May 1969

LOGICAL ANALYSIS OF PICTURES OF POLYHEDRA

D. A. Huffman

Artificial Intelligence Group

Technical Note No. 6

SRI Project 7494

LOGICAL ANALYSIS OF PICTURES OF POLYHEDRA

D. A. Huffman

Pictures of scenes which contain only plane-bounded solid objects (that is, assortments of polyhedra) are especially attractive for purposes of logical analysis. On the one hand there are arbitrarily many objects of that type since there can be for each object as many vertices as desired and its planar surfaces can meet each other so that in the region of a given vertex the object can be locally convex or concave in a variety of ways. Thus there is a rich and interesting set of objects which can be viewed. On the other hand the number of ways in which individual edges can be viewed and can obscure each other is limited. For these reasons an environment containing polyhedra is a sensible choice as an initial one into which to put a robot which looks at the scene with its television camera "eye" and which moves about the environment, deciding what it has seen and how best to look for what it has not seen.

In spite of the comparative simplicity of an environment of this type, researchers have not seemed to exploit this simplicity satisfactorily. Often, for example, they seem not to know whether a given view of a scene is inherently ambiguous or whether their heuristic analysis methods should be "patched up" yet another time. The object of the study proposed here is to achieve a definitive procedure for the analysis of pictures of sets of complex polyhedra. Such a procedure should be judged by its ability to lead us to the following goals:

1) The identification of two arbitrary regions of a picture with the same or with two different physical objects whenever that is possible.

2) The clear demonstration that ambiguities are present when that is the case.

3) A way of specifying exactly what the ambiguities are.

4) A logical basis for determining what additional camera views are most likely to provide information which will allow the ambiguities to be resolved.

Significant progress has already been made toward these goals. The appendix contains a description of some of the techniques which the author has already generated, together with several simple illustrative examples. The main features of these techniques are summarized briefly here for the reader who cannot take the time to read the appendix.

In the initial phases of the work it has been assumed that exactly three plane surfaces come together at each vertex of the polyhedra. (The techniques for less restrictive assumptions can be shown to be relatively simple extensions of this special case.) An exhaustive listing of all (only four) inherently different vertex types is made, together with an exhaustive listing of the essentially different ways they can be viewed. (Certain views of different vertex types will look the same to the camera even though they "mean" different things.) Similarly, straight lines in a picture can have any of four possible interpretations. Picture analysis progresses as an alternation between decisions about the correct interpretations of lines and about points where lines come together in the picture. Correct interpretation of these lines and points yield correspondingly correct deductions about the true nature of the associated edges and vertices in the actual scene.

The single most important insight which makes this analysis possible is the realization that there are exactly four possible interpretations of a line. It represents either a "convex" or a "concave" edge with both associated planes in view, or it represents a (convex) "hiding" edge which can obscure more distant parts of the scene either to one side of the edge or the other. A given line segment cannot have two different "meanings" in two different parts of the picture. A law of conservation has been found which is applicable to the hiding edges and which is reminiscent of Kirchoff's current law for electrical networks.

The advantages of having a clear and precise understanding of the grammatical rules associated with this kind of picture "language" are analogous to the advantages of understanding the grammatical rules for any language. If part of a picture is missing or garbled or partially obscured it may be possible to reconstruct it if one knows what redundancies are possible in the language. Ambiguities can be better understood and appropriate new pictures of the scene can be taken to resolve them.

An accompanying sheet summarizes the support which is being asked for for the period of one year. A possible sequence in which some of the problems might be studied is the following:

1) An expansion and completion of the picture analysis results for the case of vertices with exactly three associated surfaces.

2) A study of the interrelated algebraic formalisms and data structures most natural to and appropriate for this analysis.

3) An extension of the results to other closely related types of scenes; for example, to an environment with "thin" or shell-like structures with planar surfaces or to polyhedra with more general vertices.

4) A study of similar analyses applied to objects with non-planar bounding surfaces.

5) A study of the additional information available when shadows from one part of a scene are cast onto another part of the scene.

APPENDIX: A Summary of The Main Analysis Tools

The environment is assumed to contain an assortment of solid polyhedra which have exactly three planar surfaces at each of the vertices and, of course, two surfaces associated with each edge. We shall call the actual collection of polyhedra the (3-dimensional) scene; the projected (2-dimensional) view which the camera sees will be called the picture. The visible edges and vertices of the scene are associated with the lines and points, respectively, of the picture. On each side of each line in a picture is an area which may or may not be associated with one of the surfaces of the same polydedron associated with the line. In any case an area of a picture is bounded by lines and is minimum in the sense that a given area, by definition, has no lines passing through it.

