time, the next trial was started.

The sequence of 19 patterns was shown three times in three different random orders. An experimental session lasted approximately 50 min, with two subjects being tested at a time. Each pattern had six possible line segments that could be identified as the one to be drawn first. Each subject reproduced each pattern three times, each time beginning with a different line segment. In each group of two subjects, one subject began drawing with each of three of the possible segments, and the other subject began with each of the other three possible segments.

Results and Discussion

The data consist of three ordered lists of segment numbers for each subject for each pattern. An ordered list of segments can be converted into a matrix of all pairwise distances between segments by assigning as the distance between two segments the absolute value of the difference between their numerical places in the ordered list. Thus, segments adjacent to each other in the list are separated by a distance of one, and segments at opposite ends of a six-segment list are separated by a distance of five. In order to derive a representation for each pattern it is assumed that: (1) the relevant aspects of each pattern can be represented in memory by a weighted free tree, with line segments at some of the nodes of the tree; (2) segments close together in the tree will tend to be close together in the reproduction sequence; and (3) all subjects have the same memory representation for any given pattern.

One matrix of empirically estimated intersegment distances was obtained for each pattern by averaging all of the distance matrices resulting from correct reproductions of that pattern. The best weighted free tree representation for each pattern was then found by using the procedure described by Cunningham (1978). Figure 2 contains the tree representations for each of the 19 patterns, and beside each tree is the product-moment correlation between the empirical distance estimates and the derived distances. Note that in general, the trees provide an excellent account of the data, with only one correlation below .90. (For comparison purposes, if the tree recovery procedure is applied to matrices of random numbers, correlations below .90 are obtained in more than 95% of the cases.)

A few comments about the appearance and interpretation of free trees might be helpful since many readers may be unfamiliar with these structures. The trees in Figure 2 were drawn in an arbitrary way with no interpretation imposed. Since the internode distances are not changed by the orientations of the links or the angles between the links, any other way of drawing the trees is acceptable, as long as it does not change the link lengths. In the case of trees this small, most of the infor-

mation in the tree is conveyed by the link lengths and the assignment of segment numbers to nodes. In order to see that this is true, consider all possible free trees with six terminal nodes, ignoring link lengths and labels on nodes. There are then only two different kinds of trees possible: Those similar to the tree associated with Pattern A and those similar to the tree associated with Pattern D.

One criterion for determining if the derived trees are plausible models for the memory representations of the patterns is whether or not the tree implies an intuitively reasonable organization of the pattern. For instance, the top node in Figure 1b seems to divide the associated pattern into two intuitively reasonable parts: Angle 1-4 and Angle 2-3-6-5. The location of that top node in the version of the tree in Figure 1c would be in the center of the longest link between two nonterminal nodes. A similar analysis of each of the trees in Figure 2 could be made. For each tree, the longest link between nonterminal nodes is marked by a small x. The reader can verify that, in almost all cases, the link seems to divide the pattern into two intuitively reasonable parts. In some of the cases in which the division seems less reasonable (Patterns C, O, and R), the tree contains two links between nonterminal nodes that are of approximately equal length, which could mean that it is not clear in which link a "top" node should be inserted or that at this level of analysis the pattern should be divided into three parts rather than two. On the other hand, for Pattern Q, it appears that the "top" node should be inserted in the long link attached to the node representing Segment 5. This "face validity" of the trees is evidence in favor of their use as models, since the lack of this validity would lead to serious questioning of their use as models.

A remaining question is the extent to which the trees reflect the underlying memory representation as opposed to rules for drawing with a pencil and paper. Two aspects of the experimental procedure were intended to reduce the influences of any such simple rules. The fact that subjects only had to fill in dotted lines, rather than draw the patterns freehand, should have decreased the use of heuristics intended to increase the accuracy of the drawing of parallel lines, straight lines, right angles, or lines of equal lengths. By forcing subjects to begin drawing with a different particular segment each time and by averaging over many reproduction sequences for the same pattern, it was hoped that the information remaining in the intersegment distances would principally reflect the memory organization. Another question concerns the effect of having subjects record the number corresponding to the sequence in which the line segments were drawn. Did this cause the subjects to artificially focus on the digits rather than a visual pattern and thus reproduce the segments in an order different from that which would have occurred under more normal circumstances? The current

experiment cannot address this point. However, that means that the trees are modeling only the behavior exhibited in this particular situation. Whether or not they can model behavior in other situations is addressed in the next two experiments.

The conclusion from this experiment is that trees provide plausible models of the memory representations for simple visual patterns. Furthermore, for each pattern, a specific tree was found that provided a good quantitative fit to the data and implied an intuitively reasonable organization of the pattern. The extent to which the trees really reflect memory representations can be assessed by the following additional experiments.

EXPERIMENT 2: PART GOODNESS

Experiment 1 provided information about the memory structures for a relatively large number of visual patterns. Experiment 2 began an in-depth analysis of the representation of four of those patterns. The subjects in this experiment reproduced patterns from memory and then rated the "goodness" or "naturalness" of all the two-segment parts of each pattern. One interpretation of a "good" part of a pattern is that it is one which is represented by a coherent unit within the representation of the whole pattern. The closer together two parts are within a representation, the more likely it is that they are in the same coherent unit. If this interpretation is correct, it should be possible to transform goodness ratings into distance estimates. Distance estimates obtained in this way are used to infer a weighted free tree representation for each pattern. The results of a reproduction task are then compared with the trees to see if they are consistent with each other. If they are consistent, this is evidence that the same representations underlie performance on both tasks and that those representations could be the free trees that are proposed. Since tree representations of the same pattern derived from two different tasks are compared with each other, several ways of making this sort of comparison are introduced.

