Miller, W.F., & Shaw, A.C. (1968) Linguistic methods in picture processing – a survey. *Proc. 1968 Fall Joint comput. Conf.*, pp. 279–90.
Rosenfeld, A. & Pfaltz, J.L. (1966) Sequential operations in digital picture processing. *J. Ass. comput. Mach.*, **13**, 471–94.

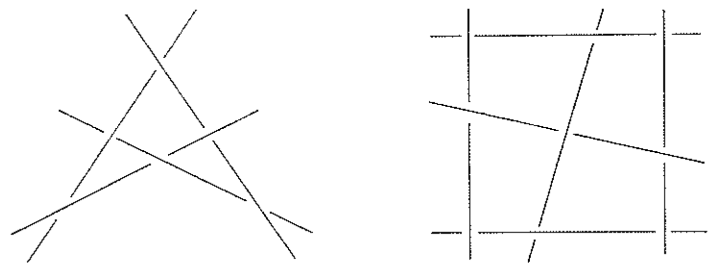
19

Impossible Objects as Nonsense Sentences

D. A. Huffman
Board of Studies in Information and Computer Science
University of California at Santa Cruz

INTRODUCTION

To every 3-dimensional scene there correspond as many 2-dimensional pictures as there are possible vantage points for the camera. It is, however, possible to construct pictures for which there is no corresponding scene containing physically-realizable objects. Pictures of such 'impossible objects' can be useful in giving insight into the constraints or grammatical rules associated with the 'language' of pictures, just as nonsense sentences can be useful in illustrating the rules of other languages. Impossible objects have been used by psychologists (Penrose and Penrose 1958) to create visual illusions which successfully challenge the ability of our perceptual systems to synthesize a 3-dimensional world from 2-dimensional information. The incompatibilities among the various portions of pictures of these objects are a novel way of testing our picture analysis procedures. The purpose of this paper is to demonstrate some possible decision procedures and to test them on pictures of both possible and impossible objects.



(a) (b)

Figure 1. Two pictures of sets of line-segments

Examples of pictures of objects for study are shown in figures 1, 2 and 3. The first of these are from a paper by the author (Huffman 1968) which dealt with objects consisting of straight line segments only. It was assumed that straight lines in the pictures were representative of straight-line segments in the 3-dimensional scene; the sense of 'passing' at the various intersections was also indicated. With these assumptions the object shown in figure 1(a) is not possible; the one in figure 1(b) is.

