THE ANALYSIS, SNYTHESIS, AND DESCRIPTION OF BIOLOGICAL IMAGES

Lewis E. Lipkin, William C. Watt and Russell A. Kirsch

THE ANALYSIS, SYNTHESIS, AND DESCRIPTION OF BIOLOGICAL IMAGES

Lewis E. Lipkin
National Institutes of Health
Bethesda, Md.

William C. Watt and Russell A. Kirsch
National Bureau of Standards
Washington, D. C.

Introduction

We will begin by asking if the prevailing notion of what computer systems can achieve in biological image-processing may not be overly modest. This question seems to us fundamental, for such an underestimate might produce not only a limitation of objectives, but also a restriction of the class of problems considered amenable to computer processing. Indeed, the very concept of mechanized processing of biological images might have been needlessly narrowed.

We believe this narrowing in fact to be needless, for we cannot accept what we hold to be the basic presupposition which underlies it. That is, we do not believe that the structure of the general-purpose digital computer confines us to the sort of methodology which is characterized by: (1) immediate reduction of information; (2) information reduction dependent on peculiarities of the information sample; and (3) treatment of information in purely numerical terms.

No matter what his object of study, the biologist has become resigned to construing the tasks he sets to the computer in such a way as to suppress the complexity of biological images in favor of such amenable attributes as optical density, size, and number of disjoint objects in the field. Now, he often does have some interest in properties which are best expressed in terms of measurements yielding numerical values; and it is certainly true that in the present state of the computer arts such measurements are widely considered least intractable to computer implementation. Some impressive examples of such implementation are currently, in fact, much in our consciousness; it is far from our purpose to disparage them. We only remind the biologist of all the structural properties he has gotten used to the computer's ignoring, but to which he himself habitually devotes a great deal of attention. And we ask if he might not, in analyzing these properties no less than in studying "quantizable" ones, welcome the computer as an ally.

(We do not claim that all structural properties are not quantizable, for we grant that the characterization of structure may be accomplished statistically, in the summation of one or more, more or less complex

984

numerical properties. In opposing "quantizable" [in quotes] to structural properties, we are making a rather different point, and one which we will have cause to return to below: that structural properties are generally *not* expressed numerically among biologists, because of the unwieldiness of such expressions. It is consistent with our general outlook that we believe it therefore incumbent on us to avoid such unwieldiness if we can.)

Our work has as its end, analysis; as its means, analysis bolstered by synthesis and description. At the heart of our research is the view that, in processing the architecture of complex images, any information-reduction must be preceded by an information-*amplification*: and this, as we will show, in two different though related senses.

By "analysis" we mean, put most simply, what any biologist means, except that we plan for this function ultimately to be performed by a computer system. Our usage of "synthesis" is, again, the obvious one. A computer program which can synthesize biological images, can provide pictorial instances which match (nonpictorial) specifications, whether these specifications determine a naturally-occurring image or an image contrived by the experimenter. Our use of "description" is also like the everyday use: a "description" of a biological image is just such a description, in English sentences, as one biologist might offer another or might offer a computer.

The information derived from the process of iteratively improving the syntheses, through relevant descriptions, provides the amplification required for the class of sophisticated analyses which forms our objective. The kinds of amplification required, and the potential means for their utilization, form the main subjects of this paper.

The Analysis of Biological Images

Scientific biology rests upon the reproducibility of its results. We must be careful however, not to confuse reproducibility with quantifiability. One can be fully as precise in explicating pattern and structure as in performing measurements which yield numerical quantities. Moreover, the explication of these articular properties, as we term them, may be made without recourse to such measurements. The distinction between quantitative and qualitative characteristics of biological images is closely related to the distinction which we are about to examine: that between metric properties on the one hand and articular properties on the other. This examination may best be introduced by an example.

If one were interested in investigating the relationships between nucleic acid content and nervous system structure and function, one might take a number of different approaches. We classify these approaches into four levels. The first of these levels might be termed *diffuse;* it is typified by chemical analysis of an entire brain or a large portion thereof. The result is a set of numerical values with a high degree of possible precision.

From the articular viewpoint, the only discrimination involved would be the identification of the object of analysis as being a brain, or brain-portion.

The second level we call *disarticulate*. In terms of our example, if one wanted to enumerate DNA-containing bodies, one might homogenize a portion of the central nervous system which had been disrupted by physical or chemical means, and then count nuclei. The resulting numerical data would be fairly precise; and insofar as the enumerated bodies represented local concentrations of nucleic acid, the results would perhaps be more indicative of the role played by DNA in the central nervous system. From the articular aspect of our dichotomy, one would have relied on a physicochemical process to disarticulate the tissue into component parts, in such a way as to isolate component objects, i.e., the nuclei. This form of chemical discrimination allows the counting to proceed in a meaningful way. Obviously, however, a large body of structural information which would doubtless be of great help in further explicating the role of DNA in the central nervous system, is irretrievably jettisoned in the process of disarticulation.

The third level of analysis, which we term *metric*, may be typified by numerical work on histologic sections. At this level every effort is made to preserve a close mapping of the object to be directly studied upon the presumed structure of the untreated tissue. To continue our example, nuclei might be counted or sized *in situ*. From the structurally-preserved image, however, only numerical values, or sets of numerical values, are obtained. From the articular viewpoint, some of the pattern discriminations necessary for the analysis are accomplished by staining; other parts of the discrimination process are accomplished through individual selection by the examining biologist. At this metric level of histologic analysis increasing numbers of objects of different types are potentially made available for measurement. And the measurements become more meaningful as they are presented in terms of articular relationships existing among the measured structures.

The fourth level, which represents our main interest, we term *articular*. It is characterized by the reproducible investigation of structural properties—by seeking answers to questions about positional relations among objects rather than to questions of "how many" or "how much." At this level of analysis any measurements yielding numerical values are subordinated to the results which present instances of such structural relations. The (primarily nonquantitative) results of an articular analysis characterize the arrangement of objects within an image and the arrangement of subobjects within them, etc., so as to present organized information about the structure of the whole, in terms of its proper components. In our DNA-measuring example, we might determine DNA content from opacity measurements on only those parts of a cell which are within the

nucleus but outside of the nucleolus and concentrated in a subnuclear membranous location or in a fine reticular network.

Analysis at the diffuse level does not deal with images and therefore does not concern us. At the disarticulated level, existing technology (e.g., the Coulter Counter[1]) is sufficient. At the third level, that of metric analysis, there is currently a great deal of activity; and in view of this we are perhaps justified in digressing from our main interest in articular analysis to consider in some detail what metric analysis can (and cannot) accomplish. In the section which follows, we will use as an example of metric analysis a set of analytic programs (originally for metallographs) which we have experimentally applied to objects of biological interest. After this discussion of metric analysis we will proceed, in the sections which follow, to a point-by-point examination of a prospective system for mechanized articular analysis.

Metric Analysis of Patterns

To perform a metric analysis we must first of all have a transducer (i.e., a scanner) able to sense the pictorial information and to convert it to a computer-usable form. The scanner used for the experiments reported here is of the rotating-drum type first described in Kirsch, Cahn, Ray and Urban[2] and in a more recent version by Moore, Stark and Cahn.[3]

Since it is used for making an off-line magnetic tape recording, the scanner in question is forced to make an exhaustive scan of the picture to be processed, rather than the more selective scan possible with a scanner directly connected to and programmed by a computer.

Our present device scans a photographic print with 192 resolution elements per lineal inch. The average optical density of each of these resolution elements is represented by a three binary digit number. Although the resulting eight levels of optical density are indeed rather few, very much higher optical density precision as in CYDAC[4] would in our case produce largely spurious accuracy. This is because we are presently dealing with a photographic intermediate, rather than with a directly scanned image as CYDAC does. With such a presentation of a biological image, and without any structural information, it is possible to perform several kinds of metric analysis. One common kind of such analysis is cell-counting and -sizing; the results of a simple counting and sizing operation are shown in FIGURES 1, 2 and 3. FIGURE 1 shows a photograph of a brain section in a Nissl stained preparation. The image is substantially black and white, as a result of photographic "information-reduction." When scanned by the prototype scanner and redisplayed on a cathode-ray tube by the SEAC computer, the otherwise unprocessed image appears as in FIGURE 2. It is then necessary to determine which of the cells in the machine-scanned image are to be analyzed. We solved
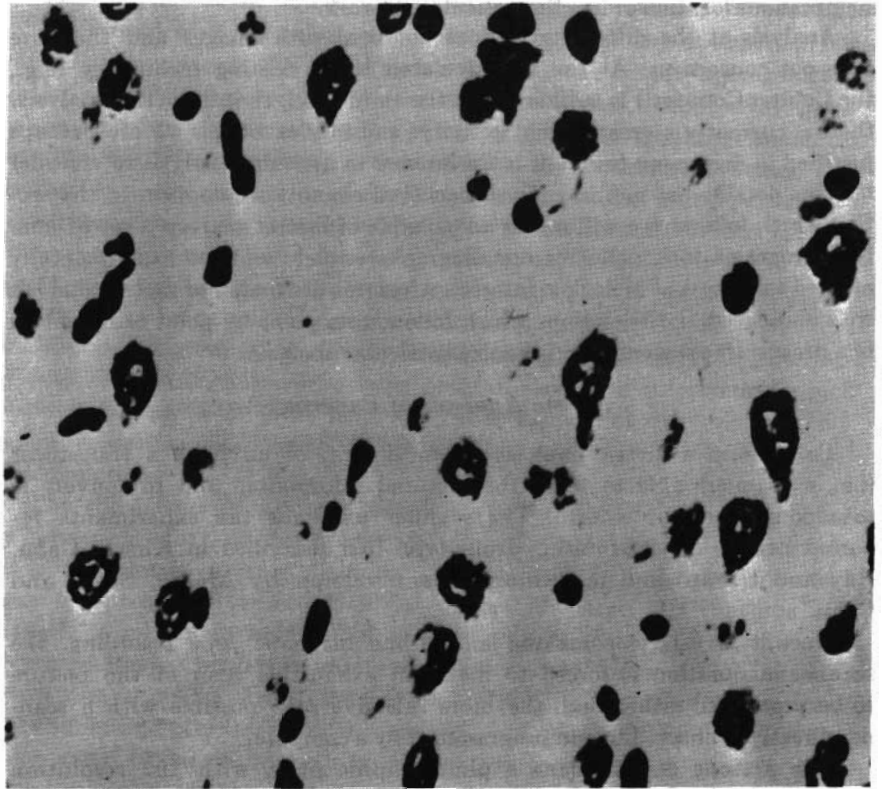
FIGURE 1. From a gallocyanin stained section of normal monkey cortex. The image has been photographically preprocessed to maximize contrast. The area indicated within the rectangle, corresponds to the memoscope display of FIGURE 2.

