Simulations of human cognitive processes often employ discrimination nets to model the access of permanent memory. We consider two types of discrimination nets—EPAM and positive-property-only nets—and argue that they have insufficient psychological validity. Their deficiencies arise from negative properties, insufficient sensitivity to the discriminativeness of properties, extreme sensitivity to missing or incorrect properties, inefficiency in representing multiple knowledge domains, and seriality. We argue that these deficiencies stem from a high degree of test contingency in utilizing property information during acquisition and memory search. Discrimination nets are compared to other models that have less or no test contingency (e.g., PANDEMONIUM) and that thereby avoid the problems of discrimination nets. We propose that understanding test contingency and discovering psychologically valid ways to implement it will be central to understanding and simulating memory indexing in human cognition.

In reviewing information processing models in the Annual Review of Psychology, Simon (1979, p. 378) stated, "The safest conclusion at the present time is that human LTM can probably be represented as a node-link memory with an EPAM like index, but that various alternatives are still open for the detailed structure and organization of that memory." Subsequently, Langley and Simon (1981, p. 363) identified "recognition by means of discrimin-
nation nets" as one of several "possible invariants of human cognition."
Cognitive scientists have used EPAM-like discrimination nets for the last 25
years to model how humans access information in long term memory. After
describing EPAM nets and some of their applications, we argue that they
have a number of deficiencies as models of long term memory access. We
next consider simple variants of EPAM discrimination nets and argue that
they too lack psychological validity. We then show that simple parallel in-
dexing mechanisms such as PANDEMONIUM avoid the problems of dis-
crimination nets. Finally, we consider each model's use of test contingency
in property acquisition and memory search, and we argue that fundamental
differences in test contingency underly these models' differences in perfor-
mance.

EPAM

EPAM (Elementary Perceiver and Memorizer) was developed by Feigen-
bauk and Simon to account for "elementary human symbolic learning pro-
cesses" (e.g., Feigenbaum, 1963, p. 297; Simon & Feigenbaum, 1964). The
central component of EPAM is an extremely efficient mechanism—a dis-
crimination net—that sorts stimulus patterns to their correct actions on the
basis of stimulus properties. Consider the example of an EPAM net in Fig-
ure 1. On the left is a list of stimulus-action pairs, in which letters represent
properties of stimulus patterns and numbers represent actions. These could
be the actual materials for a paired-associates task in which subjects learn to
recall the number that goes with each letter string. We will use these pairs,
however, to represent any kind of stimulus-action unit of behavior (e.g.,
object-category pairs in pattern recognition, cue-target pairs in memory
retrieval, condition-action pairs in skilled performance).

The discrimination net on the right of Figure 1 generates the correct
action for each stimulus. The non-terminal nodes in the net are tests of stim-
ulus properties. If the stimulus EF is presented, it is first tested for having
an E. Since the outcome is positive, EF is then tested for G. Failure on this
test sorts EF to the terminal node that represents the correct action, 49. For
AB, tests for E and then C fail, leading to the correct action, 62. We should
note that this EPAM net cannot discriminate between stimuli that contain
different permutations of the same properties (e.g., EG and GE). EPAM
nets can easily make such discriminations, however, by making tests posi-
tion specific. Instead of testing a stimulus for an E and a G in any position,
a net could test it for an E in the first position and a G in the second posi-
tion. An example of such a net is shown later at the bottom of Figure 5.

An attractive feature of discrimination nets is their ability to learn dis-
criminations. Consider how the net in Figure 1 would learn the new pair,
EH-50. When first presented with EH, a test succeeds at E but not at G, and
Figure 1. An example of an EPAM discrimination net.

the net erroneously produces 49. To differentiate the two stimuli sorted to the same terminal node, the net grows a new test node based on a property of EH that does not yet exist on its sorting path in the net. In this case, H has not yet been used and is added to the bottom of the net where 49 once was. The old and new actions (49 and 50) are attached to the negative and positive branches. The net is now capable of responding correctly to both EF and EH.

A crucial component of EPAM is a "noticing order," prespecified by the program's designer, that can control (a) the order in which the properties of a stimulus are tested, and/or (b) the order in which the properties of a stimulus are added to its sorting path. One possibility for stimuli that are letter strings is a left-to-right order: When the properties of a stimulus are tested, they are tested from left to right; when a property is to be added to the sorting path of a stimulus, the left-most property not currently on the path is chosen. As we shall see, there are many possible noticing orders.

EPAM nets have played a central role in computer programs whose primary goal is to simulate human cognitive processes. EPAM nets have successfully modelled a variety of verbal learning phenomena (Feigenbaum, 1963, 1965; Hintzman, 1968; Simon & Feigenbaum, 1964). Simon and Gilmartin (1973) used an EPAM net to recognize chess patterns in a simulation of chess perception, and Goldman used an EPAM net as part of a speech production simulation in Schank's (1975) conceptual dependency system.

Although the efficiency of EPAM nets makes them attractive to artificial intelligence projects, we have some reservations about their psychological validity. We now present the reasons that underly our lack of confidence.¹

¹J. R. Anderson and G. H. Bower (1973) also provide a critique of EPAM, but it largely addresses shortcomings related to their specific interests.
Negative Properties

The first problem with EPAM nets is their heavy reliance on negative properties as seen in the net in Figure 2. The stimulus for action 67 is defined in the net as not having an A, not having a B, and not having a D. It is defined solely by negative properties, that is, by properties it does not have. In any EPAM net, half the properties are negative, and only one stimulus can be recognized on the basis of all positive properties. We find this counter-intuitive and psychologically implausible for several reasons.