Method

Subjects. The subjects were nine right-handed undergraduates at the University of California, San Diego, who participated in the experiment as part of a requirement for an introductory psychology course.

Stimuli. The basic stimuli were Patterns I, K, L, and N, which were used in Experiment 1 and are illustrated on the left in Figure 3.

Procedure. In the first part of the experimental session, the subjects reproduced the patterns in an abbreviated version of Experiment 1, using the same materials and procedure. However, only four patterns were used, each pattern was reproduced only once, and there was no specification of where the reproduction was to begin.

The subjects made goodness ratings during the second part of the experiment. Each subject was given a four-page booklet. Each page had one pattern drawn at the top and the 15 two-segment parts of that pattern filling the rest of the page. The
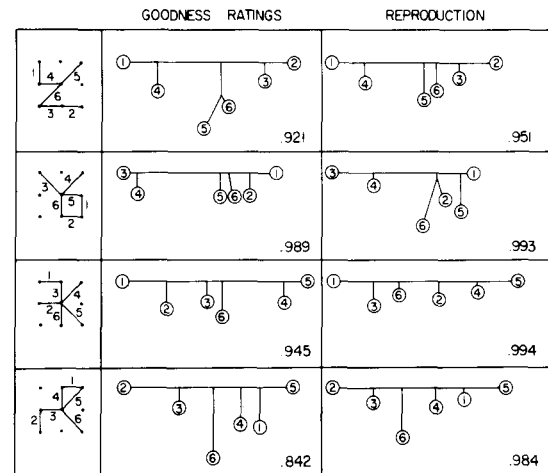

Figure 3. The four patterns used in Experiment 2 and, for each pattern, the weighted free tree representation derived from the mean goodness ratings and from the reproduction sequences, along with the correlation between the internode distances in the tree and the corresponding data values.

subjects were instructed to rate "how good or natural or obvious each part is within the figure at the top of the page." They were instructed to use a scale ranging from 1 ("very bad") to 10 ("very good") and to use the entire range of the scale for each pattern. The ratings were written in a space provided below each two-segment part. The two-segment parts were randomly ordered on the page, and the pages were ordered randomly for each subject.

Results and Discussion

Recovery of trees. First, trees were derived from the subjects' average goodness ratings and from the subjects' reproduction sequences. It was assumed that the goodness of a two-segment part of a larger pattern was inversely proportional to the distance between the two segments in the representation of the larger pattern. Each subject's goodness ratings were transformed into intersegment distance estimates by a negative linear transformation, yielding distances ranging from 1 to 10. The distance estimates were then averaged over subjects and the optimal free tree representation was found for each pattern. Those trees are shown in Figure 3. The correlations between the tree distances and the data values ranged from .842 to .989.

Intersegment distance estimates were also generated from the reproduction sequences by the procedure used for the analysis of the results of Experiment 1. Weighted free tree representations based upon these distances were found and are also shown in Figure 3. The correlations measuring the correspondence between the trees and the data ranged from .951 to .994.

As in Experiment 1, an examination of the organization of each pattern implied by the tree representations of that pattern can be made. In all cases, the implied organization seems to be intuitively reasonable, indicating that the trees have face validity.

Comparison of goodness trees and reproduction trees. An informal comparison of the trees generated from the goodness ratings and those generated from the reproduction sequences indicates definite similarities, but only for Pattern N are the two trees topologically identical. (Two trees that are topologically identical differ at most in the lengths of links between nodes and are said to have the same "structure.") But if the two trees for each pattern reflect the same underlying representation, then they should differ only because of random fluctuations in the data.

Before a more systematic comparison can be made, a formal measure of the amount of "structural difference" between two free trees is needed. A few preliminary definitions are required in order to state the measure. The insertion of a link into a tree is defined as the replacement of one of the nodes of the tree by two new nodes connected by a link. The deletion of a link from a tree is the removal of a link and the replacement of the two nodes that it connected by a single node. A move is defined as either the insertion or deletion of a single link. The structural difference between two free trees can then be defined as the minimum number of moves required to transform one of the trees into the other. (Boorman and Oliver, 1973, discussed several other possible measures.)

An indirect way of evaluating whether or not the trees differ only by chance is to compare the extent of the structural differences between the pairs of trees with the expected structural differences between pairs of randomly chosen trees. The average number of moves between the trees based on the ratings and the trees based on the reproduction sequences is 2.5. The probability of the mean difference between four pairs of randomly selected trees with six terminal nodes being 2.5 or smaller is .0025.[1] This indicates that the trees generated from the different data sets are structurally similar and may be reflecting the same underlying representation. (They are "significantly similar.")

There is a somewhat involved way to directly test the hypothesis that the differences between corresponding trees are small enough to be due to chance. Some of the details of this test are given in the Appendix. The result of this test is the probability that a structural difference of the observed magnitude or larger could occur by chance. For Patterns K and L this probability is very small ($p < .01$), so the structure of the trees based on the goodness ratings should be considered to be significantly different from the structure of the trees based on the reproduction sequences for those two patterns. For Patterns I and N the probability is large ($p \geq .32$ and $p = 1.00$, respectively), so for those patterns, the structure of corresponding trees should not be considered to be significantly different.