the problem trivially by asking a neuropathologist to use a pointing device, in this case a light-pen. The cells which were pointed to by the light-pen are shown, as redisplayed on the computer in FIGURE 3. Since the photograph of FIGURE 3 is represented internally in the computer as a binary matrix, it is possible in this highly simplified case to use a simple definition to separate cells from each other. We merely specify that if two of the white resolution-elements in the picture are horizontally or vertically adjacent to each other, then they are part of the same object. For the purpose of size measurement, the disjoint sets of such resolution elements constitute by definition the objects measured by the simple machine program. In the case of FIGURE 3, the sizes of the cells in terms of resolution elements (proceeding from left to right and top to bottom) are 88, 113, 125, 63, 99 and 81. We note yet again that, for such

a measurement program, it is not necessary to describe any of the structures of the cells to the machine.

Many biological images do not present an articular structure as favorable for machine measurement as this example, as can easily be seen from FIGURES 4, 5 and 6. FIGURE 4 is a high-contrast photograph of a spread of pigmented epithelium of the retina of a chick embryo (obtained through the courtesy of Dr. A. J. Coulombre). FIGURE 5 is the scanned, quantized, and redisplayed computer version of the original image. If we now select, by the same light-pen technique, a set of cells for analysis the objects isolated by the same program used for FIGURE 3 are as shown in FIGURE 6. It is clear that these are not the objects in the preparation which a biologist would want to analyze. It is just as clear that the failure of a machine procedure to identify the proper object for analysis is attributable to the fact that the machine has been given no information about the *structure* of the images. There remain, of course, instances of images specially prepared (histologically and photographically) so as to permit needed quantitative information to be obtained through simple procedures for isolating the object. The computer identification procedures (in contrast to computer *analysis* procedures) then become the same as for disarticulated analysis. Other examples of information obtain-
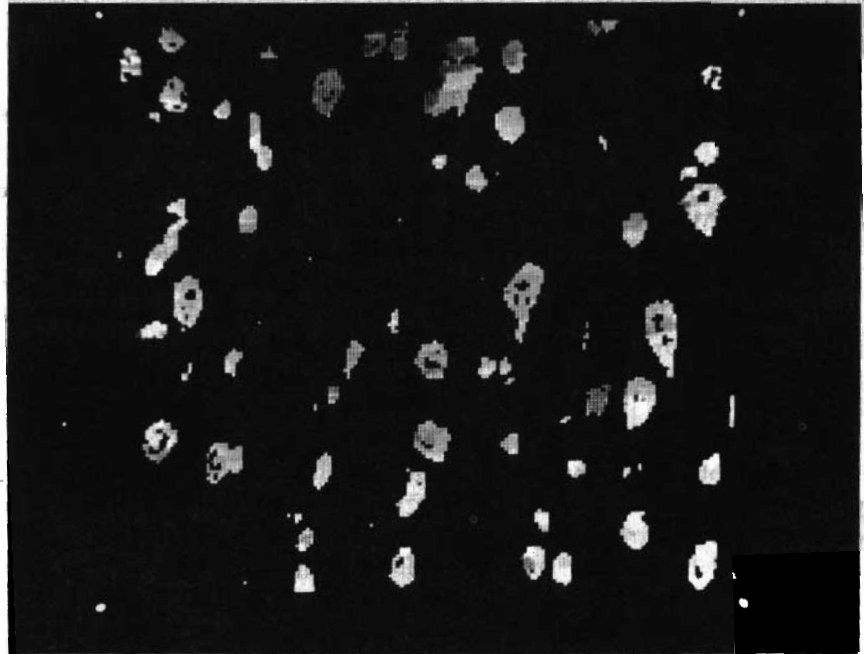

FIGURE 2. Memoscope display of SEAC memory corresponding to scanned input from FIGURE 1.
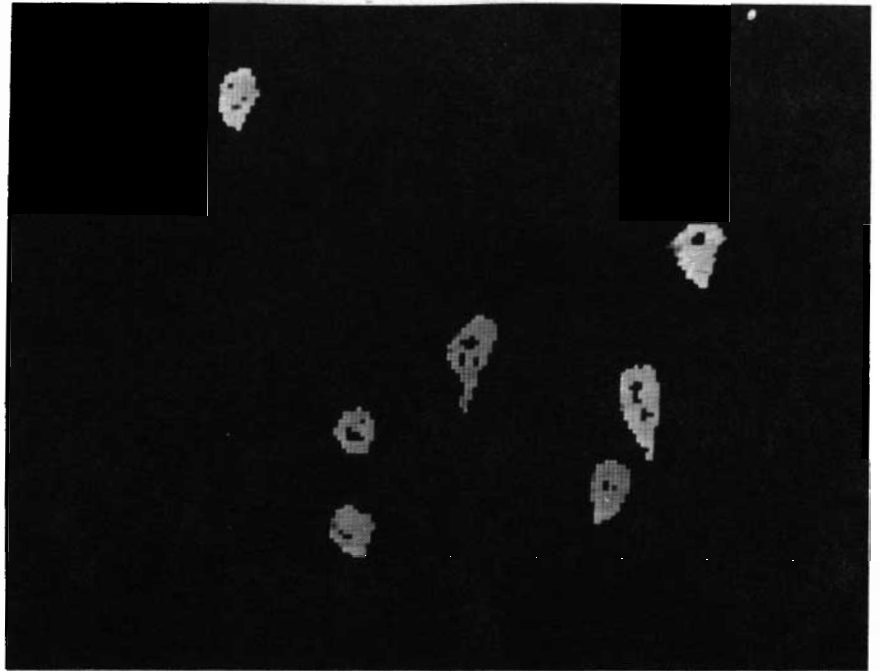
FIGURE 3. Memoscope display following light-pen pointing experiment. See text.

able by a variety of forms of metric analysis of microphotographs are described by Moore.[5]

It is evident that if we are to perform a more complex analysis of images such as those in FIGURES 1 and 4, and of other images even more complex, it will be necessary to somehow acquaint the machine with the structure of these images. We must describe to the machine their architectonics: how they are constituted of their pictorial components, how these in turn are constituted of others, and so on to the deepest level of articulation.

In attempting to convey such structural information to a computer, we meet unexpected obstacles. Among the first things we realized in the course of our research in this area, was that the information "in" the images would have to be amplified before it could be meaningfully processed further; we realized that a process which biologists perform unconsciously, for the most part, would have to be made explicit before it could be made performable by the computer. This alone is proving an exacting job; but there are other tasks equally demanding. We have concluded, surprisingly perhaps, that it will not be difficult in principle to

arrange for the computer to perform elementary structural analyses; and that a steady increase in complexity of analyses (and of images analyzed) can be anticipated with some confidence. But, we have also found that it is far from easy to know when the machine analyses are indeed approaching adequacy, and hard also to strengthen the machine's capabilities in any systematic way unless we make possible a satisfactory interchange between the computer (revealing what *it* knows) and the biologist (explaining what *he* knows). To solve both of these problems—or rather to contribute to their solution— we plan to rely on the presentation by the machine of images which it has synthesized in accordance with its state of knowledge at the time. In this way the state of its knowledge can be quickly and accurately gauged, and with suitable work advanced. We have also come to see that these synthesized images will not become valueless once the computer-system is fully operational; that their usefulness is likely to have more continuity than we had first conjectured.