We shall use the term "picture" to refer not only to projected views of possible 3-dimensional scenes but also to 2-dimensional line drawings which purport to be views of scenes but for which there may be no corresponding physically realizable set of polyhedra. Such pictures of "impossible objects" furnish a novel way of providing instructive tests for analysis procedures. They may also come about naturally if, for example, because of the lighting conditions an edge in the scene generates no corresponding line in the picture.

There are four basic ways in which three plane surfaces can come together at a vertex. All four can be illustrated by the picture of a fireplace and raised hearth shown in Fig. 1-a. In Fig. 1-b, pictures of these four vertex types are shown associated with four abstract scenes. The identifying numbers are determined in accordance with the following reasoning. The three planes which meet at a vertex partition the surrounding space into eight octants. (Even if the three planes and associated three edges are not mutually orthogonal the term "octant" will be used and the general comments offered here will nevertheless apply.) The number of octants occupied by solid material at the vertex is chosen as the type-number for the vertex. The "T-points" are not associated with physical vertices and will be commented on separately later.

A vertex can be viewed from any one of the octants which is not occupied by solid material and all views from a given octant give essentially the same "configuration". (The exact meaning of this comment will be apparent later.) For instance, a type-3 vertex can be viewed from the complementary five octants in which the eye or camera may be placed. Except in the case of a type-1 vertex, where certain rotational symmetries reduce the number of possibilities, the view from each of the octants gives a different way in which the visible and (if there are any) invisible edges meet.

The possible views of each of the four types of vertices are summarized in Fig. 2. Note that although there are always exactly three edges incident at a vertex only two may actually be visible as lines in the pictorial view; the third edge is, in these cases, associated with a line which is hidden below one surface of the polyhedron. Such edges are shown dotted in the figures.

We shall also assume, for the purposes of these notes, that each picture is taken from a "general viewpoint". That is, that the camera or eye of the viewer is not in the same plane as any of the surfaces represented in the picture. Equivalently, we assume that a small change in the vantage-point of the camera or eye of the viewer would not cause the lines in the picture to come together in a basically different configuration. (This comment, also, will become quite apparent once we have listed the possible configurations.)

The lines in each of the views of the vertices are labelled as follows:

(i) a "+" line represents an edge for which the two corresponding planes bound the solid material in such a way that a "convex" edge results

(ii) a "-" line represents an edge for which the two corresponding planes bound the solid material in such a way that a "concave" edge results.

The other two types of lines are associated with convex edges but have associated planes which are both on the same side of the edge as viewed by the camera. An arrowhead is used as the label for such lines with the convention that as one follows in the direction indicated by the arrowhead the pair of associated planes is on the right. Since an arrowhead may point toward or away from the nearby vertex we can label lines with

(iii) the "to-arrow" or

(iv) the "from-arrow".

The notations +, - and → for lines of a picture are especially convenient since, as seen from a given vantage point, a line in a picture of a real scene is of fixed type throughout its length. (Otherwise the associated pair of planes would have different orientations in different parts of the scene.) Note especially that an "arrowed" line which goes from point "a" to point "b" is labelled with a from-arrow as far as "a" is concerned but is labelled with a to-arrow as far as "b" is concerned (see Fig. 3). The notation for these lines has been chosen carefully so that whichever one of the adjacent pair of points is thought of as reference the pair of planes associated with the edge are represented to the right of the arrowed line in the picture.

The convention for hidden (dotted) lines in a picture is the same as for exposed lines. Imagine for such a line that the plane between it and the viewer is removed and note the orientations of the pair of planes associated with the line. If they are both on one side of the line it is an "arrowed" line with one orientation on the other. If not, the line can be clearly identified as "+" or "-". Examples of all four types of dotted lines are shown in Fig. 2.

All visible edges of a scene correspond in a picture to lines which can be thought of as having zero <u>depth</u>. The dotted lines in Fig. 2 have unit depth. In general the depth-index for a line is the number of surfaces which would have to be removed to expose the corresponding edge in the scene. Each (visible or invisible) line in a picture is of one of four <u>types</u>. It may have a depth index which is quite large, especially for complex convoluted polyhedra. We shall, henceforth, in these notes deal only with visible, or zero-depth lines. Thus we are giving here only a surface analysis. A deeper analysis will not be necessary for our present purposes.

From Fig. 2 it can be seen that the arrowed lines in each of the twelve possible vertex views satisfy a "conservation rule": the number of arrows into and out of a point are equal. (This result is also true of points which represent vertices with more than three associated surfaces and it applies as well to pictures with non-planar surfaces.)