Consistency of goodness trees and reproduction sequences. A detailed comparison of the reproduction sequences themselves with the trees derived from the

goodness ratings may provide better information concerning whether or not the same representation underlies both tasks. In order to do this, it is necessary to define what it means for the drawing of a certain line segment in a sequence of segments to be consistent or inconsistent with a particular tree representation. The general idea is that reproducing a pattern involves the systematic traversal of the nodes of the tree representation of that pattern. A reproduction sequence is consistent with a given tree if traversing the nodes of that tree in the order specified by the sequence "makes sense." For the present purposes, the drawing of a certain line segment in a particular sequence of line segments is said to be consistent with a given tree representation if it is the first segment in the sequence or if the path through the tree from the previously drawn segment to this segment does not pass through any nodes attached by a single link to a not-yet-drawn segment. (Basically, this means that a sequence containing only consistent segments does not "skip over" any intermediate nodes.) Note that this definition of consistency does not guarantee that there is only one line segment that is consistent with a given tree and given partial sequence of segments.

A measure of the overall consistency between a group of reproduction sequences and a set of associated trees is the overall proportion of segments that are consistent but could have been inconsistent. Since only the middle four positions of each sequence could be inconsistent, only those positions are considered. The proportion of consistent segments can then be compared with the probability of a chance consistency in order to evaluate whether or not it is significantly larger. The probability of a chance consistency is computed by comparing the number of possible consistent segments to the number of undrawn segments for each position in each sequence.

For each subject, a weighted free tree representation for each pattern was derived from the goodness ratings of that subject. The reproduction sequences of each subject were then compared with the trees for that subject. The proportion of consistent segments was $93/132 = .705$. This proportion is significantly larger than the chance probability of .494 ($z = 4.85$, $p < .00001$). Therefore it can once again be concluded that the reproduction sequences and goodness ratings reflect at least similar underlying structure. When the reproduction sequences of the individual subjects were compared with the trees derived from the averaged goodness ratings (those shown on the left in Figure 3) rather than with the trees for individual subjects, the proportion of consistent segments increased to $116/132 = .879$, which is a significant increase ($z = 3.48$, $p < .0005$). This increase would be expected if most subjects had the same underlying representations. If that was the case, then the average goodness ratings would contain less error than the ratings of individual subjects, so the trees based upon those averaged ratings would more

closely approximate the true underlying representations. (When the proportion of consistent segments was broken down by pattern, the results were: 28/28 = 1.00 for Pattern I, 31/36 = .861 for Pattern K, 30/36 = .833 for Pattern L, and 27/32 = .844 for Pattern N.)

In order to demonstrate that the notion of consistency that has been defined is a reasonable one, the reproduction sequences can be compared with the trees derived from those sequences (the trees on the right in Figure 3). The proportion of consistent segments should be very high in this case, since the trees were derived from the information contained within the sequences. When this comparison was made, the proportion of consistent segments was found to be 120/132 = .909, which is the highest proportion found. This indicates that the definition of consistency and the method of converting reproduction sequences into distance matrices are compatible with each other.

The major conclusions from this experiment are: (1) the order in which the parts of a pattern are reproduced and the rated goodness of parts of a pattern may both be based upon the same underlying representation, and (2) likely candidates for the underlying representations of the particular patterns used in this experiment are the weighted free trees shown in Figure 3.

These conclusions are strongly supported for Patterns I and N, since for those patterns the structures of the trees based on the ratings and the trees based on the reproductions are significantly similar and not significantly different. The conclusions are less strongly supported for Patterns K and L, since for those patterns the structures of the corresponding trees are significantly different as well as being significantly similar. However, a more detailed comparison of the reproduction sequences themselves and the trees derived from the ratings revealed that each tree tended to be highly consistent with its associated reproduction sequences, even for Patterns K and L. This apparent discrepancy could be a result of the way in which the reproduction sequences were collected. In this experiment subjects drew each pattern only once and were not instructed to begin with a particular segment. As a result, for Patterns K and L there was considerable agreement between subjects in their reproduction sequences (many subjects producing the same sequence). Since the information in the sequences comes from the way in which they differ from one another, little information is obtained if most subjects produce the same sequence. Thus, the trees derived from the reproduction sequences of Patterns K and L should not be taken too seriously.

EXPERIMENT 3: PART RECOGNITION TIME

It is not clear whether the trees derived from the results of Experiment 2 should be considered to be models of memory representations, perceptual representations, or both. The drawing task involved memory over a relatively short duration (approximately 1 to 15 sec), while the goodness ratings did not necessarily involve memory for an entire pattern. Experiment 3 is addressed to the question of whether or not these trees can be used to model the memory representation of these simple patterns in a task that involves memory over a relatively long duration. The results from this task can be used to estimate the distances among the representations of the line segments of a pattern in memory. The trees derived from these estimates can be compared with the trees derived from goodness ratings and to reproduction sequences. As in Experiment 2, if the representations derived from the various tasks are consistent with each other, this is evidence that the same representations are underlying performance on all three tasks.

In this experiment, subjects began by learning a set of six-segment patterns. Then they participated in a task in which the time required to recognize that two line segments were both contained within a specific pattern was used as a measure of the distance between those line segments in the memory representation of that pattern. Weighted free trees were then derived from the distance estimates and compared with trees obtained from the other tasks.

Method

Subjects. The subjects were four graduate students and staff members from the Department of Psychology at the University of California, San Diego, who volunteered to participate in the experiment. All subjects were right-handed.