In view of their value as vehicles for testing the machine's accretion of information, and also their value as continuing supports of analysis, these synthesized images merit a brief exposition of their own.
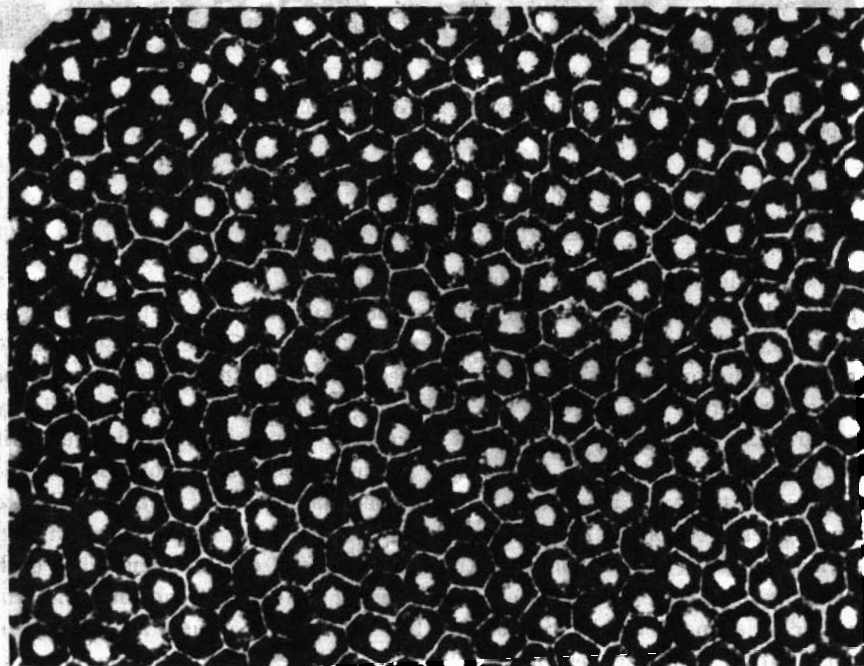

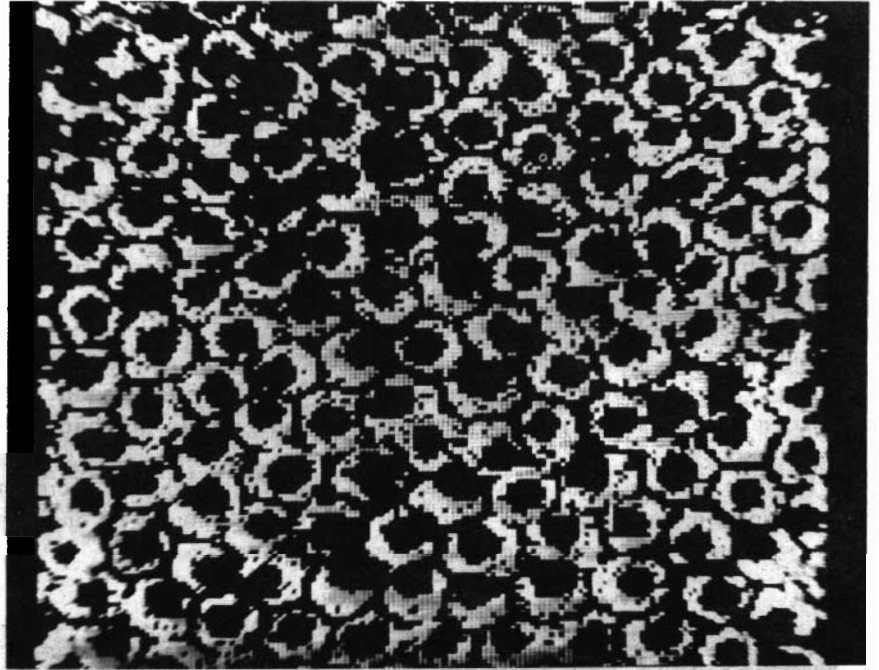FIGURE 4. Spread of chick embryo retina unstained.

FIGURE 5. Memoscope display of SEAC memory corresponding to scanned input from FIGURE 4.

The Utility of Synthetic Images

To begin with, we will admit as "adequate" a given analysis of a pictorial object only if the picture can be resurrected from that analysis. That is, we demand *at least* (not at most) that the machine be able to resurrect the original, in the sense of presenting a faithful rendition whose departures from the original are acceptable with respect to some fidelity criterion. This, however, while a necessary condition of adequacy, is scarcely a sufficient one. That this is true becomes obvious when one realizes that any of a number of quite peculiar analyses would suffice to permit resurrection of the original, in our sense; for example, a characterization of the picture obtained by recording the optical density of each of its resolution elements would suffice for this under normal (macroscopic) viewing conditions. Clearly, this will not do; such a travesty of articulation (viewing resolution elements as meaningful components) is not what we are after. Yet, how are we to explicate what we *are* after except by putting forward some "intuitive" notion of what are the artic-

ular structures imposed by the biologist? The answer lies in making a further demand of pictorial synthesis: that any given image must be resurrectable, not only as a faithful rendition of the original, but also as a *generalized instance* of the *class-membership* of the original. That is, we demand that the computer be able to supply us, for any presented picture, with synthesized images which depart from the original only insofar as is possible without shifting the original from one class of images to another: where the classes are defined by the morphologist. We could begin by demanding, for a picture of a given large bipolar neuron, a set of synthetic images: one would duplicate the original exactly; but the others would simply present other examples of large bipolar neurons. The ability to produce such a set of images from an unidentified original using knowledge based on analyses of a set of images containing objects of the same class represents a significant advance in depth of analysis; for such an ability can rest only on an articulation of relevant pictorial elements and on a differentiation between what is essentially an element of any particular neuron, making it "large, bipolar," and what is idiosyncratically present in one particular cell. The way is then open for the next
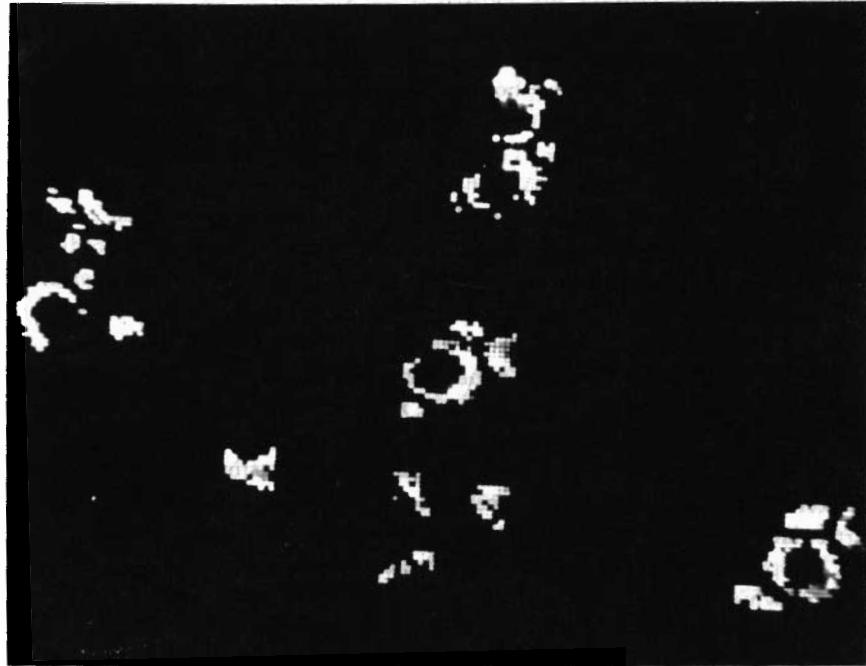

FIGURE 6. Memoscope display following light-pen pointing experiment. See text.

step: that of having the computer on demand, present not a set of images but only one: one image which has all of the properties common to the set of large bipolar neurons, without necessarily having all of the properties of any one such neuron.

A machine which can present a "generalized instance," in this sense, proves by so doing that it can analyze the presented image in terms of its articular (or architectonic) primitives and the structures which relate them. The qualities which make the cell a neuron, in our example, and which make it large and bipolar will have been singled out and reproduced more abstractly; obviously, if we are to demand such a singling-out for *any* "large bipolar neuron" which we choose to present, at whatever magnification and in whatever orientation, obviously no characterization of each such image in terms of its resolution elements will provide a very effective way of meeting this demand.

The first use of synthetic images then is that they constitute graphic representations of the machine's state of knowledge at each stage of development, thus providing us with a direct way of judging how close the machine is to meeting our criteria for adequacy. The second use of synthetic images is related to the first. For we have no real reason for believing that our a priori criteria are sufficient to guarantee machine-production of only well-formed images; the machine may meet all of our criteria and still fail to satisfy us. In this way, the display of synthetic images will serve to elicit from the morphologist additional criteria as he sees they are needed (and this point need not come only after all of his a priori criteria have been satisfied). Thus, thankfully, there will be no necessity for the morphologist to prepare in advance, before he begins to convey information to the computer, all of the information that will be required in the end. Even before the system is operational then, it may produce a side-benefit, in evoking, during the course of its development, criteria for e.g., "neuronhood" which the biologist has hitherto been unaware of or ignored.

The process of providing generalized instances will also serve this elicitative purpose. For example, a highly generalized instance of a brain-section might replace all cells by letters, e.g., "N" for "neuron," "O" for "oligodendroglia," etc. A pictorial presentation in these or similar terms might well, we think, serve to evoke more readily the biologist's articulations of intercell structural relationships insofar as he sees these as independent of individual cell shapes. Here again, insights may be evoked which have heretofore been passed over.

Pictorial Augmentation

If the machine has truly been enabled to articulate cells into their proper components, and to structure these components in a meaningful way, then these synthesizing powers can be used for yet another, ancil-

lary, purpose. That is, images which are deficient in minor ways can be synthetically improved: or, what is the same thing, the machine can analyze deficient images as if they were not. This is a minor point, but one of significance to any serious research in machine-handling of biological images; for most of the images one would like analyzed are "deficient" in one or more ways. One deficiency which the machine can correct in a generalized way resides in the fact that a brain-section is never complete in the sense of showing as complete objects all of the cells it contains in part. Axons and dendrites invariably trail off into adjacent sections (which one may or may not have available); a given section of a neuron may not have captured the nucleus; and so on. Moreover, sections commonly contain artifacts of other sorts; for example, a cell may have been separated from one of its dendrites by the process of sectioning.