First, people do not seem to have extensive knowledge of negative properties explicitly represented in memory (although they can compute them when necessary). When subjects list the properties of stimuli, they rarely, if ever, provide negative properties (see the norms of Ashcraft, 1978; Hemenway, 1981; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1975). Instead, people appear to primarily represent negative properties that deny normative expectations (e.g., the property of no vision for the concept of blindness).

Second, it seems counter-intuitive that people typically recognize things using negative information. This would be similar to walking into a room and identifying something as a chair because it does not fly and does
not have gills. People primarily appear to recognize things using positive information. Notably, the major models of stimulus structure proposed by psychologists do not employ negative properties (e.g., the geometric model of Shepard, 1962a, 1962b; the feature model of Tversky, 1977). Moreover, models like these do an acceptable job of accounting for pattern recognition data (e.g., Appleman & Mayzner, 1982; Gibson, 1965; Keren & Baggen, 1981; Krumhansl, 1982; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982).

A third problem with negative properties is that nets using them do not credit people with as much knowledge of stimulus patterns as they undoubtedly have. Consider the stimulus for action 67 in Figure 2. The net could tell us nothing positive about the properties of that object; yet it would be surprising if someone who had correctly learned its action could not recognize any of its positive properties. The use of negative properties is so computationally powerful that not all positive properties in a knowledge domain need to be represented. As the line labelled positive properties in Figure 2 shows, this net represents only 7 of the 19 positive properties in the domain. Yet we would be quite surprised if someone who had learned all the pairs in Figure 2 could recognize less than half the properties. We suspect most of the properties would have become established in memory and could be recognized as having occurred within the domain.

Attempts have been made to some extent to remedy this lack-of-knowledge problem. The nets in Feigenbaum (1965) initially sorted stimuli to stimulus images instead of to actions. These images were presumed to build up with repeated experience until they eventually contained all a pattern's positive properties. To produce an action, a stimulus image was subsequently presented to the net and sorted to its action. But this appears to be quite redundant. Why should pattern recognition include both the classification of a percept and the subsequent classification of its memory representation, which is in some sense a copy? Representing a stimulus three times—one in the net that classifies the percept, once as a stimulus image, and once in the net that classifies the stimulus image—seems unnatural and inefficient. It would be more reasonable to use a single representation for identifying a pattern and for storing knowledge about its properties.

Hintzman (1968) proposed a second solution to the lack-of-knowledge problem. In one version of his SAL model, overlearning was permitted such that all a pattern's positive properties could eventually be encoded as tests in its net. But this addition violates the spirit of EPAM nets. It undercuts their claim to fast, efficient knowledge indexing based on very little information. Clearly, if all the positive properties of a domain are eventually represented, there is no reason in the first place to represent negative properties. It would seem preferable to use some other, more efficient representation that contains only positive properties. Notably, human subjects, on occasion, do
learn more information than is necessary to discriminate stimuli (Trabasso & Bower, 1968). Consequently, some form of overlearning is necessary to account for the acquisition of redundant discriminative properties.

A fourth problem with negative properties is that nets with such properties fail to capture the distinction between generating an action to a previously perceived stimulus and generating the same action to a completely novel stimulus. The net in Figure 2, for example, would produce the same action to unfamiliar stimuli such as WXY, GJK, and FEG as to the familiar stimulus QRS, and would have no way of determining which is the familiar pattern. Yet people would probably be much better at recognizing familiar patterns than unfamiliar ones that lead to the same action.

A fifth problem with negative properties is that nets with these properties form unintuitive equivalence classes. The patterns that could be sorted to 67 in Figure 2, for example, surely form a peculiar category. This category could include familiar patterns (e.g., QRS), new patterns whose properties have occurred in the domain before (e.g., FEG), and new patterns whose properties have never occurred in the domain (e.g., WXY). Quite often these categories violate the fundamental classification principle of maximizing within-category similarity and minimizing between-category similarity. As an illustration, consider how the net in Figure 2 would categorize the new stimulus GJK. According to the classification principle just stated, GJK should be classified with AJK, but the net instead classifies it with QRS. Thus, the use of negative properties leads to predictions about transfer performance that can be unintuitive and inconsistent with psychological data.

Insufficient Sensitivity to the Discriminativeness of Properties

Focusing on the distinctive properties of a stimulus makes it possible to discriminate it from other stimuli. Focusing on the properties it shares with many other stimuli would lead to confusion and errors. We consider the following demonstration experiment (Barsalou & Bower, 1980). Subjects were asked to learn clusters of medical symptoms. Each cluster was associated with an imaginary disease. The individual symptoms (e.g., high fever, dizziness) varied in discriminativeness, being associated with either one, two, or four diseases. After subjects learned to recall each of the disease names when given its respective cluster of symptoms, they were unexpectedly tested in the reverse direction, that is, they were asked to generate the symptoms of each disease. The basic finding was that a symptom’s discriminativeness was strongly correlated with the likelihood that subjects could recall it from the disease name. If a disease had a symptom that did not occur for any other disease, subjects invariably learned that symptom first and later recalled it better than any of that disease’s other symptoms. If a disease had
no unique symptoms, subjects still learned its most discriminative symptoms. Subjects appeared to learn distinctive symptoms best because these were most useful in identifying the diseases. Discriminativeness has been shown to be central to human performance on numerous other occasions (cf. Fisher, 1981; Krumhansl, 1982; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Winograd, 1981).