Stimuli. The basic stimuli were the four six-segment patterns used in Experiment 2. Each of the patterns was paired with a unique digit ranging from 1 to 4. For the first part of the experiment, each pattern was printed in black ink on the right-hand side of a piece of paper, with its associated digit on the left-hand side. In the part of the experiment involving the recording of reaction times, the stimuli were displayed on an oscilloscope screen (Tektronix 602) in a dimly illuminated soundproof booth and appeared as light green patterns on a black background. The stimuli were 3.2 cm in height and the subject viewed them from approximately 45 cm, but the subject's head was not fixed.

Procedure. Each subject was first given four pieces of paper, each containing a pattern and its associated digit. The subject was instructed to learn which pattern was associated with which digit. When the subject said that the patterns and associations had been learned, a test was administered. The test consisted of giving the subject the digits from 1 to 4 in a random order and having the subject draw the pattern associated with each digit. The experimenter observed and recorded the order in which the line segments were drawn. If a mistake was made, the subject was given another opportunity to learn the patterns and associations, followed by another test. The procedure was repeated until the subject performed perfectly in the test. This portion of the experiment required less than 10 min.

The second part of the experiment consisted of a sequence of trials on which the subject was shown a digit followed by a two-segment pattern and was asked whether or not the two-segment pattern was contained within the six-segment pattern associated with the digit. Specifically, an experimental trial had the following form. The 3 by 3 matrix of dots that was a part of all patterns was displayed on the oscilloscope screen continuously between trials. When the subject was ready to start a trial,

(s)he pushed a button; after a 500-msec delay, the dot matrix was replaced by a digit. The digit was displayed for 3,000 msec, during which time the subject was supposed to "bring to mind" the pattern associated with the digit. After the digit disappeared, the screen remained blank for 1,000 msec, and then the test stimulus was displayed. The test stimulus was a 3 by 3 dot matrix with two line segments connecting adjacent dots. The test stimulus was displayed until the subject responded either "yes," indicating that the two line segments were contained in the pattern associated with the digit, or "no," indicating that they were not. The subject responded by pressing one of two buttons on a table in front of him (her). When the subject responded, the test stimulus disappeared, and either a Y or an N was displayed for 3,000 msec, indicating the correct response for that trial. The time from the onset of the test stimulus to the subject's response was the recorded reaction time.

For each of the four patterns, each of the 15 possible two-segment parts of the pattern was used as a test stimulus for that pattern four times. So, for each pattern, there were 60 trials for which the correct response was "yes." There were also 60 test stimuli for which the correct response was "no" for each pattern. Thirty of the test stimuli were randomly chosen from the 60 possible test stimuli having one segment in common with that pattern, and 30 were randomly chosen from the 45 possible test stimuli having neither segment in common with that pattern. The 480 experimental trials for each subject were randomly ordered and then divided into three blocks of 160 trials.

The third part of the experiment involved making goodness ratings of various two-segment patterns. Each subject went through the same rating procedure as the subjects of Experiment 2. In addition, they rated the goodness of the 20 possible unique (up to reflections and rotations) two-segment patterns, independent of any context. That is, they rated the goodness of each pair of segments as an independent pattern.

Each subject participated in three experimental sessions on 3 separate days, each lasting approximately 1 h. In the first session, the subject learned the patterns and associations and then had 30 practice reaction time trials, followed by the first block of 160 experimental trials. Each of the next two sessions began with the subject's refreshing his memory for the patterns and associations, followed by 20 practice reaction time trials and a block of 160 experimental trials. When a subject made an error on an experimental trial, that trial was repeated after one half of the remaining trials in that block had been executed. At the end of the third session each subject made the goodness ratings.

Results and Discussion

Effect of the part itself on part recognition time. The first question addressed is the extent to which the time required to recognize that a two-segment pattern is part of a larger pattern depends upon the two-segment pattern itself, rather than on its relationship to the larger pattern of which it is a part. This question is addressed in two ways. The first way is to consider cases in which the identical two segments are parts of more than one larger pattern. In particular, Patterns L and N have four line segments in common and thus six two-segment parts in common. A three-factor analysis of variance with the factors of part (six levels), pattern (L or N), and subjects showed that there were no significant differences in reaction time due to which part was being recognized $[F(5,15) = .67]$, but that there were differences due to which larger pattern had been "brought to mind" $[F(1,3) = 14.51, \ p < .05]$. This suggests that the important determinant of the recogni-

tion time is the relationship of the part to the larger pattern, and not the structure of the part itself.

The second way of addressing this question is to compare the part recognition times with the two different goodness ratings made for each part. Each two-segment pattern was rated for its goodness as a part of a six-segment pattern and for its goodness as an isolated pattern. The correlation between the goodness-as-a-part ratings and the normalized median reaction times for the corresponding segment pairs was computed for each subject. The average value of this correlation was $-.589$, indicating that good parts were recognized faster than bad parts. This correlation was significant for all subjects [with t values ranging from $t(58) = -11.08$, $p < .00001$, to $t(58) = -3.68$, $p < .001$]. The correlation between the goodness-as-a-pattern ratings and the corresponding normalized median reaction times was also computed for each subject and had an average value of $-.262$. However, if this relationship was assessed by computing the correlation between goodness-as-a-pattern ratings and normalized median reaction times with goodness-as-a-part ratings partialed out, the average correlation became $-.035$. The value of the partial correlation was significant for only one subject [t values ranging from $t(58) = 2.33$, $p < .05$, to $t(58) = -.59$, $p > .50$]. An average of 36.9% of the variance of the median reaction time can be accounted for by the goodness-as-a-part ratings, and only an average of 3.9% more variance is accounted for by the goodness-as-a-pattern ratings. In other words, a part of a larger pattern tends to be recognized quickly if it is a good part of the larger pattern, but being a good pattern unto itself tends to be unimportant. This is additional evidence that the primary factor that determines the time required to recognize a two-segment part is the relationship between the segments in the memory representation and not the two segments' relationship in the physical world.