One could accomplish a great deal, to be sure, even if one were forced to provide only perfectly-prepared sections (or photographs thereof); but one would certainly chafe at such a restriction. And no section, as pointed out above, is "complete." It is then highly desirable for the machine to *augment* the pictures presented to it, performing synthesizing operations of a rather minor nature (we now believe) in order to make possible a more valuable analysis.

An image analysis of an articular sort requires a prior recognition of the structure of the image. It is not sufficient to be able to name (and thus classify) an image into one of a small number of categories. The reason for this is that the analytic properties of an image are usually not properties of the whole image, but rather of articulated subparts. Consequently, character-recognition-like techniques are of very limited use in biological image processing. As examples of such more specialized (and less useful) techniques we may consider two which have demonstrated worth in character recognition. The first procedure of character recognition is useful in cases where images are template or prototype derived.[6] For such images, a matching procedure against the prototype is sufficient to identify them. Unfortunately, biological images do not generally have, in any graphic sense, templates against which they can be matched. If there is to be a matching, it is at a more abstract level: the level of syntactical description.

A second somewhat more elaborate procedure still appears inadequate. This is the use of the decision tree.[7] A decision tree can be used to predict recognition tests upon the results of previous tests and thus to allow recognition of a more complex sort. It still fails, however, to provide what a full syntactic description provides, namely the ability to resurrect parts of the image which are manifestly absent, parts which, though they may be detected, cannot be recognized (to use Unger's distinction[8]) without invoking a priori knowledge about the class of images of which the one under test is an instance. To have a syntactic description of a biological

image is to be able to synthesize it according to that description, to be able to analyze according to the synthesis procedure, to be able to augment with information about class membership of its structural parts, and finally to be able to augment it by supplying parts which from biological considerations are present though pictorially manifestly absent.

The "syntactic" treatment we have spoken of is to be underlain by an *iconic grammar*. We will return to a brief prescription for such a grammar after the following section, which is devoted to describing the PLACEBO series of *linguistic grammars*.

The PLACEBO Microgrammar

We have dwelt at length on the complexity of the biological images on which our interest centers, and we have alluded *(see The Analysis of Biological Images)* to the fact that, due to their complex and structural nature, they are not easily characterized, for human understanding at least, in the form of detailed numerical measurements. We believe that the consequence is clear: that if the morphologist is to describe such images in such a way as to do justice to their structure, some sort of *language* must be offered him. In this section we will take up the question of what language is best for this purpose, all things considered; and the question of how such a linguistic facility is to be provided.

It may be surprising that we devote much attention to such questions at all. For a more normal reaction would be to select a large general-purpose language for this purpose—say IPL-V[9]—and adapt it where necessitated by our special demands. A more unusual solution, but still in keeping with general practice, would be to select some language of formal logic—say the first-order predicate calculus[10]—for this purpose. That we have not done so warrants some explanation; but an adequate explanation is rather peripheral to the main points here at issue, and would require, to make much sense, more time than would be appropriate. We, therefore, beg the reader's indulgence in resting our case on a banality: that formal languages of any kind have, from our standpoint, four clear disadvantages. They represent an extra investment of time on the part of their prospective users—and of such time as may well be resented; they must in any case be greatly adapted to fit the needs of biological description, thus losing, in forfeiting ready availability, one of their chief advantages; and no one, no matter how adept in the use of these languages, becomes as easy with them as he already is with his native tongue. Lastly, if we force the morphologist to describe the structure of biological images in the painfully detailed way that characterizes formal languages, we may expect to have greatly compromised his description.

Thus, for our language we have chosen English, or more accurately a semblance thereof. In doing so we undertake a difficult task, it is true; but if we succeed we will have provided the biologist, in one stroke, not

only with a language he has used for description throughout his professional career, but also with a language which is adequate, insofar as any language is, for his descriptive ends. The advantages of using English will be fought by few; but the possibility of providing it will seem farfetched to many. Yet, as it turns out, we are actually closer to this goal than to any other.

A procedure for interpreting the sentences of a simple language is described in Kirsch,[11] where it is shown how some of the information content of sentences derives from those sentences "syntatic analysis," or "parse." That is, the very way in which sentences are organized, manifesting the regularities of English grammar, can be analyzed and implemented as conveying information. English sentences, then, may be processed in somewhat the same way as, e.g., IPL-V statements. Seen in this way, the apparently manifold differences between IPL-V and English are somewhat reduced.

The language described in Kirsch (*ibid.*), though fitted with interpretive routines which make it more like IPL-V, is much too small to serve the biologist. A larger and more flexible language will be needed: and one is under development. A language containing about eleven hundred rules is reported in Watt[12]; the same language, expanded to include about twenty-one hundred rules, is reported in Watt,[13] forthcoming. These two papers discuss consecutive stages of development of essentially the same tool; the stages are each independent and "complete" in some sense, however. Named "PLACEBO IV" and "PLACEBO V," they represent what we think two long strides toward our goal; and "PLACEBO VI," now well under way, is another such step. In our current planning the final language, "PLACEBO VIII," is expected to contain on the order of 10,000 rules, several times the number contained in the aforementioned "PLACEBO IV." These figures are meaningful only if it is realized how limited a tool PLACEBO IV really is—only if it is realized that PLACEBO VIII will be orders of magnitude smaller than would be necessary to accommodate all of English. Thus, only a small subpart of English will be offered the biologist; and it is natural to ask, then, how we hope to provide, with such a language, a descriptive tool of any usefulness. It is this point, in fact, this necessity of providing the facility of English with much less than English, which forms a crucial part of our planning; we will return to it just below.

PLACEBO IV and its sequels are "microgrammars" in Watt's[14] coinage; this term has been informally defined to cover subsets of English grammatical rules, roughly speaking, which analyze and synthesize ("determine") a coherent set of English sentences. A microgrammar is then a linguistic object having as its correlate the kind of iconic or pictorial grammar presented below. A "coherent subset of English sentences" is a term which must be defined with reference to prospective *users* of the

microgrammar—being coherent is an extrinsic property—and we mean by it that a microgrammar must be such as to be more or less uniform in embracing speakers' coverage of a limited field of linguistic activity. A large microgrammar might determine a professional dialect, for example, such as the dialect in use among neuropathologists in their professional capacities; a somewhat smaller microgrammar might determine just the subdialect which neuropathologists use in describing sections of brain-tissue. A still smaller microgrammar, such as PLACEBO VIII, might determine a large subset of this subdialect—say, almost all of the English sentences which neuropathologists use in describing brain-tissue in a highly circumscribed situation, as when addressing an audience of medical students.

Let us look very briefly at what sort of "rules" these microgrammars contain. The PLACEBO microgrammars are written in what is known as the "Context-Free Phrase Structure Model," as defined and formalized by Noam Chomsky[15]; because we may also use this model to determine image, we have taken to terming it simply "Context-Free Constituent," or CFC. In the CFC model, rules take the form of very simple "rewrite" rules, where a single symbol is "rewritten" or "instantiated" as one or more other symbols. These other symbols are rewritten or instantiated in turn, one by one, until so-called "terminal" symbols are instantiated, at which point the rewriting halts. These "terminal" symbols are in our case words, parts of words, or brief idioms. Thus, starting from a single initiating symbol, say "S" (for "Sentence", of course), the rules operate so as to produce strings of words forming English sentences.
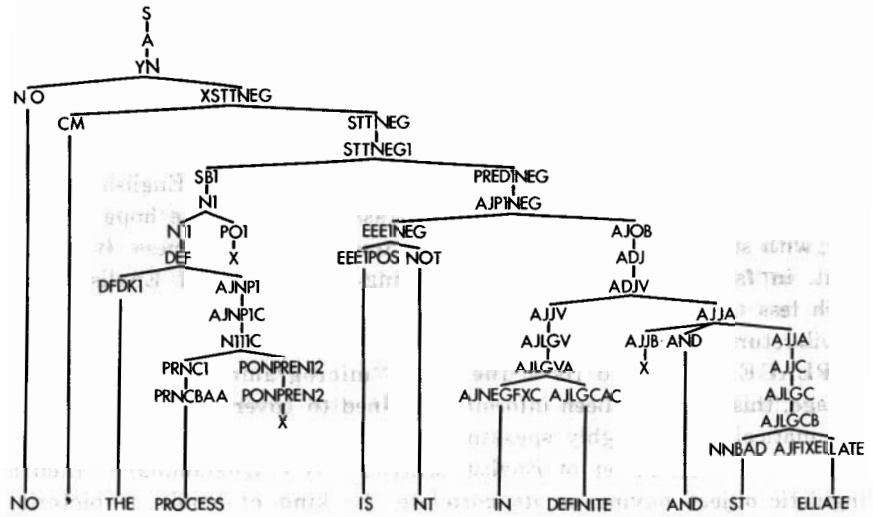