An additional kind of point can occur in a picture which is associated with no corresponding vertex in the scene. These are the "T-points" noted earlier. Each gives direct evidence that the bar of the "T" is an arrowed line in which the direction of the arrow is from right-to-left (when the "T" is in the standard upright position). The line type for the obscured line may be "+" or "-" or arrowed (in either direction).

A listing of all of the line structures which are possible at points in a picture is given in Fig. 5. Note that it is convenient to call each either a "V", "W" or "Y". The integer shown is the vertex-type.

As an example of the use of the list in Fig. 5 for the analysis of a picture consider the object of Fig. 6. Since one can see the entire "object" we know that the entire periphery consists of arrowed lines proceeding clockwise. Furthermore, the T-point also gives evidence of arrowed-lines. These have been shown. In Fig. 6-b are shown additional line types which are deduced from observing that some of the points on the periphery are "W". The list tells us that "W's" with arrows on the outer lines must have center lines which are the "+" type. Figs. 6-c through 6-f show that a unique and consistent labelling of lines with their types is possible for this picture. The labelling itself is an aid to us in picturing what the object is: a rectangular parallelopiped with a cavity which is a skewed parallelopiped.

Another example of a picture to be analyzed is given in Fig. 7-a. The peripheral lines have already been labelled with arrows, which are shown. Fig. 7-b has additional labels which, according to the listing of Fig. 5 must apply. The result, however, is that one "V" point which is interior to the picture has its two lines both labelled "+". This is not one of the possible configurations at a point. Thus the picture is of an impossible object. (Impossible, that is, in the realm of solid objects with vertices at which exactly three planes meet.)

The type of analysis we have demonstrated is based on a set of conditions which, given our assumptions about the scene, are necessary in the associated picture. These conditions are, in general, not sufficient to guarantee that a picture will have a possible corresponding 3-dimensional scene. For instance, the edge labels for the "object" given in Fig. 8 can be determined quickly and are mutually consistent. Nevertheless there is no such plane-bounded object. This can be seen easily by noting that the two areas identified as A and B can have only one (possibly interrupted) common line. Yet the picture gives evidence that there are two.

Other examples of "projective-geometric" kinds of constraints are known and their close relationship to the line-labelling techniques illustrated here is being further investigated. Much is already known, but these results would be outside the scope of this appendix.