EPAM nets exhibit serious deficiencies in capitalizing on the discriminativeness of information as is seen in Figure 2. The row labelled unique properties lists properties that occur in only one stimulus pattern. But the reader should note that the net represents only four of these (D, E, F, and G). Categorization, non-optimally, occurs primarily on the basis of much less discriminative properties. Many highly discriminative properties are not represented, even though the net exhibits "perfect" discriminative performance.

Part of EPAM's insensitivity to discriminativeness has to do with its noticing order. The noticing order in Feigenbaum's (1963) model dictated that properties at the beginning or end of a letter string should be chosen as discriminators before properties in the middle. In Hintzman's (1968) SAL model, properties were selected from left to right. Both procedures were apparently designed to emulate the order in which humans notice new properties. If EPAM nets are to be used as simulations of human cognitive behavior, however, their noticing order has to be flexible enough to reflect the discriminativeness of information. Quite clearly, a fixed noticing order (e.g., left-to-right) is generally not conducive to encoding into a net the properties that are most discriminable. Learners cannot know in advance which properties are going to be most discriminative, except in unusual cases when discriminativeness covaries with noticing order (e.g., discriminativeness decreases from left to right).

But what happens when a discrimination net tries to encode as much discriminative information as possible? We constructed a version of Hintzman's (1968) model to which we added a discriminativeness-oriented noticing order. Prior to each learning trial, a routine computed the discriminativeness of each property, where a property's discriminativeness was an inverse function of the number of different patterns in which it had occurred. Whenever our simulation erred in categorizing a pattern, the most discriminative property of the pattern, not yet a part of its sorting path, was always chosen to construct a new discriminating test branch. Despite this change, we consistently observed unsatisfactory asymptotic nets across many simulation runs: Highly discriminative properties were often omitted. This occurred for two reasons. First, it was impossible to accurately assess the true discriminativeness of properties early in learning since not enough patterns had been encountered. At this point, properties were randomly chosen when no one property was more discriminative than another. Low discrimi-
native properties initially acquired for a stimulus in this manner, along with negative properties acquired during the processing of subsequent stimuli, may suffice to discriminate a pattern. Consequently, it may never have its more discriminative properties encoded. More important, once the true discriminativeness of all properties is known, it is often too late to encode highly discriminative properties that were omitted. The second way highly discriminative properties were left out was that the stimulus sorted to the left-most terminal node in a net was always defined solely by negative properties. Even when this stimulus had unique properties, they were not encoded. To summarize, EPAM nets omitted highly discriminative information even when conditions were optimal for acquiring it.

Furthermore, problems existed even when highly discriminative properties were encoded in a net. First, when a disease had a unique symptom, subjects in our symptom-disease experiment rarely learned the symptoms it shared with other diseases. Yet the nets produced by our distinctiveness-oriented simulations often contained shared properties for such stimuli (as for ACG in Figure 2). Second, shared properties (e.g., A and C for ACG) in our simulations were always tested before the unique property (e.g., G), despite the fact that the unique property alone was sufficient to generate the correct action. Moreover, once a test sequence had been encoded for a stimulus, it became permanent and could not be reorganized to be more efficient. Not only did EPAM nets fail to encode highly discriminative properties, they also failed to take full advantage of those encoded. In general, then, EPAM nets are not well-suited to reflect the discriminativeness of information.

**Extreme Sensitivity to Missing or Incorrect Properties**

Another problem, originally pointed out by Neisser (1967), is that discrimination nets are too sensitive to missing or incorrect information, as in Figure 2. Normally, pattern ACG gets sorted to response 77. But if A is missing or is misidentified as an H, for example, the pattern fails all tests and will be sorted incorrectly to 67.

The reader should note that when A is not detected or is misidentified, ACG is confused with a stimulus with which it has nothing in common (QRS) instead of with a more similar pattern (ACH). Such predictions are inconsistent with the confusion matrices derived from numerous experiments on pattern recognition: When information is missing or misidentified, a pattern is most likely confused with patterns that share its remaining properties (e.g., Conrad, 1964; Gibson, 1965; Shepard, 1962b). But as we have just seen, EPAM's confusions often do not reflect similarity among stimulus patterns. In general, the earlier in a discrimination net an error occurs,
the more likely it is that the resulting confusion will violate this similarity principle.

Yet EPAM nets have been credited with successfully accounting for similarity (e.g., Simon & Feigenbaum, 1964). For example, when an incorrect response is given to a stimulus during learning, the confusion is often between very similar stimuli (as in the example of learning presented for Figure 1). These similarity based confusions, however, primarily occur when all the properties of a stimulus are correctly perceived. In many cases in which a single property is misidentified or not encoded—certainly a common perceptual phenomenon—search is misdirected away from the terminal node that represents the action for the most similar stimulus. Under such conditions, EPAM nets often do not account correctly for similarity.

**Inefficiency in Multiple Knowledge Domains**

The next two problems arise when an EPAM net represents multiple knowledge domains. In general, the more knowledge domains in which an EPAM net discriminates, the more inefficient and psychologically implausible it becomes.

*Extended Negative Properties.* Imagine subjects learning stimulus-action pairs from several different knowledge domains, where no two domains share any properties. The resulting EPAM net would look something like Figure 3, where each subnet (i.e., A, B, C, and D) represents a knowledge domain. Each right-branching path in this net represents a positive property, and each left-branching path represents a negative property.