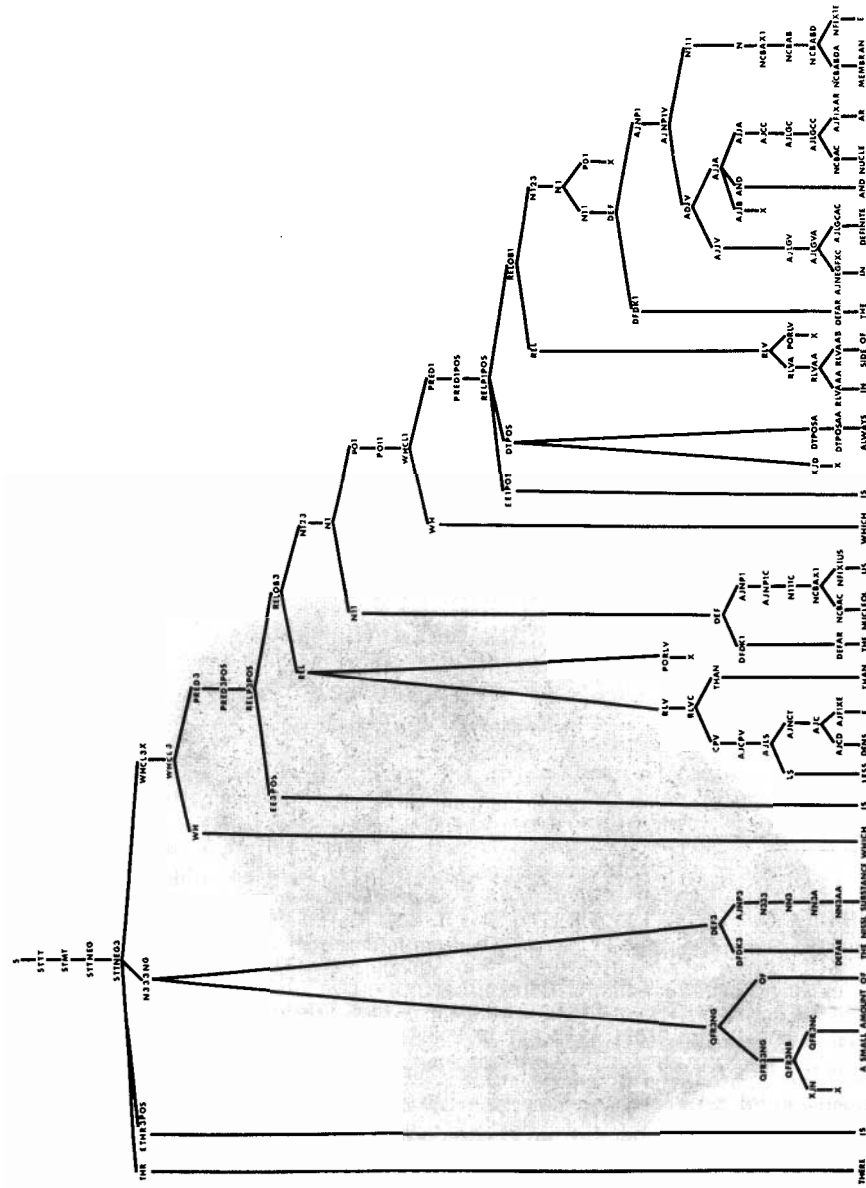FIGURE 7a. The structure of a PLACEBO IV sentence.

FIGURE 7*b*. The structure of a PLACEBO VI sentence.

FIGURE 7*a* shows the sentence "No, the process isn't indefinite and stellate." The way in which this sentence is produced by the rules of PLACEBO IV[12] is also shown in FIGURE 7*a*, in a "tree" representation; "S" has been rewritten, and its rewrites rewritten, until a sentence has eventuated. Also, since "determination" goes both ways, the way in which the sentence is produced is the way in which the sentence, if presented cold, will be analyzed. The "rewriting" process then works in reverse, rewriting the words and word-parts backwards until "S" is reached and the sentence, so to speak, is accepted as an example of "S", or as a well-formed sentence.

FIGURE 7*b* shows the sentence "There is a small amount of the Nissl substance which is less dense than the nucleolus which is always inside of the indefinite and nuclear membrane." The derivation of this sentence, from the rules of PLACEBO VI, is also exhibited in "tree" form.

In the course of its development, the PLACEBO microgrammar is to serve a definite function in the building of our system. That is, it will serve as the descriptive medium in which the articular structure of
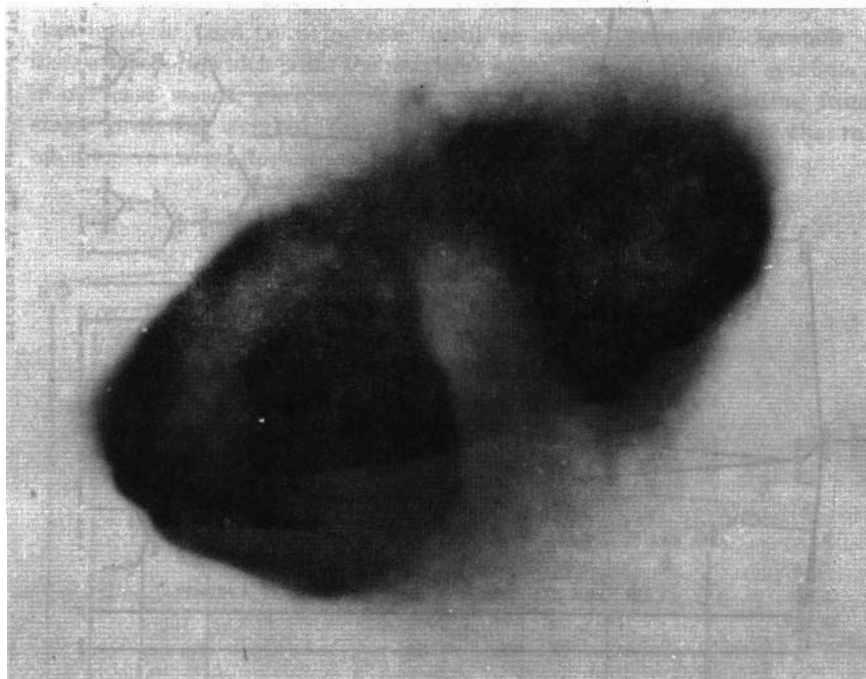

FIGURE 8. From one of a sequence of optical serial sections of gallocyanin stained normal cortex. The area represented in FIGURE 9 is outlined.

```
        1        11111111111122222211
      1111    11222222222222233333222
      1111111233333433333334444333
      11112333344555555444455555554
      11222334566777776665555666666666
      11112233345677777777776666666667777
      11223344455777777777776666666677777
      11223445555677777777777666666777777
      111233455556667777777777766666777777776
      1123445555666667777777777776677777776666
      11123455556566666677777777777777777776776
      1133455666666666666777777777777777777776766
      11335556666666666666677777777777777777776666
      12345666777666666605666677777777777777777776667
      13566777776666666666555566677777777777777776677
      13577776665556666666555555556777777777777766777
      1367777665554566666665444444567777777766677766767
      136776666655555665666665443334456777777776666776666
      13777765555555666666666554433334567777776677766666
      1377766555566666676666556543333334567777777777777766
      13677765555556667777776666543333334455677777777777777
      1367776555555566777777777776066443333445567777777777777777
      13577765555555556777777777777776443333445567777777777777777
      13677776655556667777777777777776443333445556777777777777777
      3677777655555567777777777777777775433344455567777777777777777
      25777776555567777777777777777776433344555567777777777777777
      13777776555556777777777777777777765333445555667777777777777777
      13677777655666777777777777777777776444455555556677777777777777
      135777777766566777777777777777777777644445555555667777777777776
      135777777776555667777777777777777777777765445455555555677777777765
      135777777776655667777777777777777777777775445455555555577777777754
      236777777766556667777777777777777777777777765555544555455777776543
      23677777766556667777777777777777777777765555544454445677765432
      13677777766666666677777777777777777777765555554444443335555554321
      13677777777666667777777777777777777777776555544443333334443321l
      1367777777776666777777777777777777777777765544433332223333322211l
      15777777777777777777777777777777777777777765544433332222221111l
      13777777777777777777777777777777777777777765555444333321112111
      15777777777777777777777777777777777777777765544433332221111
      36777777777777777777777777777777777777776555444333222111111
      3677777777777777777777777777777777777765443332222221111
      2577777777777777777777777777777777777776544332222221l
      267777777777777777777777777777777765443332211111
      3777777777777777777777777777777776544333222111l
      2577777777777777777777777777777776544333222111
      13577777777777777777777777777544433332211111
      12367777777777777777777777755433333222111l
      123577777777777765433332222111111
      1134567765443332222212111 1
      112222222221211111111l
        1 11 111111
```

FIGURE 9. Print-out with 8 levels of intensity representation. The distortion is due to the difference in aspect ratio occasioned by the type front of the printer.

biological photographs to be machine-analyzed can be made explicit to the machine through the mechanisms of interpretation. As the system nears completion, this function gradually will phase out, and will have only occasional use as it becomes necessary, from time to time, to provide the machine with additional information.

The microgrammar will have a somewhat similar function after the system becomes operational. At that point, users of the system will communicate with it via the last PLACEBO microgrammar, querying the machine in PLACEBO-analyzable questions, directing image-processing in PLACEBO-analyzable directives. This second function, obviously, requires a somewhat different set of sentences from those required to fulfill the first function; but all-in-all it is more efficient to embody both functional facilities in one device rather than in two, in view of the appreciable intersection between the two sets of required sentence-structures.
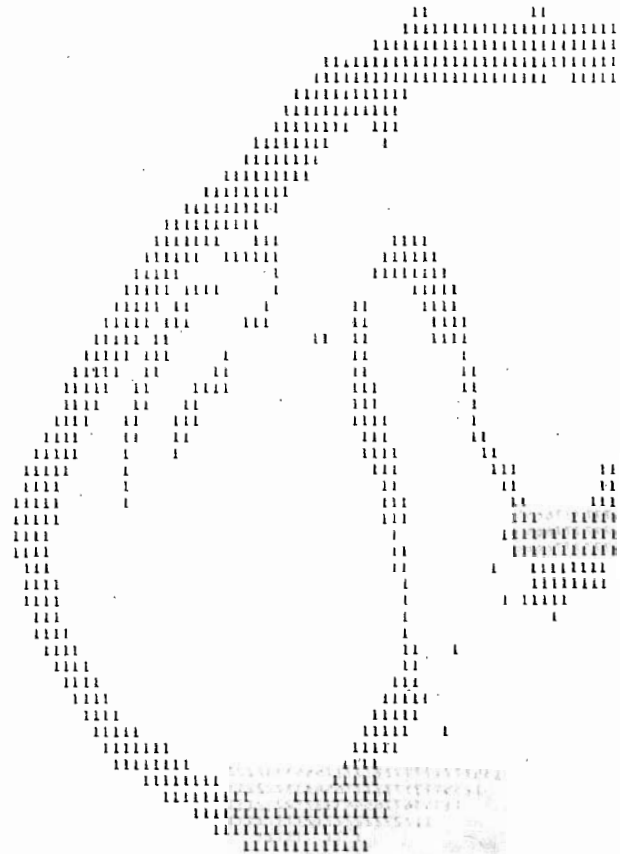
FIGURE 10. Print-out of the results of a spatial differentiation of FIGURE 9.