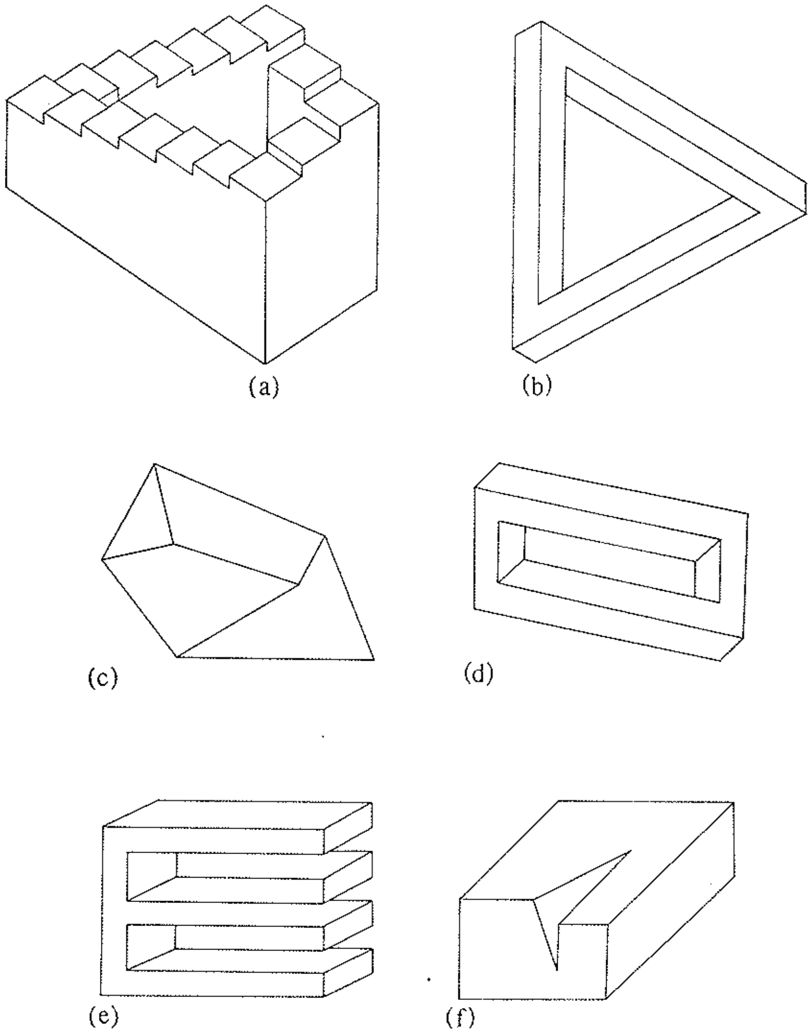


Figure 2. Examples of impossible polyhedra

The pictures in figure 2 purport to be of plane-bounded polyhedra. [The first two of these are adapted from Penrose and Penrose (1958)]. It is primarily with such polyhedral objects that this paper will deal. Under certain natural assumptions none of the objects in figure 2 is possible.

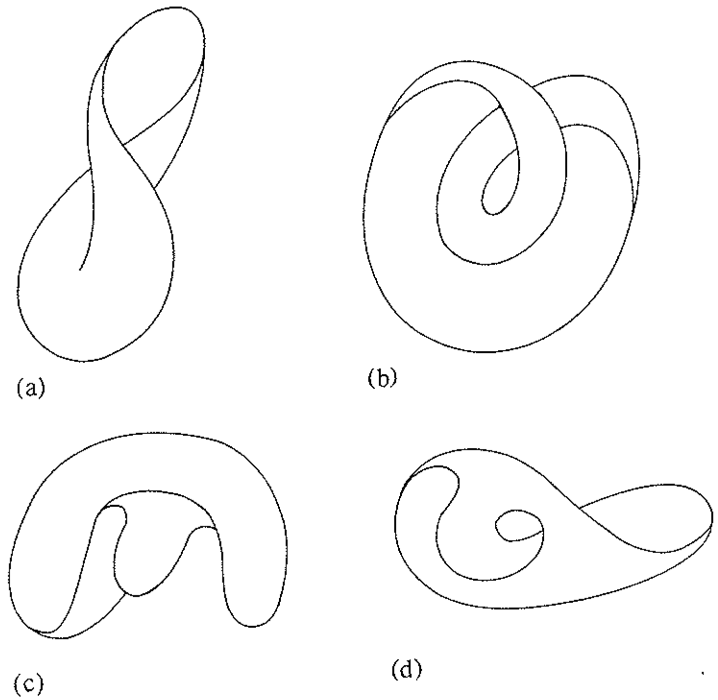


Figure 3. Examples of 'smooth' objects

In figure 3 are shown pictures of four different 'smooth' objects. One of these is impossible; the other three are possible. The constraints which must be taken into account in making decisions about such objects will be the topic of the final sections of this paper.

Of course the possibility or impossibility of the objects in any picture depends upon what assumptions we make about the physical nature of the 3-dimensional elements which we can, in the picture, view in only two dimensions. Typical assumptions we *could* make are: that what appear to be straight lines are actually straight, that lines which appear to be parallel are parallel, that objects are either 'thin' or have appreciable thickness, that all lines which appear to meet at a point actually meet at that point, that all edges which are on the side of the object toward the camera actually are represented in the picture, and so forth.

One assumption we shall make throughout this paper is that all pictures are taken from a 'general position'; that is, that a slight change of the position from which the picture is taken would not change the number of lines in the picture or the configurations in which they come together. In the case of pictures of polyhedra this eliminates the possibility of pictures in which two vertices of the objects in the scene are, by coincidence, represented at the same point in the picture, or two edges in the scene are seen as a single line in the picture, or a vertex is seen exactly in line with an unrelated edge.