The problem to be discussed arises from the long path of negative properties down the left edge of the net. To make discriminations in the D domain, it is necessary to go down a very long path of negative properties to get to the relevant tests. Not only is this inefficient, it results in several counter-intuitive predictions. First, items in the domain learned first, A, should always be discriminated faster than items in the most recently learned domain, D. This runs counter to the common observation that more recently studied material is retrieved faster than material not practiced in a while. The second problem stems from an assumption, not made by EPAM theorists, but one that certainly seems reasonable; namely, each application of a test node strengthens it in memory such that it is processed more quickly on future applications. It follows from this assumption that when stimuli are discriminated in domain D, nodes for tests along the left-hand path through A, B, and C should be strengthened. But this amounts to claiming that in the process of recognizing furniture, the nodes for has feathers and swims could become stronger, assuming that domain D is furniture, and that do-
mains B and C are *birds* and *fish*. It seems psychologically implausible that properties a stimulus does not have could be strengthened by recognizing it.

Figure 3. An EPAM net that exhibits the problem of extended negative properties in multiple stimulus domains.
**Exponential Node Growth.** This problem occurs when different knowledge domains share properties. Consider Figure 4, in which each of four domains contains a common structure represented by the BCD subnet (the domains are individually defined by the properties W, X, Y, and Z). The reader will note that the overall net has multiple nodes for properties of the subset; 4 for B, 8 for C, and 16 for D. In general, if the same subnet exists for each of several domains, and if an action exists for every terminal node, then the number of nodes for a shared test is $d^2p-1$, where $d$ is the number of domains and $p$ refers to the property's position of acquisition (i.e., whether it was acquired first, second, third, etc.). For example, if four domains share a fifth test, the number of nodes needed to represent it in all patterns is $4(2^5-1)=64$. As can be seen, the number of nodes required to represent each new test increases exponentially. Obviously this is not an efficient use of storage space. More important, it does not make much psychological sense to represent properties a different number of times in an arbitrary fashion.2

**Seriality**

As Feigenbaum (1963, p. 306) states, “EPAM is a serial machine.” Yet it is now widely believed that human pattern recognition is a parallel process (e.g., McClelland & Rumelhart, 1981; Neisser, 1967; Shiffrin & Gardner, 1972). Furthermore, there is a plausible neurophysiological basis for parallel processing in pattern recognition; namely, feature detectors in sensory systems have parallel projections to the brain. The properties of a stimulus are not all that can be accessed in parallel. Representations of different stimuli can be accessed in parallel as well. Extensive practice at accessing multiple stimulus representations in permanent memory eliminates the slow-down characteristic of serial processing (e.g., Neisser, Novick, & Lazaar, 1963; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). In general, numerous results suggest that access to permanent memory during well-practiced activities such as reading, object recognition, and other more specialized skills is a parallel process. Besides arguing for parallel retrieval of well-learned information, many theorists additionally argue for parallel retrieval of information processed on only a single occasion; namely, retrieval of episodic memories (e.g., J. A. Anderson, Silverstein, Ritz, & Jones, 1977; J. R. Anderson, 1976; Ratcliff, 1978).

1Feigenbaum (1965) tried to avoid this problem by having only one type node in memory for any test. When a test had to be represented more than once, the multiple tests were considered tokens of the same type. However, this does not handle the problems of using storage space inefficiently and of creating different numbers of test tokens indiscriminately. It seems more parsimonious and psychologically valid to have only one test per property unless there is good reason not to (e.g., for a property common to different senses of a word).
Figure 4. An EPAM net that exhibits the problem of exponential node growth in multiple stimulus domains.

The Paradox of the Expert. As Smith, Adams, and Schorr (1978) note, modern memory models often predict that increasing amounts of information in a memory should increasingly interfere with accessing any one
piece. This runs counter to the everyday observation that experts—people with extensive knowledge of a domain—access this knowledge very efficiently. EPAM nets, which have been used to simulate the knowledge of skilled chess players (e.g., Simon & Gilmartin, 1973), also have this deficiency. As an EPAM net's "expertise" increases, so does the number of patterns it learns to discriminate. This is invariably accompanied by an increasing number of serially performed tests, thereby increasing the average time required to categorize a pattern. Consequently, serial models like EPAM are not well-suited to account for the efficient memories of experts. In contrast, theories of skilled performance often use production systems to capture the automatized and parallel character of the expert's performance (e.g., J. R. Anderson, 1982).

TEST CONTINGENCY IN EPAM

We believe that a high degree of test contingency is the fundamental problem with EPAM nets, and that all the problems we have discussed so far stem from this more basic problem. We next discuss how test contingency dominates two fundamental processes in EPAM: property acquisition and memory search. In acquisition, the probability of encoding a test for a pattern is highly contingent on which properties have been previously encoded for other patterns. In search, the probability of a particular property being tested for a pattern is highly contingent on which of its other properties were previously tested. As we shall see when we discuss PANDEMONIUM, some alternative models possess no test contingency at all in acquisition and search.

One form that test contingency takes in EPAM's acquisition of properties is in the extensive use of negative properties. As we saw in Figure 1, when two stimuli were erroneously sorted to the same node during learning (EF and EH), a new test node (H) was grown to discriminate them. For the stimulus having this new test property (EH), the test was for a positive property, but for the stimulus not having the property (EF), the test was for a negative property. Notably, the stimulus not having the property was defined relative to the stimulus having it—one of its tests was contingent on the stimulus with which it was confused. If EF had been confused with a different stimulus, it could have easily been defined as not having a different property. In general, a given stimulus can be defined in an indefinitely large number of negative ways depending on with which of an indefinitely large number of stimuli it is confused. Taking advantage of this powerful, but highly contingent, form of representation results in problems of psychological validity: (a) EPAM's resulting representation is credited with having extensive, explicit knowledge of negative properties; (b) negative knowledge takes part in pattern recognition and is strengthened during the process; (c)
the representation of positive properties is slighted; and (d) unusual equivalence classes are formed.