It may be worth pointing out, parenthetically, that our demand that the machine reply in PLACEBO-analyzable sentences has imposed a requirement of "closure" on the microgrammar: that is, PLACEBO VI and its predecessors have been constructed so that, if the rules are followed, only English sentences can be produced. In fact, we test the microgrammar's closure by generating sentences at random by the computer.

It must have occurred to the reader by now to ask how we expect the user to memorize the thousands of rules of PLACEBO VIII; or how we have come to believe that memorizing these rules is easier than learning, e.g., IPL-V. The answer to this just question is that we do not mean for anyone to memorize so many rules—if indeed this is possible; in fact, our intention is that the rules of PLACEBO VIII will be completely unknown to the person using them. The paradox is only apparent. For we also mean to make the microgrammar "habitable," in Watt's sense,[14]

that is, the microgrammar will be able to accept enough of the sentences likeliest to be produced in the using situation, as to reject submissions only very occasionally. We rely on two constraints to limit the variety of sentential submissions: the fact that the only users will be neuropathologists in their professional capacity; the fact that they will be seated at a computer-console and will thereby be somewhat inhibited (otherwise regrettably) in their speech. Thus both vocabulary and sentential structure will be greatly limited *just by the constraints of the using situation;* the users would have no need to call on all of English even if they could. In fact we think it evident that they will not even call on all of their professional dialect. Herein lies our hope of "habitability"—our hope of incorporating, after studies of usage, most of the comparatively few sentences which will actually be presented. If this hope seems unwarranted, we submit that even the rather weak PLACEBO V has succeeded in netting an encouraging proportion—about one-fourth—of the sentences used in a very limited descriptive situation.

In any case, if our "habitability" hypothesis can be accepted for a moment, it is easy to see why we claim that "English", in the form of PLACEBO, is superior for our purposes to, e.g., IPL-V: not only does it have greater descriptive power; but also it need not be learned at all, by any neuropathologist who speaks English.
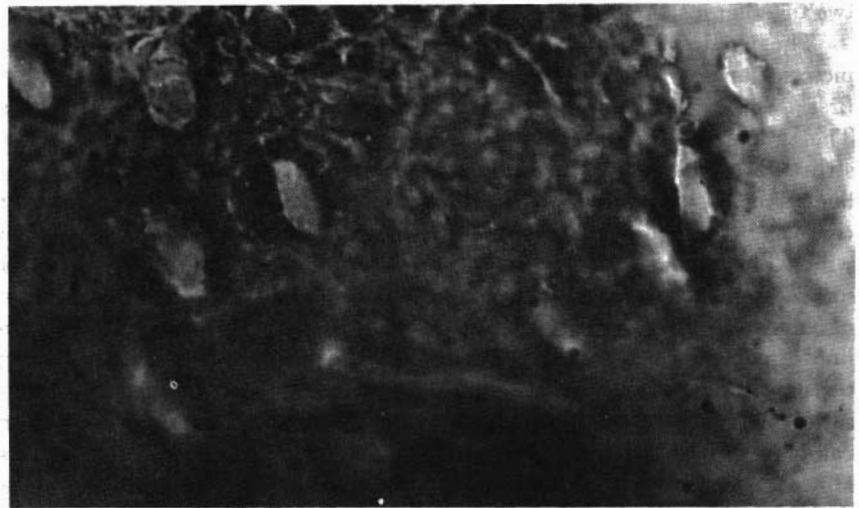

FIGURE 11. From one of a sequence of optical serial sections of unstained cortical tissue.

The Iconic Microgrammar

In earlier sections we have argued that highly articular biological images are best described in English, rather than in terms of numerical characterizations; and that they are best analyzed with respect to a pictorial or iconic "grammar," rather than with respect to reticles or templates. In the section just preceding we have told something about PLACEBO, the English-handling faculty of the system we are building; in this section we briefly describe the iconic microgrammar needed for this system. We emphasize that such an iconic microgrammar as we envision is in its earliest planning stages, though experiments with constructing simple iconic grammars have satisfied us that such devices are not only feasible, but can very likely be put together with the aid of well-proven linguistic techniques such as those used, in part, to build PLACEBO. Naturally, these techniques must be adapted somewhat to fit the needs of iconics.

It may be useful at this point to recapitulate very briefly the uses of iconic microgrammars. As we have already intimated, much of the description of images can be at an entirely linguistic (PLACEBO) level; but a large part remains which is best described, to a machine as to a human, by exhibiting instances of the images under survey. A complete system for the processing of biological images should include both linguistic and iconic facilities. For completeness, we may view the problem of the description of synthetic images as consisting of a purely linguistic part, a purely iconic part, and also of a third part which binds the first two together insofar as this is possible.

The iconic description of image structure is accomplished with the pictorial analog of a linguistic grammar. This iconic grammar has, at its first level, purely pictorial primitives: simple shapes and line-segments. The simple shapes are chosen as those which have no internal articular structure of interest to a neuropathologist or general morphologist. The line-segments are used to circumscribe areas in the image, to join primitive shapes together and to construct composite structures. The way in which line-segments may be used to construct structures of an iconic sort is well illustrated by Sutherland.[16] For purposes of handling images of varying optical density, however, it is necessary to supplement the line drawing system of Sutherland with the ability to specify arbitrary textures and arbitrary blackness values. What's more, the generation of satisfying synthetic analogs to natural images requires contours of a more complex sort than those generated by straight lines and circular arcs. However, it appears to be true for continuous tone images that line-drawing approximations contain, to a substantial extent, enough information for the syntactic structure. In order to compare the biological pictures of interest in this paper with the line-drawing approximations to

```
44333323342112333312222234543334443442343333321123333333333443444444332222
54433322321122234212722234554334433442234344323333333333334434933333222233
443333222112223321112222345544333444321344432343222333333344333322222233334
432233222122223311112222335554333355543222343223311222233432222221122233443
44222222222222221111122233444433355554332234333322222113322111111122234443
44322222222222332223333322234443344555543333354323222 1233332111122233333
44322222222222332233322222222344434444443333335532222222211223321122223333333
443233322223333212222222223344335554443233334432112232212233323333333333
4432333222333333222222221222134544665543222334332222212222222223444333333
54323333222333222222712227233454456654322234433333222333322222233333333333
54333333322232222222211111222345455565322112333333333323334443111212222222222
554333334333332222221111121122354444443211223333333333333343322222222211111
5544333444433321112211122222355544343222333333333333343433333332122211111
544444445543333212222222222223555543333333333333333343333342223333112332222
444444445554322332223322222222234555544444433333332334433333213333321233222
443444455543233232222232331235555544444434333333334332222233333322322111
4333334445554222222232222232123555554434444333533333334432233333333232221
33333333444554232222222222222223555543334333343333344443212333322212233322 1
33223334444443322222232322222355554333333334422334443222233322112333221
332333444444322222233332222223355554433333333344333333333222223332211123221222
22333444444444222333323333332333455443333333333344444433322222223333222221211222
23333444433344421233333333333344555543333443334433344444432232222333321223211222
334444443333443211233322232234554443334432333444443223333323334321233211222
45444444333334443222233321222345544443344332234444322334444333332222222222
55433444334444422222222122344554433333333334442233344444222333221122322
544334443345455543333221234455544332222333344443332333333323443333221111112322
44334444344555555555442233455554432222333444433333334443322333432211122232
544344433444555555555554344554454333334444444333333332222223333222222222
55444443334445554455555544455554443333334444433333222111222222333222322333222
554454433334444444455555555554444543333334444333332222111112333334432233333332
44355443333444333444555444444444333333333333332112333333444432222223333
44445443333343333334444444433344433222343333333333333213344443333332222112233
34444433334433334433444444443333443322343333333333333333444443333333322221123
344444333333444444444443333343222333333334444334444433333333333333222122
45554433444444333444444455444444433234443222333333343334333344433332333333332212
55555544444554444444455554444444443333333333333333333343233432222233333333222
55555544455554444554455555544444444443333333333333333334433221111123333333333211
555444444555444444444544334444444444332333323333444422        1233333443211
5444444445555545544545543333444555554443333222333334443432        12333443211
43343334455444455455456564333333444555544333212333344444321        11233342122
44443333445544445455566544443223344555544322212344444444432        1123442233
555554455555554544554455565444432222344555444433222344444444421        1123343333
55665435555444544555555544443332233444444333333334444444421  1     1123443332
45556445544455555555554443333333444433333334443234444444211        1112344333
44555553444444554455555555544332334444433333334444323344444431 11 1   111334433
444455554444444455555555554432334444443332344444334444444443111111   111233333
444444444444444555555555544323333444443332233333334444444421111   111233333
44444444444444445555555555444444444444443322333333434444444442  11  1111123333
44444444444444455556665555555554444442223335533444444444431 111111111112344
44444544444444455566555555555554444444344421133333334444445542111 111111112334
44444444444444445555544444455444444444444443111333333344444455531111111 111112344
44444444444444455554544444554444444444444443212333333334444455541111111 111112344
444444444555442233455433334444454444444444321213333334434444455431111 1111112344
55545454444553233444444443334455554444444433222334443444444444321 11111111111244
444454444454333333334444444445554444444444433323333444444444444311111111111234
43334554444443323333344444444455554443344433233434444444444442211111111111334
43334554444443222223333444444444555544444344322233333444444444442111111111112344
43334444444542222222333334433334554454434442212333334444444444444211 1112333
44333444444442222222221333333333334445555444442122233334444444444443112111112222
454434344444443333332233223223334444544455554333333323444444434444444332211222
555444444444444443322333322223333334555554455443334434333343333333333333222
555444444444444443322233222222233445555555544433333343333343333333333333333222
55554444434444444322222222221122334555555544332222332233333333333222223333322
555555444444444443322222222111112344555555444332112232223333334333332222333332
5555555555444544443222222222111112344455554444332211233222232333333332222333322
5555555566655555444322222221111111112345555544443332223332222222222232222333222
4555555676555544443322222111111111123455555444333333332221122322222222222333322
4555556677655554443332222111111111123455555444333332233333321223333222211222333222
455555567766554433332221111111121235556655443322223333433333333332222211222222
44555555566666655433332211111111222355555555543222333333343333333333332221112222
44445455556666666543333221111111111235555555554322233333334333333333333222222222
44445555555566666654233322111111111112355555555543222333333333313333333333277777171
```