Reaction time trees and goodness trees. If the time required to recognize a two-segment part is determined by the relationship between the segments in the memory representation, then those reaction times can be used to make inferences about the memory representation. In order to make those inferences, it is assumed that the patterns used in this experiment are represented in memory by weighted free trees, with the line segments at some of the nodes of the trees. It is also assumed that the time required to recognize that a two-segment pattern is part of a six-segment pattern is a function of the distance between the two segments in the tree representation of the larger pattern. The farther apart the two segments are in the memory representation, the longer the time required to recognize that they are both part of the memory representation. This means that the "yes" reaction times can be used as intersegment distance estimates. For a given subject, there are four replications of the "yes" reaction time for each pair of segments, and the median of those four replications was

used as the estimate of the distance between those segments. This yielded a distance matrix for each pattern for each subject.

For each subject, a weighted free tree was derived for each pattern. In general, the trees agreed very well with the data, with correlations between derived distances and empirical estimates ranging from .762 to .981, with a mean of .910. The individual subjects' reaction time distance matrices were averaged, and an overall weighted free tree representation was found for each pattern. The four trees are shown on the right-hand side of Figure 4, along with the correlations between the derived tree distances and the original mean distance estimates.

Weighted free trees were derived from the average goodness ratings in the same manner as was used for the analysis of Experiment 2. Those trees are shown on the left-hand side of Figure 4, along with their correlations. The trees based on the group reaction time data were compared with trees based on the group goodness ratings. The average structural difference between a goodness rating tree and the reaction time tree representing the same pattern is 2.75 moves. The probability of a value that low or lower occurring by chance is approximately .0064. This indicates that although the pairs of corresponding trees are not identical, they are similar and may be reflecting the same underlying structures. Tests of the hypotheses that the differences between corresponding trees were due to chance variation yielded nonsignificant results for all four patterns (Pattern I, $p = .33$; Pattern K, $p = .13$; Pattern L, $p = .28$; Pat-

tern N, $p = .34$). This is further evidence that the same tree representation may be underlying performance on both tasks.

When the trees are compared with the patterns, the following interpretations can be applied to both the goodness rating trees and reaction time trees. For Pattern I, the angle 1-4 is separated from the angle 5-6-3-2. The angle 3-4 is separated from the square 5-6-2-1 in Pattern K. Pattern L consists of three units, the parallel lines 1-2, the straight line 3-6, and the angle 4-5. The goodness rating tree for Pattern N separates the triangle 4-1-5 from the rest of the pattern, but the reaction time tree is uninterpretable. This difficulty in interpretation is caused by individual differences in organization of the pattern and is discussed later in this section.

If the trees based on goodness ratings from this experiment are compared with the trees based on goodness ratings from Experiment 2 (shown in Figure 3), remarkable agreement can be seen. In all but one case the corresponding trees have identical structures and similar link lengths. It appears that trees based on average goodness ratings are very robust between experiments. Even though the subjects of this experiment were more sophisticated and had more exposure to the patterns than did the subjects of Experiment 2, the trees based on their average goodness ratings are almost identical.

Another question of interest is how representative of the trees of individual subjects are the trees based upon the group data. The average structural difference between the tree of an individual subject and the group tree for the same pattern is 2.56 moves for the trees based on goodness ratings and 2.25 moves for the trees based on reaction times. These differences are about the same as the differences between tree representations of the same pattern based upon different dependent measures. This indicates that the trees based upon the group data are similar to each of the trees of individual subjects, suggesting that they may be reflecting the same (or at least similar) memory structures.

The similarity between the tree based upon the group data and the individual trees is least for Pattern N. For that pattern the average structural difference between the individual trees and the group tree is 3.25 moves for the trees based upon goodness ratings and 4.00 for trees based upon reaction times. (Not surprisingly, Pattern N has the lowest correlations between the derived distances and the data, with the reaction time tree having a correlation of only .678.) This lack of agreement is caused by the lack of agreement between individuals. As examples, the trees derived from the median reaction times of Subjects S2 and S4 are shown in Figure 5. Pattern N seems to have been organized in one of two ways, depending upon the subject. One way was as a pair of "steps" made out of Segments 2, 3, 4, and 1, with a "brace" made out of Segments 5 and 6; the
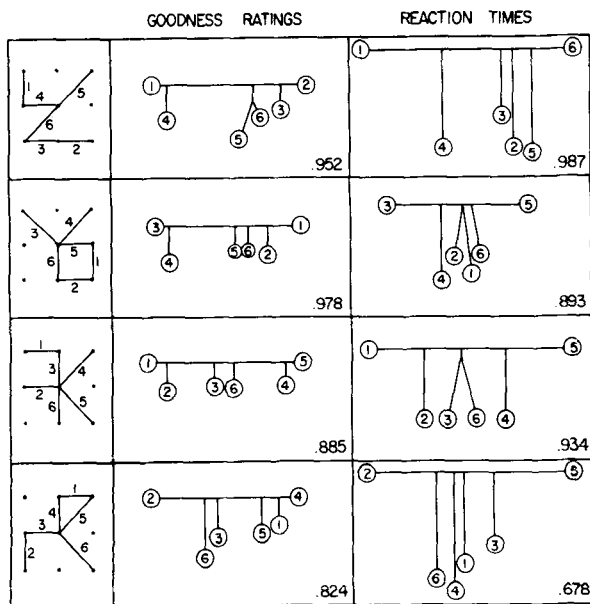