Test contingency also underlies EPAM's insensitivity to discriminativeness during property acquisition. As we have seen, negative properties are a highly contingent form of representation. Notably, representing stimuli in this way can result in the omission of their unique properties. This occurs when negative properties, along with low discriminative properties acquired early in learning, suffice to discriminate a pattern having unique properties. By employing the discriminative power of negative properties, EPAM often avoids having to ever encode the unique properties of a stimulus. This is clearly inappropriate, however, with respect to psychological validity.

Moreover, test contingency causes problems even when unique properties are encoded. A stimulus with a unique property may often have its shared properties unnecessarily represented on its sorting path in a net. This occurs because shared properties can be encoded into a pattern's sorting path during previous learning trials involving other stimuli that share these properties. This highly contingent acquisition process also constructs permanent test sequences in which unique properties are tested after shared properties. In general, EPAM does not optimally represent discriminativeness because the properties encoded for a stimulus are contingent on what was learned previously for other stimuli.

Test contingency underlies EPAM's construction of extended negative properties during property acquisition. This problem occurs because the representation of a new domain is contingent on the representation of old ones: When a new domain does not share properties with old domains, this contingency is unnecessarily and inefficiently represented in EPAM nets as a long path of negative properties.

So far we have been discussing the role of test contingency in EPAM's problems with property acquisition. Test contingency also underlies EPAM's problems with memory search. The basic problem is that a property can only be tested if it has been preceded by a particular sequence of test outcomes. Later parts of a correct test sequence are contingent on earlier tests being performed correctly. This characteristic of search underlies EPAM's extreme sensitivity to missing and incorrect properties. If a property is absent or misidentified, the correct sequence of tests will never be performed, as later tests in the sequence are contingent on all previous tests being performed correctly. Test contingency in search produces disastrous "non-robust" results when EPAM tries to identify noisy patterns.

The impact of test contingency in search is also seen in exponential node growth. As stated previously, a property can only be tested following a particular sequence of previous test outcomes. Notably, if many different paths require a test for the same property, each path must have its own test,
since using a common test would invalidate any differences in possible test sequences up to that point. For example, prior to a test for P, tests for A and B could have been met, or tests for X and Y could have been met. If both paths converged on the same test for P, the results of these previous tests would be lost. By having a separate test for P on each path, information from prior tests is preserved. But as more paths require a test for the same property—something that increases as multiple domains share more properties—the number of property tokens can increase exponentially.

Seriality in discrimination nets is another outcome of test contingency in search. Since the tests performed at one level in a net are contingent on the outcomes of tests at previous levels, tests at deepening levels must proceed serially. The inability to account for the paradox of the expert stems directly from this form of serial processing: As EPAM develops increasing knowledge of a domain, it must perform a greater number of contingent tests to perform a given discrimination. Only representations capable of testing all the properties of a stimulus simultaneously appear capable of modelling the fast and efficient performance of experts.

DISCRIMINATION NETS WITHOUT NEGATIVE PROPERTIES

Given the problems associated with negative properties, discrimination nets without them might fare better as simulations of human cognition. We will refer to such nets as positive-property-only (PPO) nets. De Jong's (1979a,b) FRUMP program is an example of how PPO nets can be employed in simulations of cognitive processes. FRUMP sorted newswire stories through a discrimination net to scripts relevant to understanding the stories. Only tests for positive properties (e.g., the actions, objects, and agents of a story) occurred in this net. We next explore how PPO nets like those in FRUMP handle the problems we raised for EPAM.

As shown at the top of Figure 5, PPO nets differ ostensibly from EPAM nets in that (a) they do not have tests for negative properties, and (b) multiple tests can emanate from a given superordinate test (e.g., from START and G). Several interpretations can be made of PPO nets. One is to implement the decision at each n-way branching node as a series of n-1 "yes-no" decisions or what we will refer to as "local serial search." Under this interpretation, the two- and three-way decisions in the PPO net at the top of Figure 5 are realized as corresponding subsets of serial "yes-no" decisions in the EPAM net at the bottom on Figure 5. If the PPO net is searched in this serial manner, then a stimulus is first tested for A. If that fails, control is returned to START and then sent to B which is the equivalent of taking the negative path from A in the EPAM net. As can be seen by further comparing the two nets, other paths in the PPO net can be decom-
posed similarly into a positive path plus a negative path. Tests are performed one at a time, with failures equivalent to negative properties. Negative properties are therefore implicit in PPO nets that use "local serial search" at each n-way branching node.

The other ostensible difference between EPAM and PPO nets—multiple tests emanating from a previous superordinate—also vanishes when PPO nets use local serial search. A stimulus either passes a test in a PPO net and traverses deeper into the net, or the stimulus fails the test, takes the circuitous route back to its superordinate, and proceeds down to the next test. There are still only two paths at each test.

Since PPO nets that use local serial search are extremely similar to EPAM nets, they suffer from many of EPAM's difficulties. Consequently, we will not further consider this interpretation of PPO nets.