The assumption of a general position for the camera is a practical one since it reduces the number of types of local configurations of lines which we must deal with in the picture. Furthermore, if this assumption leads us to judge as impossible an object or set of objects which we know to exist (and therefore by definition 'possible') we can conclude that the camera was probably not in a general position (or that some other assumption was unjustified). In that case we can either move the camera slightly and retake the picture, or go to an augmented list of local configurations which are possible and reanalyze the picture accordingly.

The assumption of a general position for the camera, strictly interpreted, would prevent us from considering pictures having several line-segments which, if extended, met at a common point unless the corresponding edges in the scene would also meet at a common vertex. A special case of this issue arises if several line-segments in a picture are mutually parallel (and would therefore meet at a point-at-infinity if extended). Because it is annoying to have to concern ourselves about this issue in drawing pictures to be analyzed we shall make whatever special commentary is necessary if several line-segments in a picture would meet if extended.

DERIVATION OF LOCAL CONSTRAINTS FOR PICTURES OF DEGREE-3 POLYHEDRA

Introductory Remarks

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. Other authors [for example, Guzman (1968)] have also considered the problem of the analysis of scenes containing polyhedra.

In the work reported here no account has been taken of the effects of gravity or of the possible mutual interaction effects among various members of a set of polyhedra. If we did not make this assumption we could soon find ourselves immersed in such (for us, secondary) problems as determining angles of repose for piles of tetrahedra, or of deriving the result that a stack of identical unit cubes is just barely stable if the amount by which the nth cube down in the stack extends beyond the one below it is equal to $1/n$.

We shall assume 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 will progress as an alternation between decisions about the correct interpretations of lines and of the points where sets of lines are incident in the picture. Correct interpretations of these lines and points yield correspondingly correct deductions about the true nature of the associated edges and vertices of the objects 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 in a picture. It can represent either a 'convex' or a 'concave' edge with both associated planes in view, or it can represent 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 and it is this constancy of interpretation of what a line must mean in a picture which is the key to the method of analysis presented here.

Definition of terms

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 straight *lines* (actually, line-segments) and *nodes*, respectively, of the picture. These lines and nodes determine a picture which is actually a graph, which we may refer to, from time to time, as the *picture-graph*. 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 polyhedron associated with the line.

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' may be constructed with a view to exercising our analysis procedures in a special way. They may also come about naturally if, for example, an edge in the scene generates no corresponding line in the picture because of the lighting conditions under which the picture was taken.

Types of vertices and edges in the scene

There are four basic ways in which three plane surfaces can come together at a vertex. All four can be illustrated in the picture of a fireplace and raised hearth shown in figure 4(a). In figure 4(b), pictures of these four vertex

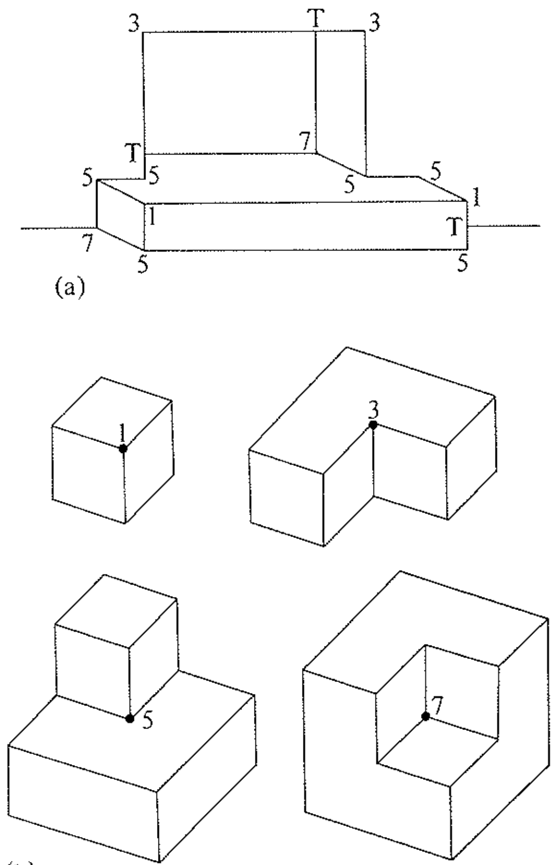


Figure 4. Illustrating the four types of vertices: (a) a fireplace and hearth; (b) the four vertex types