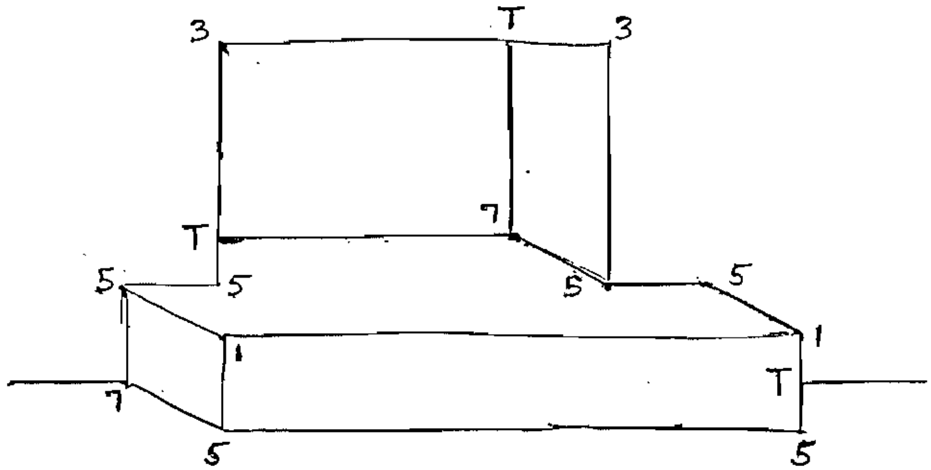


Fig. 1-a: A fireplace

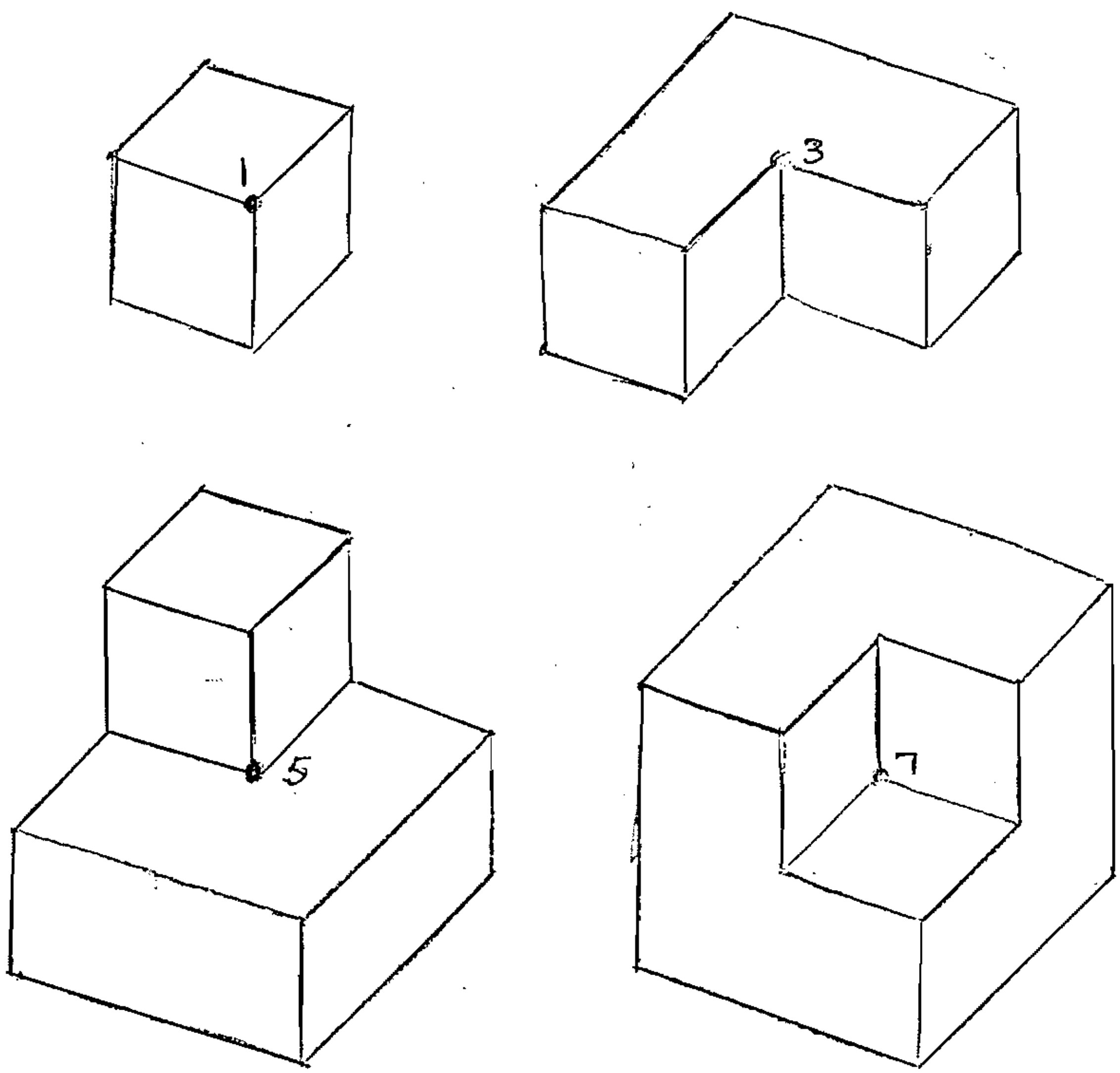
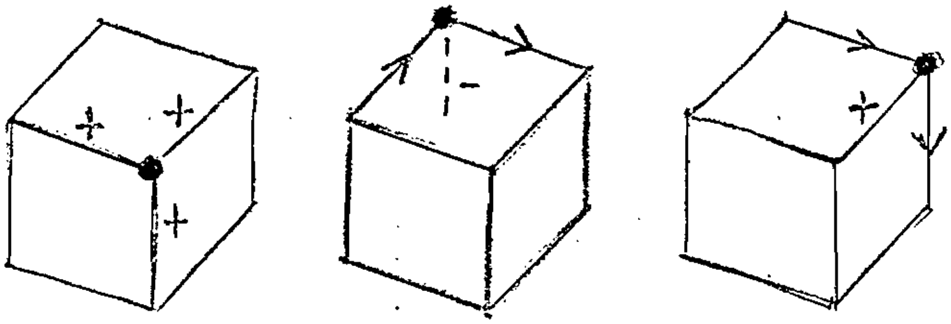