Another interpretation of PPO nets is that the n tests emanating from each node are processed in parallel or what we will refer to as "local parallel search." After a stimulus is sorted to G at the top of Figure 5, for example,
it is simultaneously tested for C, E, and R. All tests emanating from the previously successful test are performed in parallel, perhaps by hash coding the properties of the stimulus being tested. Consequently, the decision at each node is arrived at directly and is not implemented by binary "yes-no" decisions. Under this interpretation, PPO nets have somewhat different properties than EPAM nets. In the following sections, we see how PPO nets that use local parallel search fare when confronted with the difficulties raised for EPAM.

**Negative Properties.** Since PPO nets primarily grow through the addition of positive properties, they avoid the problems associated with negative properties.

**Insufficient Sensitivity to the Discriminitiveness of Properties.** Since PPO nets store more positive properties than do EPAM nets, they are more likely to include highly discriminative properties. But to the extent a noticing order is not oriented towards discriminativeness (e.g., it is left-to-right), abstraction of discriminative properties will not be optimal. Optimal abstraction will only occur when a memory that contains the discriminativeness of properties is kept, and when this information is used to guide the encoding of properties.

Even under optimal conditions for abstracting discriminative information, PPO nets have problems. Like EPAM nets, PPO nets often begin constructing the representation of a stimulus prior to processing it. Consequently, a pattern with a unique property may unnecessarily have properties on its sorting path that it shares with other, previously processed stimuli (e.g., G for GRY in Figure 5). Also, like EPAM nets, PPO nets always inefficiently test a pattern's shared properties before its unique ones when both kinds are represented. This problem is aggravated by the fact that once a test sequence is encoded for a pattern, it is permanent and cannot be made more efficient.

PPO nets can be reorganized in various ways to make them more efficient. For example, a separate list of all the condition-action pairs presented during learning could be maintained and intermittantly used to completely reorganize the net more efficiently at various points in its development. But such proposals, however efficient they may eventually be, seem psychologically implausible.\(^3\)

\(^3\)We should note that under certain conditions, the form a PPO net takes is independent of its training history. For example, when a PPO net (a) attempts to learn a fixed stimulus domain, (b) uses a left-to-right noticing order in both test and acquisition, and (c) develops a net in which each level corresponds to a particular letter position, the resulting net will take the same form, regardless of the order in which stimuli are processed. A PPO net takes on different forms primarily when it uses flexible noticing orders during property acquisition (e.g., a noticing order that is oriented towards discriminativeness).
**Extreme Sensitivity to Missing or Incorrect Properties.** PPO nets do not fare better on this problem than EPAM nets. If a property is missing, it may not be possible to search any further in the net (e.g., if G were lacking from GCK in Figure 5). Some simulations using PPO nets have ways of inferring missing properties, thereby making it possible to overcome this problem on some occasions (cf. Kolodner, 1983a, 1983b, in press). But, if special heuristics are unable to infer missing information, search will still fail. Nonoptimally, even when the remaining properties are sufficient for a correct discrimination, they can not converge on the correct action once search has been misdirected.

Misperceived properties and incorrectly inferred properties are more serious. On the basis of a single incorrect property, correct categorization is impossible. This can even occur when numerous unique properties have been encoded that are clearly sufficient to discriminate the stimulus from all others.

One way to deal with these problems is to implement redundant PPO nets. The important property of these PPO nets is that a given stimulus can be sorted down more than one path to the same terminal node. Each of these multiple paths contains a *different* sequence of tests sufficient to discriminate a stimulus. Besides varying in the order in which they test properties, redundant test sequences can also vary in the properties they test. Consequently, when search down one path fails because of a missing or misidentified property, search down another path can still succeed if its sequence does not test the missing or incorrect property. Kolodner's model of episodic memory, CYRUS, is an example of how redundant PPO nets can be used as simulations of cognitive processes (Kolodner, 1980, 1983a, 1983b, in press; Schank & Kolodner, 1979).

**Inefficiency in Multiple Knowledge Domains.** Since PPO nets contain only positive information, they do not have the problem of extended negative properties. When searching for a stimulus in one knowledge domain, they do not have to first search through irrelevant knowledge domains. Like EPAM nets, however, PPO nets exhibit the problem of exponential node growth (which is beginning in Figure 5 with C and E). As increasing numbers of properties are shared by subnets within a larger net, the number of nodes needed to represent each additional shared property can increase exponentially. This problem is even more severe for redundant PPO nets, since multiple sorting paths for a stimulus typically require multiple tests for at least some properties. In general, we find it psychologically implausible that a test can be represented an arbitrary number of times in memory and that a pattern can have an arbitrary number of sorting paths.

**Seriality.** Although tests directly emanating from a previous test node are locally processed in parallel, search down through a PPO net from level to level is serial. This is true of both nonredundant and redundant PPO
nets. Even when PPO nets exhibit local parallel search, they nevertheless exhibit global serial search. If several properties of a pattern must be considered during recognition, they are processed serially, so that processing time will increase with the number of tests performed. This prediction clearly runs counter to much of the work in pattern recognition, which generally assumes and has found evidence for parallel processing of a stimulus pattern's properties (e.g., parallel processing of letters in word recognition; McClelland & Rumelhart, 1981).

Since processing in PPO nets is serial at the global level, they mistakenly predict longer access times as expertise in a domain increases. As a net discriminates more stimuli, it generally becomes deeper, and accessing a given response necessarily takes longer. PPO nets do not seem to have any way of overcoming the slow-down due to increasing knowledge.