FIGURE 12. Representation of a part of FIGURE 11 with eight intensity levels
There is the same aspect ratio distortion as in FIGURE 9.

them, we may consider the three pictures in FIGURES 8, 9, and 10. FIGURE 8 is a photomicrograph of two cells in a specimen of monkey cortex. FIGURE 9 is a computer redisplay of the scanned photomicrograph with the blackness levels of the original represented by a blank and the digits 1 through 7, corresponding respectively to an arbitrarily white value in the original through an arbitrarily black value. It is possible to obtain a spatial derivative of the scanned image which simultaneously rejects those parts of the original image which are out of focus and preserves the boundaries where the value of the derivative exceeds some threshhold. Such a derivative or differentiated picture is shown in FIGURE 10. FIGURE 10 is substantially a line-drawing representing some of the information in the original photograph. If a line-drawing similar to that of FIGURE 10 were supplied to the machine as part of the iconic description of the structure of a nerve cell, the machine would be able to identify the differentiated photomicrograph with the synthetic sketch, and thereby be enabled to recognize the structure of the nerve cell. The procedure is not quite so simple, however. Some of the difficulties to be encountered are shown in FIGURES 11, 12 and 13. FIGURE 11 is a photograph of similar monkey cortical tissue, but here left unstained; the photograph was prepared using ordinary light microscopy. FIGURE 12 is the computer print-out of the scanned image where now the structure of the original is not quite so manifest to the eye. In FIGURE 13, the same differentiation process has been applied to the unstained scanned image as was applied in FIGURE 10. Some gross structure is evident in FIGURE 13. But what is also clear is that no sample matching between a line drawing of, say, a single cell and the embedded cell in such a tissue "sketch" as FIGURE 13 is immediately possible. The structure of FIGURE 13 would just have to be decomposed before a matching could be undertaken.

Iconic grammars with properties like those mentioned earlier do not yet exist beyond such simple ones as the one given by Kirsch.[11] However, an iconic grammar appears to be a prerequisite to a satisfactory structural analyses of images.

The ITSELF System: Overview

If we have been at all able to make our position clear, it must be obvious by now that we view the problem of analyzing biological images as one capable of a radically more ambitious construction than is customary. We have suggested in earlier passages that the class of images to be mechanically analyzed should include images of complex structure; and we have proposed several procedural innovations to accomplish such structural analyses. These proposals ought now to be brought together; they are as follows. First, that the description of biological images can
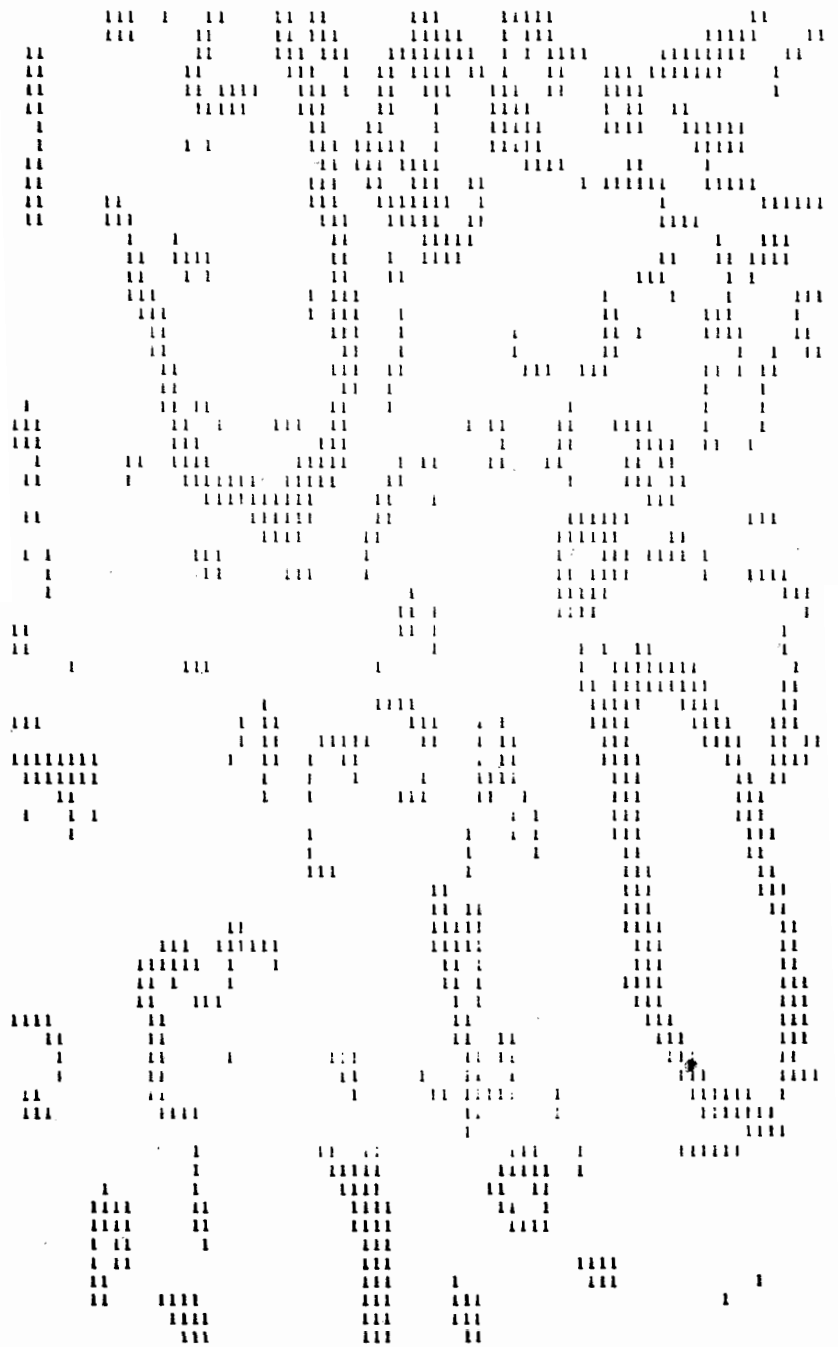
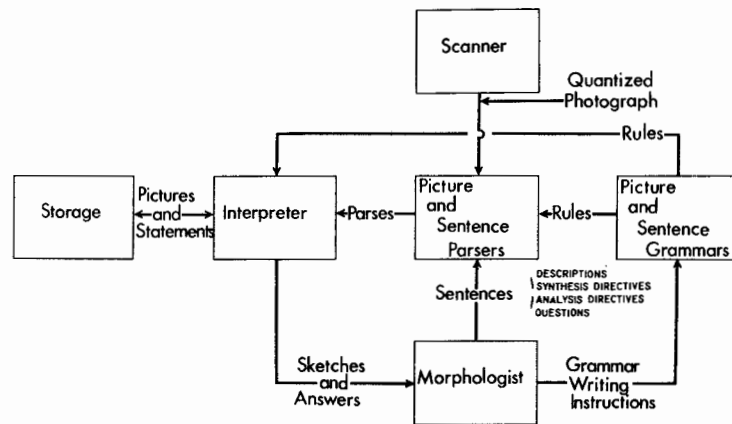FIGURE 13. Print-out of the results of a spatial differentiation of FIGURE 11.