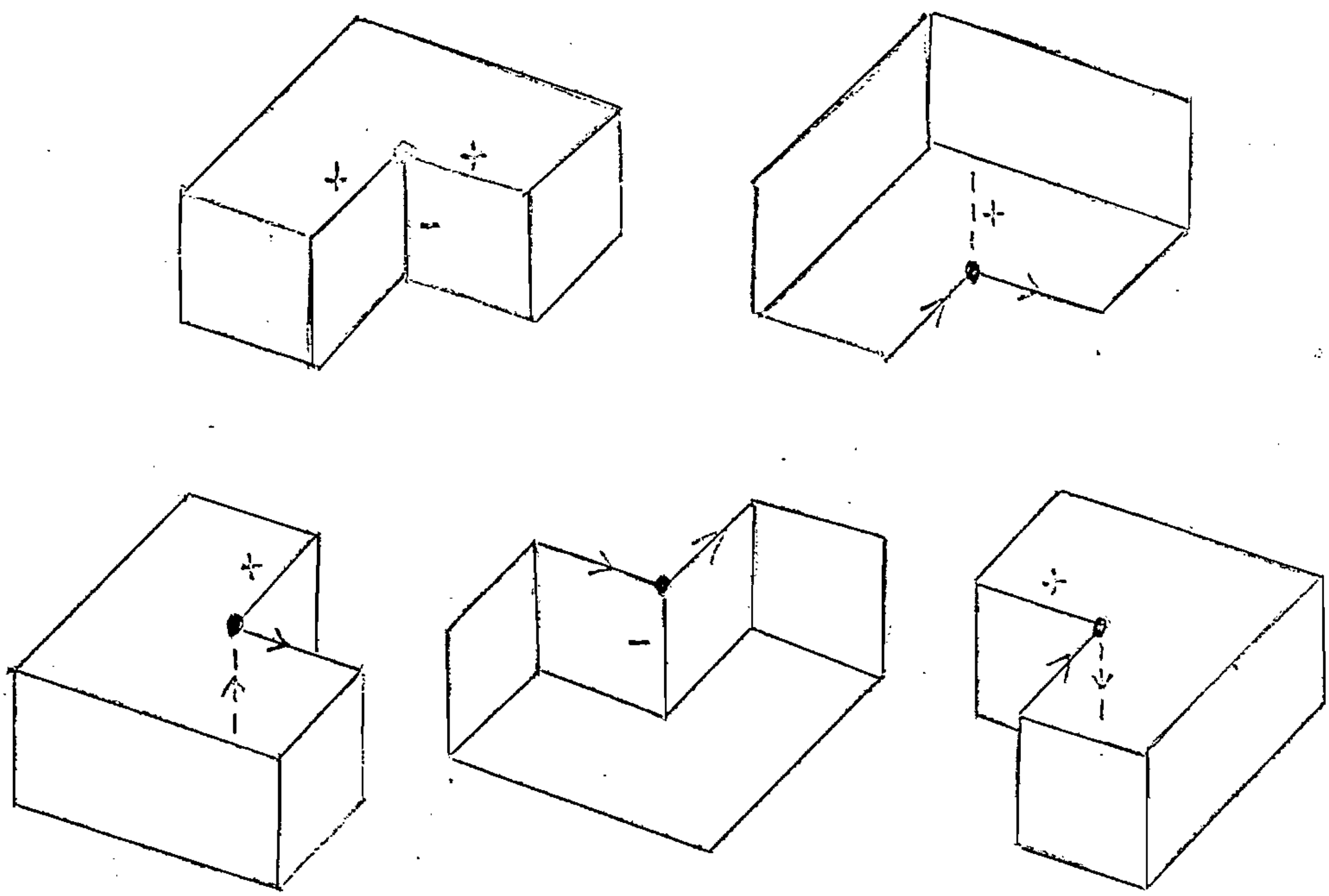
Fig. 1-b

The four vertex types.

Type 1:

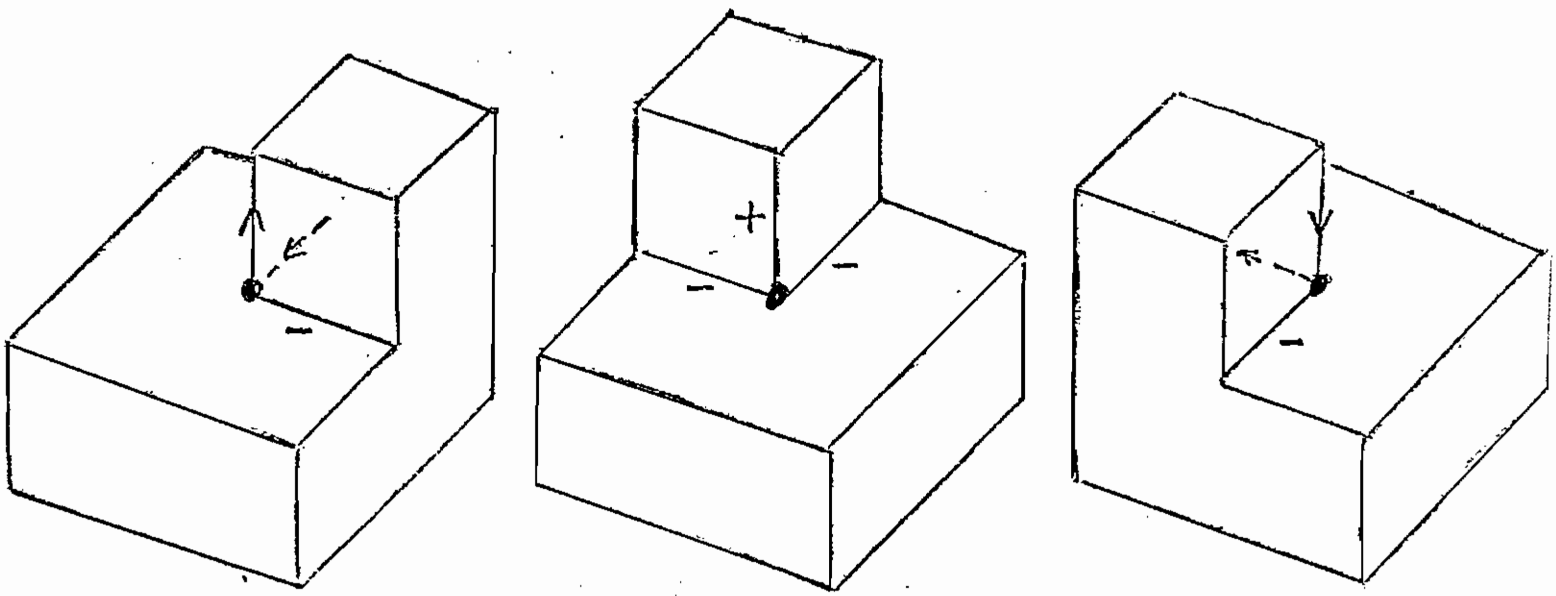


Type 3:



(Fig. 2)

Type 5:



Type 7:

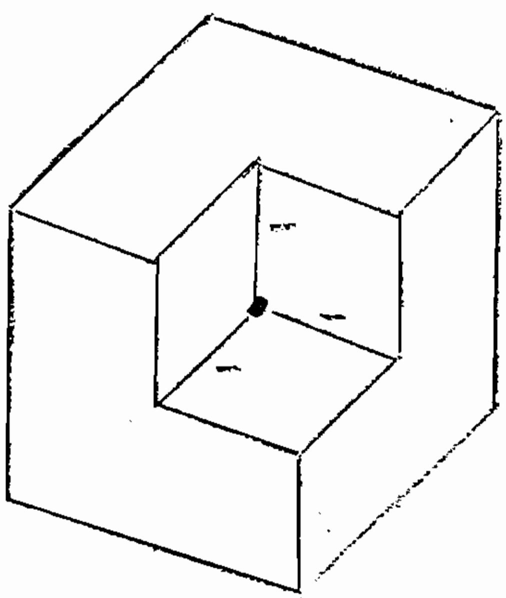


Fig. 2

Views of vertices.