types are shown associated with pictures of four polyhedra. 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 which are occupied by solid material at the vertex is chosen as the type-number for the vertex. (The 'T-nodes' are not associated with physical vertices and will be commented on separately later.)

There are only two types of edges possible for our plane-bounded polyhedra: concave and convex. Each of the four vertex types we have allowed is uniquely associated with a trio of edge-types. The edges associated with type-1 vertices are all convex; those associated with type-7 vertices are all concave. Two of the edges incident at a type-3 vertex are convex and the other is concave. Two of the edges incident at a type-5 vertex are concave and the other is convex.

Types of nodes and lines in the picture

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 rotational symmetries reduce the number of possibilities, the view from each of the octants gives an essentially different way in which the edges can meet.

The possible views of each of the four types of vertices are summarized in figure 5. Note that although there are always exactly three edges incident at a vertex, only two may actually be visible as lines in the picture; the third line is, in these cases, associated with an edge which is hidden (from the camera position). Such lines are shown dotted in the figures.

The (undotted) lines in each of the pictures of the vertices are labelled as follows:

(a) a '+' line represents a convex edge which has both of its corresponding planes visible from the camera;

(b) a '−' line represents a concave edge which has both of its corresponding planes visible from the camera;

The other two types of label for picture lines correspond to those convex edges in the scene which have both of their associated planes on the same side of the edge as viewed from the camera, one hiding the other. An arrow is used as the label for such lines with the convention that, as one moves in the direction indicated by the arrow, the pair of associated planes is *to the right*. Since an arrow may lead toward or away from a given node of the picture we conclude that, with respect to that node, we can label an incident line with either

(c) an 'in' arrow, or

(d) an 'out' arrow,

as well as with the '+' or '−' label.

It is clear that a line in a picture of a 'possible' scene must have a *single* one of the four labels associated with it along its entire length. (Otherwise the pair of planes associated with it would have different orientations in different parts of the scene.)

Our convention for labelling dotted lines (representing edges hidden from the camera) is consistent with the one for labelling lines representing visible edges. In order to determine the label for a line corresponding to a hidden edge, imagine that the surfaces between that edge and the camera are removed, and note the orientations of the pair of planes associated with the edge. The label which is to be assigned to the picture line is the same as if the intermediate surfaces had not been present. Examples of all four types of dotted lines are shown in figure 5.

Note that a hidden *line* may be labelled '+', even though the corresponding *edge* is concave, and that a hidden line may be labelled '−', even though the corresponding edge is convex. Similarly, hidden lines which are labelled with arrows may be associated with either concave or convex edges. (In each of these cases it is the parity of the number of surfaces between the camera and the edge which is the controlling factor.)

All visible edges of a scene correspond in a picture to lines which can be thought of as having zero *depth*. The dotted lines in figure 5 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 line in a picture has not only one of four *types*; it also has a non-negative depth index. We shall, in general, in this paper deal only with exposed, or zero-depth, lines. Thus we are giving here only a 'surface' analysis. A 'deeper' analysis will not be necessary for our present purpose.

From figure 5 it can be seen that the arrowed lines (both dotted and undotted) in each of the twelve possible vertex views satisfy a 'continuity' or 'conservation' rule: the number of arrows into and out of a node are equal. (This result is also true of nodes which represent vertices having more than three associated surfaces and it applies as well to pictures representing non-planar surfaces.)