Figure 4. The four patterns used in Experiment 3 and, for each pattern, the weighted free tree representation derived from the mean goodness ratings and from the mean median reaction times, along with the correlation between the internode distances in the tree and the corresponding data values.
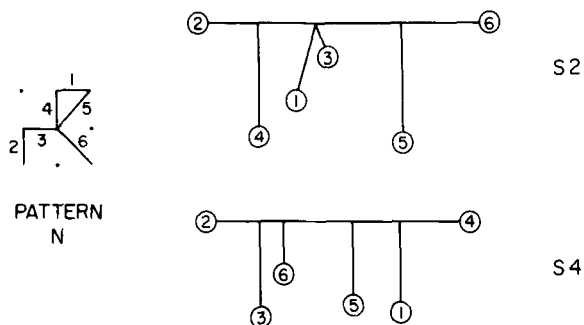
Figure 5. Pattern N and weighted free tree representations for it based upon the reaction time data of Subject S2 and Subject S4.

other way was as a "triangle" made out of Segments 1, 4, and 5, resting on a "base" made out of Segments 2, 3, and 6. Subject S2 appears to have the "steps and brace" organization, as evidenced by his tree in Figure 5 and the fact that his reproduction sequence was 2-3-4-1-5-6. On the other hand, Subject S4 appears to have the "triangle and base" organization, as shown by his tree in Figure 5 and his reproduction sequence of 5-1-4-2-3-6. The combination of these differing organizations results in a tree representation of neither, as illustrated by the reaction time tree for Pattern N shown in Figure 4. For the rest of the patterns, the group tree appears to be much more representative of the individual trees.

Reaction time trees and reproduction sequences. Each subject drew the patterns with pencil and paper after they had been learned. The order in which the line segments were reproduced can be compared with the trees based upon the reaction time data to see if they are consistent with each other. Consistency is defined in the same way as for Experiment 2. When the reaction time trees for individual subjects were compared with those subjects' reproduction sequences, the proportion of consistent segments was .812, which is significantly larger than the chance proportion of .560 ($z = 5.14$, $p < .00001$). This indicates that the reproduction task and the reaction time task may be tapping the same underlying memory structure.

Comparison of trees with other models. A natural question is whether or not there are alternative models that could do a better job of accounting for the data than trees. Accordingly, the adequacy of a spatial model and a specific network model will be compared with that of the proposed free trees. The spatial model is one in which the segments of a pattern are represented by points in a two-dimensional euclidean space. This model was fit by performing metric multidimensional scaling with the MDSCAL computer program (Kruskal, 1964). The network model is one in which the segments of a pattern are represented by nodes; any two segments that are adjacent in the pattern are connected by a link. The optimal link lengths were obtained by performing

multiple linear regression. (When there were multiple paths between nodes, the length of the path containing the fewest links was taken as the distance between the nodes.) When each model was applied to the group reaction time data for each pattern, an average of 78% of the variance in the data for each pattern was accounted for by the free tree model by estimating 9 free parameters, an average of 84% was accounted for by the spatial model with 10 parameters, and an average of 76% was accounted for by the network model with an average of 9.5 parameters (different patterns required different numbers of parameters). Thus, the spatial model fits the data slightly better than does the tree model but requires one additional parameter to do so, while the network model fits the data slightly worse than does the tree model. Since a quantitative comparison of the degree to which the three models correspond to the data does not clearly indicate which model is best, the most appropriate model would have to be chosen on other grounds. Earlier, several reasons were given for why the simplicity of trees made them preferable to general networks. A tree is also an attractive model because it has been demonstrated that a tree can represent the detailed information contained in a visual pattern (e.g., Pavlidis, 1977, Chapter 10). On the other hand, a simple spatial model of the type used here includes very little information about the exact structure of the pattern, since each line segment is represented by just two coordinate values. Of course, it is possible to propose a spatial model in which an exact copy of the pattern is stored in memory, but little explanatory power is gained by such a model. Since a purpose of this paper is to demonstrate the empirical adequacy of a tree model, the important point is that neither alternative model provided a clearly better explanation of the data.

Overall reaction time results. Up to this point, only the detailed structure of the pattern of correct "yes" reaction times has been analyzed. An examination of the overall reaction time results yields a better idea of exactly what subjects were doing in this task and why the "yes" reaction times have a detailed structure worth examining. Figure 6 plots reaction time averaged over subjects, patterns, and replications as a function of the number of segments that the test stimulus had in common with the pattern. If the test stimulus had two segments in common with the pattern, the correct response was "yes"; otherwise, the correct response was "no." Figure 6 shows that both the probability of an error and the mean reaction time for correct responses increases as the number of segments in common increases. Incorrect responses were less numerous and their times more variable than correct responses, so all that can be concluded about the pattern of their mean reaction times is that they are approximately equal to the longest mean reaction times for a correct response.