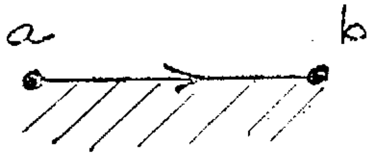


Fig. 3

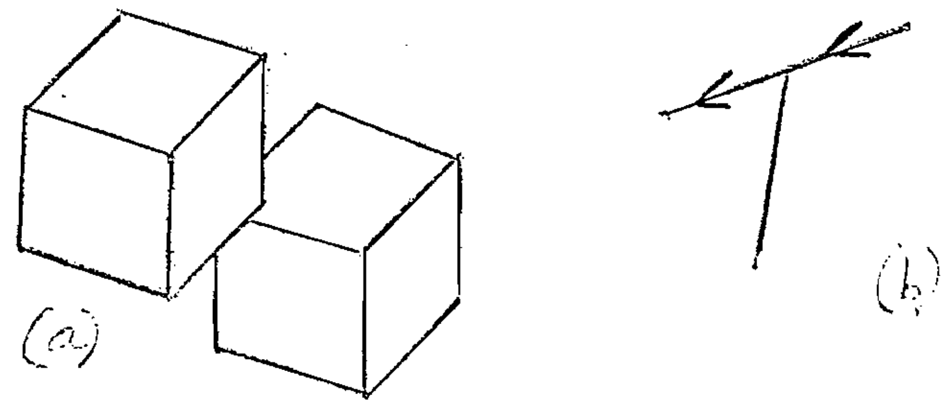


Fig. 4

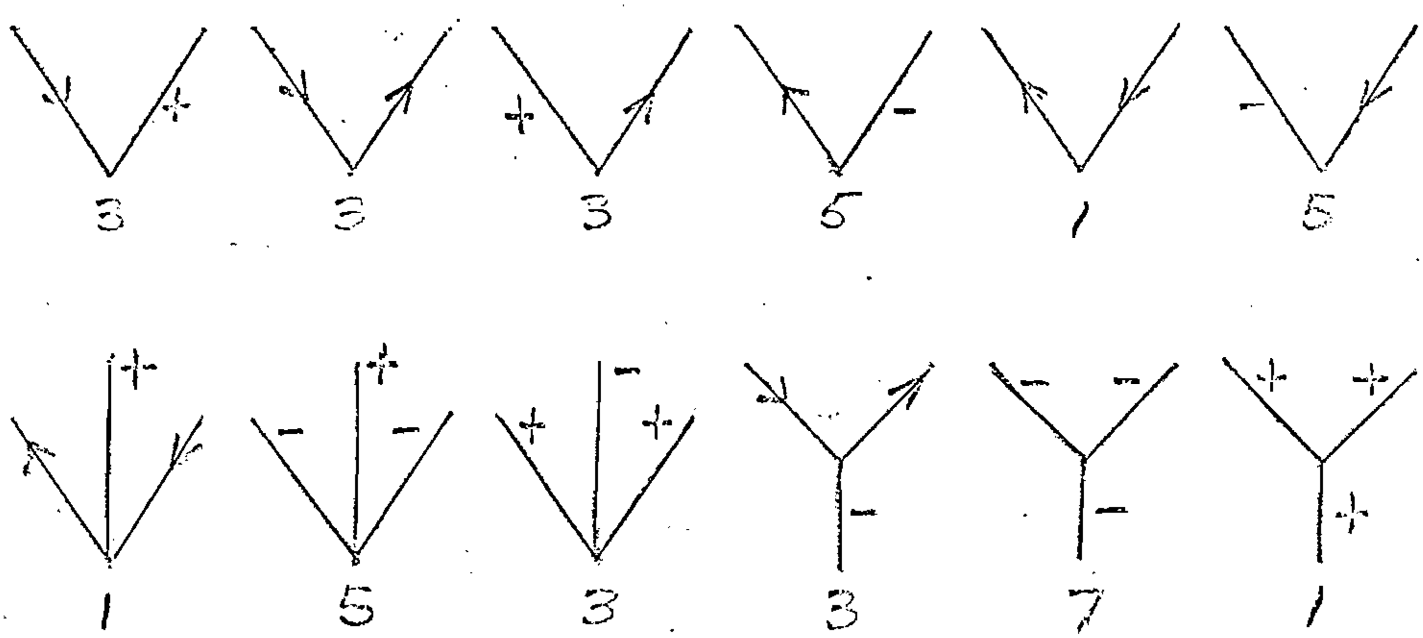


Fig. 5

A list of points with possible associated line configurations.