**Test Contingency in PPO Nets**

As with EPAM nets, the problems of PPO nets can similarly be traced to test contingency. In acquisition, the tests encoded for a stimulus are contingent on which stimuli were previously processed during learning. Because the shared properties of a stimulus are often acquired during the processing of previous stimuli, highly discriminative properties may not be encoded for a stimulus, or they may be encoded last.

Test contingency also underlies the problems of PPO nets during search. First, test contingency causes PPO nets to be overly sensitive to missing and incorrect properties. As with EPAM nets, later tests in a test sequence are contingent on previous tests being performed correctly. Consequently, search can only proceed properly for a stimulus if all its properties are correctly perceived. Once search is misdirected by a missing or misidentified property, it is impossible to perform the proper sequence of tests leading to the correct action (except for redundant PPO nets).

Second, test contingency in search is responsible for the problem of exponential node growth. Since a test for a property may be contingent on several possible sequences of previous tests, it must have a distinct representation in each sequence. Consequently, many representations of a property may be necessary. This problem is especially acute in redundant PPO nets.

Third, test contingency in search is responsible for global serial search in PPO nets. Tests at deepening levels must proceed serially, since the tests performed at one level are contingent on the outcomes of tests at previous levels.

**PANDEMONIUM**

Lest the reader conclude that the problems confronting discrimination nets are general difficulties for any model of memory indexing, we will briefly
consider how PANDEMONIUM handles these problems (Selfridge, 1959; also see Neisser, 1967). We use this model to illustrate that the problems confronting discrimination nets are solvable, not to argue that PANDEMONIUM is without problems of its own.

The PANDEMONIUM net shown in Figure 6 is sufficient to discriminate the list of stimulus-action pairs given at the top. The lettered nodes represent properties, the unlabelled nodes represent stimuli possessing these properties, and the numbered nodes represent actions. When a stimulus is presented, all its properties become simultaneously active in memory, and activation spreads along all links emanating from them. Each stimulus node then summates the activation arriving from its properties. For simplicity, we assume that activation arriving at a stimulus node does not travel back down to any of its properties. The action receiving the most activation from its stimulus node is finally selected as the classification of the presented stimulus.

\[
\begin{align*}
ABC & - 42 \\
BDE & - 96 \\
CDE & - 67 \\
BCE & - 73
\end{align*}
\]

![Figure 6. An example of a simple PANDEMONIUM net.](image)

The absence of test contingency in PANDEMONIUM nets is striking in comparison to the extensive test contingency in discrimination nets. When a PANDEMONIUM net acquires a new pattern, links are established from its properties to its stimulus node independently of the other property-
DISCRIMINATION NETS

It is of interest to consider a more recent and sophisticated PANDEMONIUM net—McClelland and Rumelhart’s (1981) interactive activation model of word recognition—which is intermediate in degree of test contingency between simple PANDEMONIUM nets and discrimination nets. Their model has three levels of tests; one each for features, letters, and words. At the feature level, all features of all letters can initially be tested simultaneously—there is no test contingency in search. But, as some features accrue evidence for their presence in a display, they begin to inhibit incompatible features. As evidence accumulates, test contingency increases to the point of completely suppressing unlikely features. Such interactive test contingency similarly occurs to some extent within and between the other levels. As visual features provide evidence for the R in RAT, for example, the RAT word-unit facilitates the R-letter unit and inhibits competing letter units for the first letter position (e.g., L and W). Notably, test contingency in this model is very different from that in discrimination nets.

Pandemonium Versus the Issues

Returning to the simple PANDEMONIUM model, let us see how it deals with the difficulties confronting discrimination nets.

**Negative Properties.** Since stimulus nodes primarily become associated to the positive properties of patterns over the course of learning, PANDEMONIUM avoids criticisms directed at the use of negative properties.

**Insufficient Sensitivity to the Discrimination of Properties.** Although the simple PANDEMONIUM net in Figure 6 is not sensitive to the discriminativeness of information, it can be made so with the addition of two standard assumptions typically found in models of human memory. First, by allowing property-to-stimulus links to vary in strength, highly discriminative properties can develop stronger links to their respective stimuli than less discriminative properties. Moreover, since stronger property-to-stimulus links contribute more activation to their responses, highly discriminative properties have more control over memory indexing. One factor that could

---

4 The PANDEMONIUM net in Figure 6 uses letter nodes that are not position specific. Consequently, it will confuse different permutations of the same letters (e.g., AB and BA). Such confusions can be avoided by having letter-nodes specific to each position.
cause strength to vary is that people may rehearse properties to the extent they are discriminative, thereby strengthening the links from highly discriminative properties more than those from less discriminative properties.

A second assumption, associative interference, also enables sensitivity to discriminativeness. One form of this assumption states that the amount of activation along a property-to-stimulus link decreases to the extent more stimuli are associated to the property (i.e., the fan effect; J. R. Anderson & Bower, 1973). In Figure 6, for example, less activation would flow along the property-to-stimulus links from B than from A, since more stimuli are associated to B than to A. To the extent many links are associated to a property (i.e., it is not discriminative), it will only play a weak role in activating a response. Unique properties suffer no fan effect, contribute more activation to their responses, and, therefore, have more control over memory indexing.

Since sensitivity to discriminativeness must take an entire stimulus domain into account, these added assumptions introduce some test contingency into PANDEMONIUM nets. These forms of test contingency, however, are much more flexible and adaptive than the form found in discrimination nets. If new stimuli are added to a domain such that the discriminativeness of properties changes, PANDEMONIUM can easily accommodate this by changes in strength and interference. Notably, discrimination nets cannot accommodate themselves in this manner, given their problems in representing strength and given their difficulty in rearranging properties in a net.