FIGURE 14. An overview of the ITSELF System.

be made at an articular level without exceeding the capabilities of computer systems such as may now be envisioned. Secondly, that for the most part this articular description is best carried, and treated, in its customary linguistic form. But that, third, direct pictorial "description" can also be made via sketches and other synthesized representations. We have grazed a fourth point, that to a certain extent the internal representation of biological images in a computer can be interchangeably linguistic or iconic, and that to this extent (which we can now only guess at) the two determining microgrammars, linguistic and iconic, can be bound together. Fifth, we have suggested that the procedure for analyzing images, whether linguistically or iconically, should be the inverse of the procedure for synthesizing them; a corollary suggestion, bearing on the pictorial aspect alone, is that a good synthesis procedure provides the largest part of a good procedure for analyzing images. We have suggested, sixth, that in providing image-descriptions to the computer the morphologist be aided by use of machine-synthesized images; for these may function as feedback to enable him to improve his description until it meets the adequacy test of an acceptable synthetic image. Seventh, we have proposed that the same microgrammar, which is in use during the construction of an image-analysis system, can also be the microgrammar which is used in interrogating and instructing that system when turned over to the user. And, we now assert that these observations are not disjointed matters of interest but rather that they form the design principles for a coherent system capable of analyzing biological images of high structural complexity. Having made this assertion, we must now try to outline a tentative overall design for such a system which we call "ITSELF."

This system demands, for its realization, various computer processors, some of which are already in existence, others of which are not. In the

discussion which follows, those computer facilities which have already been constructed, whether by us or by others, will not be distinguished from those which have yet to be constructed by anybody. Of course, it would be bad faith on our part in bringing the reader this far if we did not have ample reason to believe that those parts yet to be created are not impossible of creation: and that they can be built with present knowledge or with reasonable extensions thereof. In fact, we would like to take this occasion to urge that suitable effort be devoted to developing those parts now now available, particularly to developing those parts which as of now seem most difficult. We beg leave to point out that such processors, far from having utility only for the "ITSELF" system, can be foreseen to offer wide-ranging advantages to a variety of kindred systems, for many different disciplines.

The diagram of FIGURE 14 summarizes graphically our concepts of how the system might be structured at some future time. The linguistic informant-work leading to the formation of the structure of PLACEBO VIII (though not necessarily its entire lexicon) will have been completed. The initial iconic grammar will be finished for the most part. The grammar-writing instructions shown as proceeding from the morphologist will largely be concerned with construction of then-incomplete aspects of the system, mainly those which bind together the pictorial and linguistic analyzers. Even though PLACEBO IV has attained some power for expressing contextual material, a great deal more will be needed. For example, at the present stage of our work it is possible to recognize that some grammatical constructions applicable to cerebral cortex would not be applicable to the basis pontis. Such contextual modifications of grammatical rules represent an important class (but by no means the only one) of properties to be modified during the early on-line operational phase. (Indeed, iterative grammar modification by the morphologist, as a result of the repeated synthesis-analysis cycles, is crucial to the system.) The unit titled "Picture and Sentence Grammar" thus stands for a highly complex structure. Its complexity will hopefully be reduced by several factors, most of which are not yet seen very clearly. One of them, however, may be the considerable degree of isomorphism which we expect to obtain between the iconic and linguistic grammars.

The group of processors termed "Picture and Sentence Parsers" in the diagram would constitute a nodal point for information-transformation in the system. Via the typewriter these processors will receive the directives, descriptions, and queries of the human morphologist, transmitting to the interpreter not merely these strings of words, but also their analysis as PLACEBO sentences. As noted previously, this analysis formed in accordance with the grammatical rules contains information necessary for formal treatment of the sentence.

A parser such as this is not completely hypothetical. In fact, a parser

for PLACEBO V exists but many factors, chief among them its slowness, render it valueless for any but experimental purposes. A much faster parser will be needed to enable the interval between the entry of questions and the display of answers to be of the order of seconds rather than minutes. A truly efficient parser would be a necessary (though, of course, not sufficient) condition for the system's operability.

Since we cannot at this time show examples of any except the most primitive form of picture grammar, it is a little presumptuous to discuss the iconic portion of the parsing processor. However, this part of the overall prospect should not be regarded as totally bleak. The knowledge gained in constructing efficient sentence parsers will surely be in part transferable to the construction of iconic parsers; in fact, insofar as the expected isomorphism between sentence and picture grammars emerges, the picture and sentence parsers will prove to be similar structures.

The scanner that we are now employing for digitization of photomicrographs will almost certainly not be the one which will transmit such quantized images to the picture parser. Optimism with regard to this component of our system is more firmly based on the hard results of others. The refined techniques embodied in CYDAC,[4] for example, show the possibility of high fidelity and high precision in optical density representation. Computer control of the scanner would seem to be a technical *sine qua non,* and this not only for reasons of storage-economy. The scanner unit in the diagram does not represent a scanner alone, not even as sophisticated a scanner as that of CYDAC. We plan that considerable preprocessing occur between the initial digitization and the presentation of the image to the parser. Preprocessing as well as regional and resolution variability in scan are examples of operations which should be under computer control.

The storage unit is envisioned as being more than a very large array of magnetic cores, discs, and/or tapes. We believe that for reasons of storage economy as well as convenience, it will be desirable (and possible) to employ real pictures or even real histologic sections as additional storage media. Thus, a sort of library would be available to the morphologist for reference and instantiating purposes. Obviously, such a storage mechanism places additional requirements on the scanner components. The difficulty of meeting these requirements we do not believe to be insuperable.

Storage of statements would be not in the form of words or sentences, but (probably) in terms of their equivalents in a functional calculus, i.e., the form into which parses (pictorial as well as sentential) are placed by the interpreter and in which form they are manipulated by it.

The interpreter might appear, as so far explained, to be a miscellaneous group of functions, ranging from a translator of parses into formal statements and thence into subroutines, to an input-output control

unit. But, such a hodgepodge of logically unrelated functions would be false to the design principles as we now see them. The interpreter will be a logical whole in which the various functions referred to will be interlaced in a way corresponding to the structures of the grammars, the various levels of discourse, etc. In avoiding the pragmatic piecemeal approach in its construction, we take the immediately difficult approach, which, however, seems to us the approach which in the end will require far less total effort.

We have not specifically indicated a separate display as one of the outputs of the interpreter. Obviously, however, several types of display capabilities will be required. As is the case with the scanner complex, work by others in the area of display systems is progressing at a rapid rate, a rate sufficient to justify optimism with regard to this aspect of our system. The foregoing applies equally well to pointing devices such as light-pens and cursors.

A final note concerning the morphologists who will participate in this hybrid intelligence system. They will find themselves typing in statements, directives and questions in a language as close to their everyday professional dialect as we can make PLACEBO VIII. The language will be "habitable." They will not have to know much if anything about computers. Curiosity, a willingness to explore the bases of taxonomic and diagnostic statements, and a dispassionate attitude toward the display of the consequences thereof—these will be definite assets. We believe that the better the morphologist is as a morphologist, the greater the body of applicable information he brings, and the more articulate he can be about articular relationships, the more useful will his information be to himself within such a system.

The role of the morphologist will change. The investment of time made during the early phases of grammar construction and the outpouring of his knowledge of articular structure will be increasingly rewarded as this new instrument gains steadily in power and usefulness. His role can be expected to be more questioning and more directive as the system grows; his own growth as a morphologist seems almost assured.

Acknowledgment

References

1. COULTER, W. H. 1953. U.S. Patent No. 2,656,508.
2. KIRSCH, R. A., L. CAHN, L. C. RAY & G. H. URBAN. 1957. Experiments in processing pictorial information with a digital computer. Proc. Eastern Joint Computer Conf. Washington, D. C. : 221–229.
3. MOORE, R. T., M. C. STARK & L. CAHN. 1964. Digitizing pictorial information with a precision optical scanner. Prepared for Presentation at the Joint Annual Meeting of the American Society of Photogrammetry and the American Congress on Surveying and Mapping.
4. BOSTROM, R. C. & W. G. HOLCOMB. 1963. CYDAC—A digital scanning cytophotometer. IEEE Intern. Conv. Record. 9: 110–119.
5. MOORE, G. A. 1964. Direct quantitative analysis of photomicrographs by a digital computer. Phot. Sci. Eng. 8(3): 152–161.
6. MINSKY, M. 1961. Steps toward artificial intelligence. Proc. I.R.E. 49(1): 8–30.
7. BOMBA, J. S. 1959. Alpha-numeric character recognition using local operations. Proc. Eastern Joint Computer Conf. Boston, Mass. : 218–224.
8. UNGER, S. H. 1959. Pattern detection and recognition. Proc. I.R.E. 47(10): 1737–1752.
9. NEWELL, A., F. M. TONGE, E. A. FEIGENBAUM, B. F. GREEN, JR. & G. H. MEALY. 1961. Information Processing Language-V Manual. Prentice-Hall, Inc., Englewood Cliffs, N. J.
10. DUNHAM, B. 1957. The formalization of scientific languages. Part I. The work of Woodger and Hull. IBM J. Res. Dev. 1(4): 341–348.
11. KIRSCH, R. A. 1964. Computer interpretation of English text and picture patterns. IEEE Trans. Electronic Computers. 13(4): 363–376.
12. WATT, W. C. 1965. PLACEBO IV rules, concordance, sample computer generation. National Bureau of Standards Technical Note 255.
13. WATT, W. C. 1965 (Forthcoming). Materials for PLACEBO V.
14. WATT, W. C. 1965. A prerequisite to the utility of microgrammars. National Bureau of Standards Technical Note 258.
15. CHOMSKY, N. Formal properties of grammars. In Handbook of Mathematical Psychology. R. D. Luce, R. R. Bush & E. Galanter, Eds. 2(9–14): 323–418. John Wiley & Sons, Inc. New York, N. Y.
16. SUTHERLAND, I. E. 1963. Sketchpad, a man-machine graphical communication system. AFIPS Conf. Proc. Spring Joint Computer Conf. 23: 329–346.