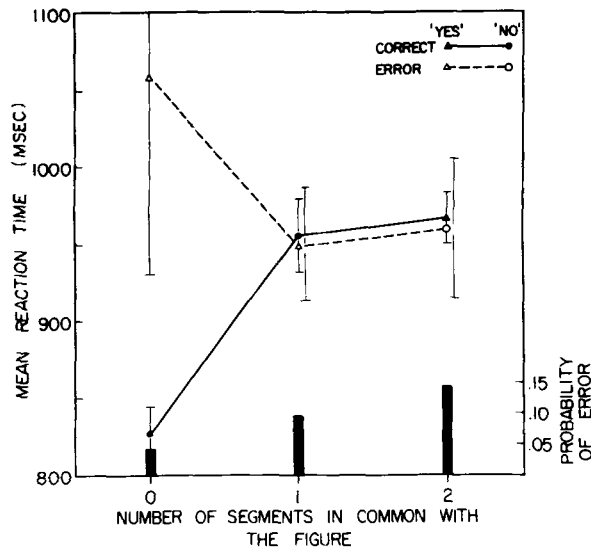The mean reaction times for correct responses are

Figure 6. Mean reaction time and proportion of incorrect responses plotted as a function of the number of common line segments shared by the memory pattern and the test stimulus. The error bars represent plus and minus one standard error.

consistent with an appealingly simple model of what subjects are doing in this task. The model proposes that subjects go through the following sequence of operations: Select one of the two segments of the test stimulus and compare it serially with each of the segments of the remembered pattern. If no matching segment is found, respond "no." When a matching segment is found, select the other segment from the test stimulus and compare it with each of the remaining segments of the pattern by searching through the memory representation starting at the node corresponding to the segment already located. If a match is found, respond "yes"; otherwise, respond "no." This model can predict the obtained ordering of mean reaction times for correct responses, and it predicts that the time required to recognize a two-segment part will be a function of the distance between the nodes corresponding to the two segments in the memory representation.

GENERAL DISCUSSION

These three experiments have demonstrated that free trees provide reasonable models of the memory representations of simple visual patterns. The results from experiments involving the drawing of patterns, goodness ratings of parts of patterns, and the recognition of parts of patterns could all be accounted for by postulating that the patterns were stored in memory in the form of weighted free trees. Trees representing the same pattern but derived from different dependent measures were very similar to each other. This similarity indicates that the behavior of a given subject in a variety of tasks could be modeled by assuming a single memory representation for any given pattern. In addition to

random error, the task requirements and the nature of the processes acting upon the memory representation might explain any minor differences between structures inferred from different tasks.

These experiments have also illustrated a type of methodology available for investigating cognitive processes in general and visual memory representations in particular. That methodology consists of (1) defining the smallest to-be-remembered units in which you are interested, (2) empirically estimating the distances between those units in the memory representation, (3) finding the tree structure that best explains those distances, and (4) trying to verify that that particular tree structure is a good model of the memory representation.

A variant of this sort of methodology has been employed by other investigators by hypothesizing that the memory structure is a continuous dimensional space rather than a tree and by using multidimensional scaling to find the best form of the representation (e.g., Henley, 1969). Previous attempts to use this approach with trees as the hypothesized memory representation (e.g., Anglin, 1970) have been limited by the fact that hierarchical clustering was the only technique available for deriving the tree structures, and thus only the use of the overly restrictive ultrametric trees was possible (Cunningham, 1978; Johnson, 1967).

Both Palmer (1977) and Reed (1974; Reed & Johnsen, 1975) performed experiments similar to Experiment 3 reported here, and yet they were unable to infer specific representations for specific patterns. Given the kind of data that they collected, the methodology available at that time was not sufficient to permit such inferences. The experiments reported here, on the other hand, always produced intersegment distance estimates of some kind, so the methodological approach discussed earlier could be used to infer the structure of specific memory representations.

While trees provided adequate descriptions of the data from these experiments, there is a lot of information in a visual pattern that is not represented in the simple trees of the type illustrated in Figure 1b. For instance, the exact location, orientation, and length of each line segment is not represented. Thus, one of the trees should not be thought of as a model of the entire representation of a pattern, but rather as a model of part of the representation of a pattern. Whether or not data from tasks that require different information or data from more complicated patterns can also be interpreted in terms of free tree representations is yet to be determined.

REFERENCE NOTES

1. Baylor, G. W. *A treatise on the mind's eye* (Tech. Rep.). Montreal: Institute of Psychology, University of Montreal, 1971.

2. Cunningham, J. P. *Similarity between trees: A statistical analysis.* Paper presented at the meeting of the Society for Mathematical Psychology, Hamilton, Ontario, August 1978.

REFERENCES

ANDERSON, J. R., & BOWER, G. H. *Human associative memory.* Washington, D.C: Winston, 1973.

ANDERSON, J. R., & PAULSON, R. Interference in memory for pictorial information. *Cognitive Psychology*, 1978, **10**, 178-202.

ANGLIN, J. M. *The growth of word meaning* (Research Monograph 63). Cambridge, Mass: M.I.T. Press, 1970.

BOORMAN, S. A., & OLIVER, D. C. Metrics on spaces of finite trees. *Journal of Mathematical Psychology*, 1973, **10**, 26-59.

COLLINS, A. M., & LOFTUS, E. F. A spreading-activation theory of semantic processing. *Psychological Review*, 1975, **82**, 407-428.