**Sensitivity to Missing or Incorrect Properties.** Since simple PANDEMONIUM nets have no test contingency in search, they behave much more robustly than discrimination nets when properties are missing or misidentified. If a property is absent or incorrect, the remaining properties can still converge on the correct response if they are sufficiently discriminative. More important, missing or incorrect properties do not prevent other discriminating properties from entering into the decision process, as non-optimally occurs for discrimination nets. Consequently, PANDEMONIUM gives graded responses to "noisy" patterns depending on their similarity to stored patterns.

**Inefficiency in Multiple Stimulus Domains.** Since PANDEMONIUM nets do not use negative properties, they avoid the problem of extended negative properties. Moreover, since PANDEMONIUM nets use the same property node repeatedly for different patterns, they avoid the problem of exponential node growth.

**Seriality.** Since PANDEMONIUM nets process a pattern's properties in parallel, they are capable of accounting for the parallel processing that is characteristic of human memory indexing and expert performance.
PSYCHOLOGICALLY VALID USES OF DISCRIMINATION NETS

Although discrimination nets may not be psychologically valid accounts of memory indexing in humans, they may, on occasion, accurately represent high level decision strategies. In some cases, these strategies appear to contain negative properties and to exhibit serial processing. For example, negative properties often appear central to high level decision strategies people use when buying a new car (e.g., has/does not have power steering, gets/does not get good mileage, seats/does not seat more than two people). Moreover, people often appear to sort models of cars serially through properties like these to terminal nodes that represent subjective preferences or general car types. Another example is psychiatric diagnosis. The Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1980) sets forth decision trees that are basically EPAM nets. These nets contain psychiatric symptoms as test nodes (e.g., panic attacks, bizarre behavior, hallucinations) and psychiatric diagnoses as terminal nodes (e.g., brief reaction psychosis, infantile autism, schizoaffective behavior). Physicians may similarly employ discrimination nets in diagnosing physical disorders. In general, it is easy to imagine people literally memorizing discrimination nets for use in domains like these. Nevertheless, PANDEMONIUM nets can also account for performance on such tasks. In fact, the data from the symptom-disease experiment we reported earlier are more consistent with PANDEMONIUM nets than with discrimination nets (Barsalou & Bower, 1980).

CONCLUSION

Test contingency in discrimination nets causes each stimulus in a domain to be represented and processed relative to all others. As a consequence, stimuli do not have unique representations, but instead share much of their representation with each other. In contrast, each stimulus in a PANDEMONIUM net has a unique representation; namely, a set of independent property-to-stimulus links. Such links are not constructed during the processing of previous stimuli and are processed independently of other links during search (although assumptions of strength and interference introduce some degree of test contingency).

The implementation of test contingency implicitly assumes that the context of a stimulus will remain fixed. Representing a stimulus relative to others assumes that it will always be discriminated from the same set of stimuli: Representations of new stimuli will not be added to memory, and representations of familiar stimuli will not be updated. But if such changes occur, as they must in realistic settings, a net may be crippled to the extent it
has not encoded independent representations for each stimulus. By making the representation of one stimulus dependent on the representations of others, the ability to recognize a stimulus will be threatened by changes in the stimulus set. When the stimuli comprising a domain can be expected to fluctuate, models that represent stimuli independently appear preferable to models that represent stimuli contingently. Under such conditions, stimulus representations should be sufficiently robust and context-independent to be useful as the stimulus set changes.

Test contingency may only be practical in highly constrained contexts such as letter and word recognition for which exhaustive knowledge of the stimulus domain exists (cf. McClelland & Rumelhart, 1981). Otherwise, when new stimuli are encountered, the pattern of contingency may change so that the old representation of contingency becomes counter-productive. Under such circumstances, some representations seem to recover more gracefully than others. As discussed earlier, for example, PANDEMONIUM nets respond to changes in discriminativeness more easily than do discrimination nets.

Much remains to be learned about the role of test contingency in human cognition. In what stimulus domains does it operate? Does it take different forms in different domains? Although test contingency may play an important role in cognitive processes, modelling it with discrimination nets is inappropriate for the reasons we have raised. Other forms, such as the fan effect and the use of inhibition by McClelland & Rumelhart (1981), appear more appropriate. Test contingency has recently become a central issue in the literature that addresses how people acquire representations of categories (Medin & Schaffer, 1978; Medin & Schwanenflugel, 1981).

Because discrimination nets are an efficient way of indexing the memories of modern serial computers, they are attractive for use in artificial intelligence. We have no argument with the use of efficient procedures in computer science applications, but we do question their use as simulations of human behavior. Although discrimination nets may play an occasional role in high level decision strategies, the confidence some cognitive scientists have in them appears unwarranted. We certainly do not believe it safe to conclude that a component of human long term memory is a discrimination net indexing mechanism. Nor do discrimination nets appear to be an invariant property of human cognition. We hope our analysis will lead to new forms of memory indexing that are more psychologically valid. In particular, we believe that understanding the role of test contingency and discovering psychologically valid ways to implement it will be central to understanding and simulating human cognitive processes.

REFERENCES


Journal of Experimental Psychology: Human Learning and Memory, 7, 355-368.
Schank, R., & Kolodner, J. L. (1979). Retrieving information from an episodic memory—or—Why computers' memories should be more like people's (Research Report #159). New Haven: Yale University, Department of Computer Science.