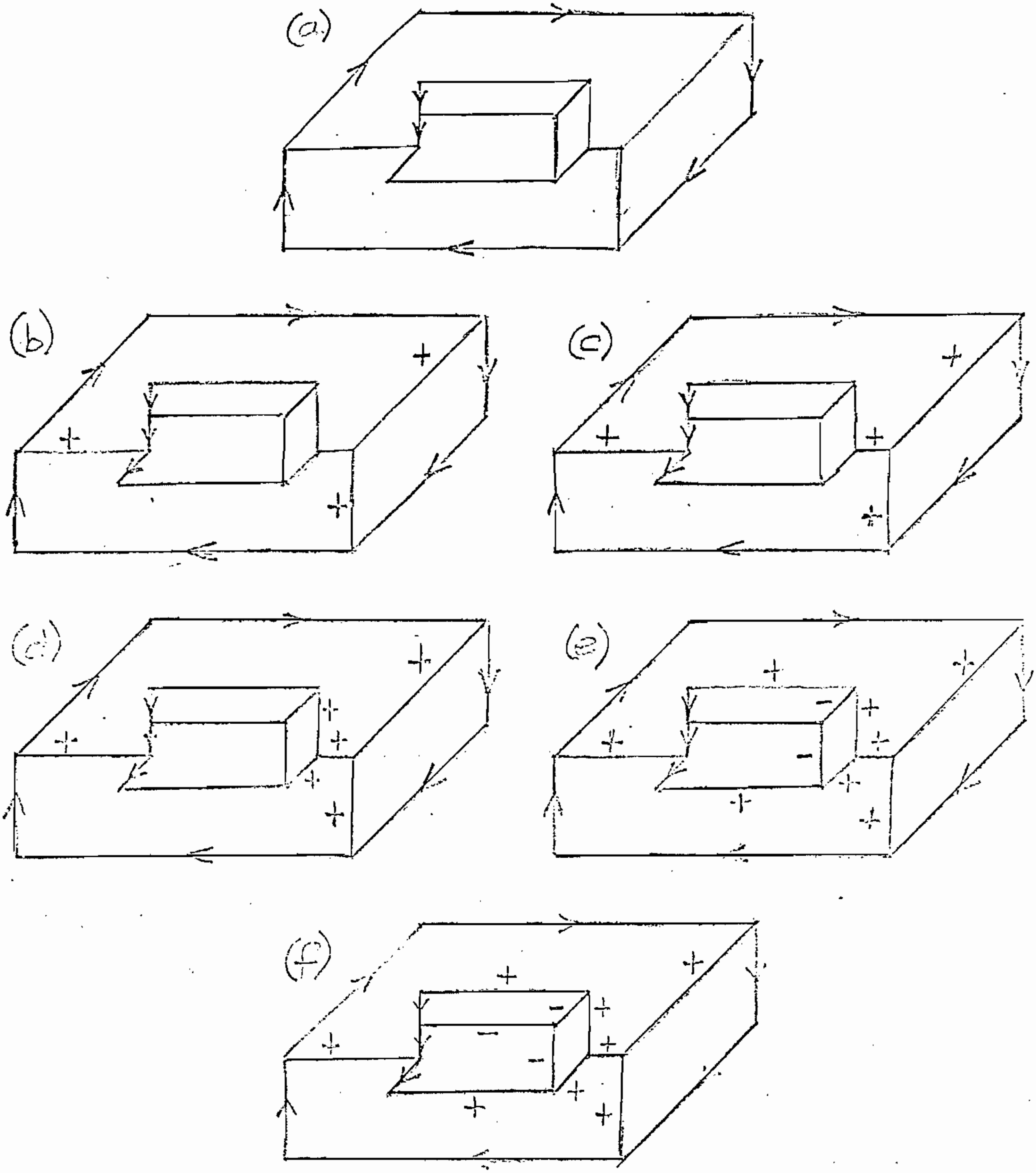


Fig. 6

Analysis of an unusual polyhedron.

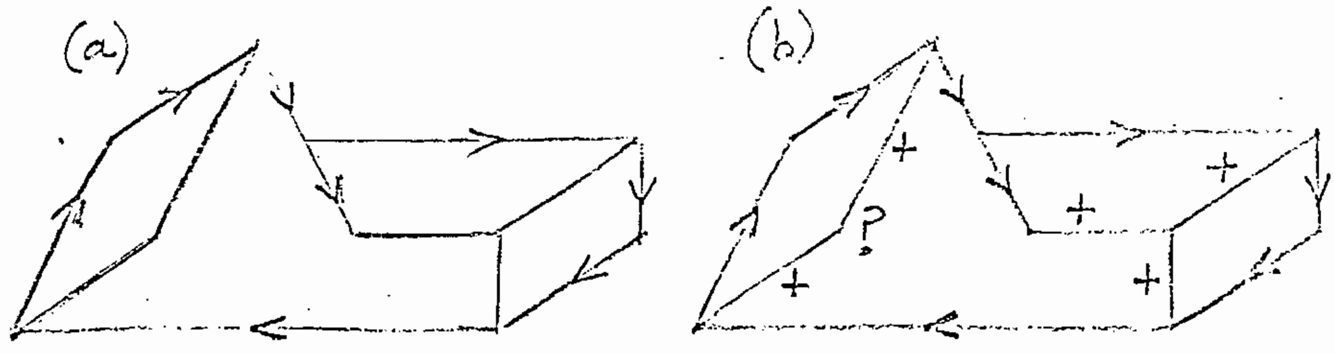


Fig. 7

Analysis of an "impossible object"

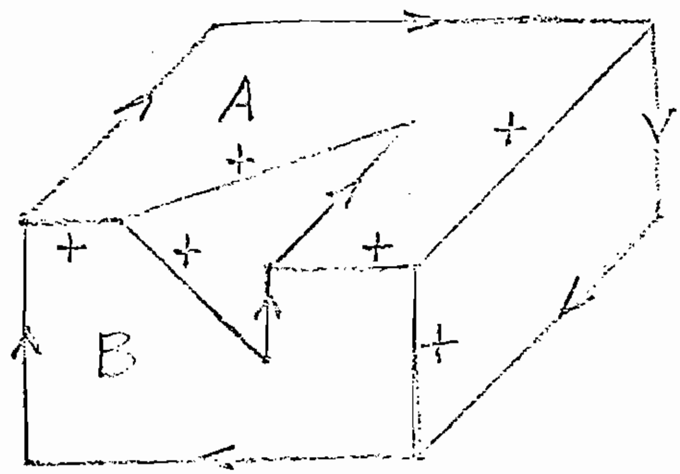


Fig. 8

An impossible object the lines of
which can be properly labelled.