COLLINS, A. M., & QUILLIAN, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 240-248.

CUNNINGHAM, J. P. Free trees and bidirectional trees as representations of psychological distance. *Journal of Mathematical Psychology*, 1978, **17**, 165-188.

DERKS, P. L. Visual recognition of similarity and identity. *Journal of Experimental Psychology*, 1972, **95**, 237-239.

FRIENDLY, M. L. In search of the M-gram: The structure of organization in free recall. *Cognitive Psychology*, 1977, **9**, 188-249.

HENLEY, N. M. A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 176-184.

JOHNSON, S. C. Hierarchical clustering schemes. *Psychometrika*, 1967, **32**, 241-254.

KRUSKAL, J. B. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 1964, **29**, 115-129.

NORMAN, D. A., & RUMELHART, D. E. (Eds.). *Explorations in cognition.* San Francisco: Freeman, 1975.

PALMER, S. E. Visual perception and world knowledge: Notes on a model of sensory-cognitive interaction. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in cognition.* San Francisco: Freeman, 1975.

PALMER, S. E. Hierarchical structure in perceptual representation. *Cognitive Psychology*, 1977, **9**, 441-474.

PAVLIDIS, T. *Structural pattern recognition.* New York: Springer-Verlag, 1977.

REED, S. K. Structural descriptions and the limitations of visual images. *Memory & Cognition*, 1974, **2**, 329-336.

REED, S. K., & JOHNSEN, J. A. Detection of parts in patterns and images. *Memory & Cognition*, 1975, **3**, 569-575.

REITMAN, J. S. Skilled perception in Go: Deducing memory structures from inter-response times. *Cognitive Psychology*, 1976, **8**, 336-356.

WINSTON, P. H. Learning structural descriptions from examples. In P. H. Winston (Ed.), *The psychology of computer vision.* New York: McGraw-Hill, 1975.

NOTE

1. The probabilities of various numbers of moves separating a random pair of trees can be calculated by enumeration. Then the sampling distribution of the mean of four pairs can be calculated by considering all possible kinds of samples of four pairs and calculating the probability of each.

APPENDIX

The problem of determining whether or not the structures of two trees are significantly different from each other can be treated as being analogous to the problem of determining whether or not two sample means are significantly different from each other. An outline of a solution to the problem is given here, and more details are given in Cunningham (Note 2).

The null hypothesis (H_0) is that the two trees derived from different sets of data differ from the single underlying "true"

tree only because of random error in the data sets. This is analogous to the hypothesis that two sample means differ from a single population mean only because of random error (or sampling variability). The alternative hypothesis is simply the negation of H_0.

In order to test H_0, the sampling distribution of the structural difference (measured in moves, as previously defined) between two trees, given that H_0 is true, must be found. This sampling distribution will depend upon the amount of random error in the data in a way analogous to the way the sampling distribution of a sample mean depends upon the variability of the population distribution. The approximate form of the sampling distributions was determined by Monte Carlo techniques. (That is, various amounts of artificially produced random error were introduced into a large number of distance matrices from weighted free trees, and the form of the relevant sampling distributions was determined by comparing pairs of derived trees and tabulating the results.)

In order to perform the test, the amount of random error in each of the two distance matrices is first estimated. These two estimates indicate to which sampling distribution the number of moves separating the two trees derived from the matrices should be compared. The result of this comparison is the probability that a separation that large or larger could occur by chance, given that H_0 is true and that the data contain the estimated amount of error variability.

In order to be more precise, some notation must first be introduced. The objects being represented by the trees can be labeled by the integers 1 to n. Let $t(i,j)$ be the distance in the true underlying tree between the nodes representing objects i and j (normalized, for convenience, to have a mean of 2 and a standard deviation of .5) and $d_k(i,j)$ be the empirically estimated distance (similarly normalized) between objects i and j given in matrix k ($k = 1$ or 2). From each distance matrix, the optimal weighted free tree is derived, and the distance between nodes i and j in the tree derived from matrix k is $t_k(i,j)$. A multiplicative model of random error is assumed, so that the relationship between the true underlying tree distances and the empirically estimated distances is $d_k(i,j) = e_{ijk}t(i,j)$, where e_{ijk} is a multiplicative random error component with a mean of 1 and a standard deviation of σ_k. The amount of random error in distance matrix k can be estimated by obtaining an estimate of σ_k, and that estimate is

$$s_k = \left(\frac{\sum_i \sum_{j<i} \left[\frac{d_k(i,j)}{t_k(i,j)} - 1 \right]^2}{\left[\frac{n(n-1)}{2} \right]} \right)^{1/2}.$$

The notation $M(\sigma_1, \sigma_2, n)$ is used for the random variable representing the sampling distribution of the number of moves separating two weighted free trees, each with n terminal nodes, that were derived from distance matrices reflecting error multipliers with standard deviations of σ_1 and σ_2. The approximate probability distribution of $M(\sigma_1, \sigma_2, 6)$ for the tree-derivation method of Cunningham (1978) was found for various combinations of σ_1 and σ_2 by employing Monte Carlo techniques. If the number of moves separating the structures of two empirically determined weighted free trees is m_{12}, then the p value for the one-tailed test of H_0 is

$$p = P[m_{12} \geqslant M(s_1, s_2, 6)/H_0 \text{ true}].$$

This is the probability reported for tests of this kind in the body of the paper.