In a picture an additional kind of node can occur which is associated with *no* corresponding vertex in the scene. These are the 'T-nodes' 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 other labelled line

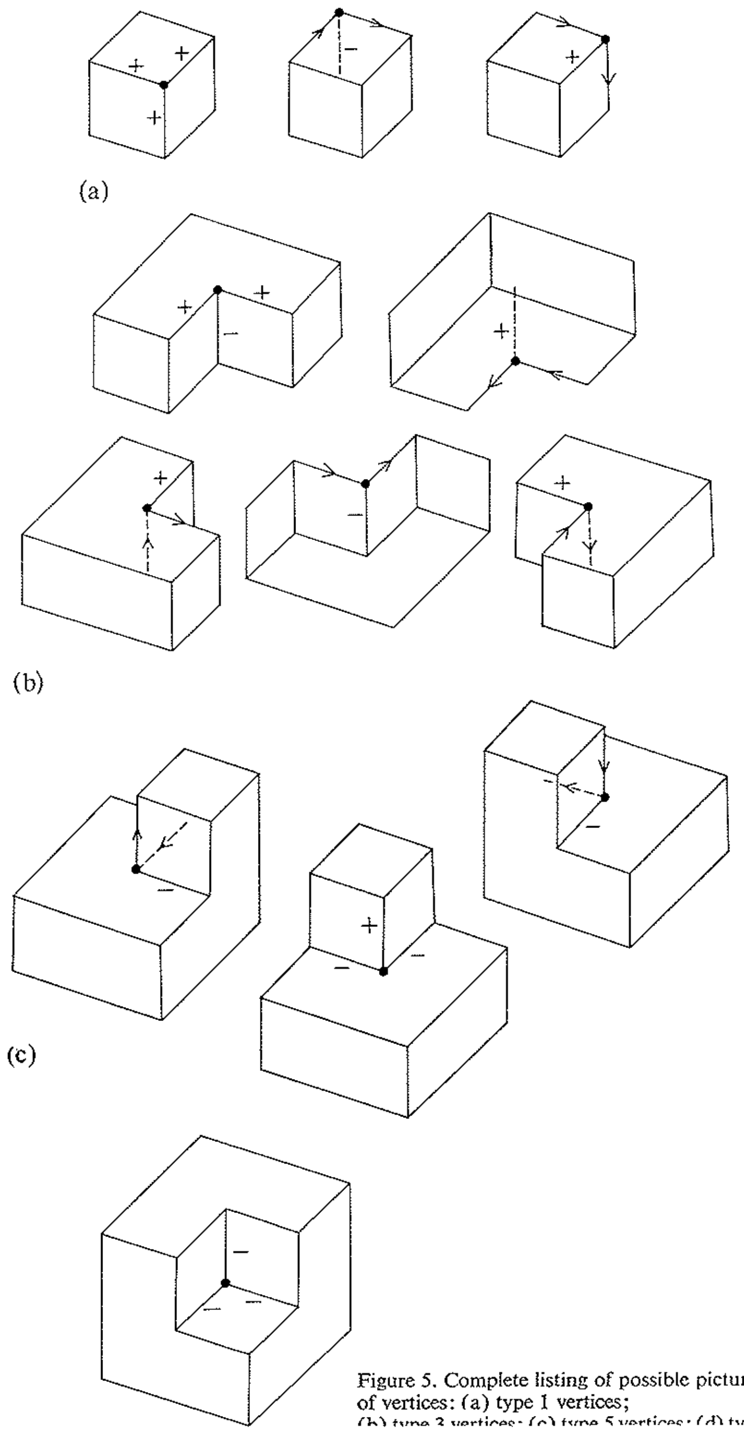


Figure 5. Complete listing of possible pictures of vertices: (a) type 1 vertices; (b) type 3 vertices; (c) type 5 vertices; (d) type

configurations possible in the vicinity of a node in a picture-graph is given in figure 6. It is convenient to call each either a 'V', 'W', or 'Y'. (Each of these configurations may, of course, appear with an arbitrary rotation in a given picture.) The integer shown is the associated vertex-type. Note that each of these three types of configuration around a node of the picture-graph can have several different 'meanings' in terms of the (3-dimensional) vertex-types it can represent. The proper interpretation of the configuration of lines incident at a given node cannot be determined unless the surrounding 'context' is taken into consideration.

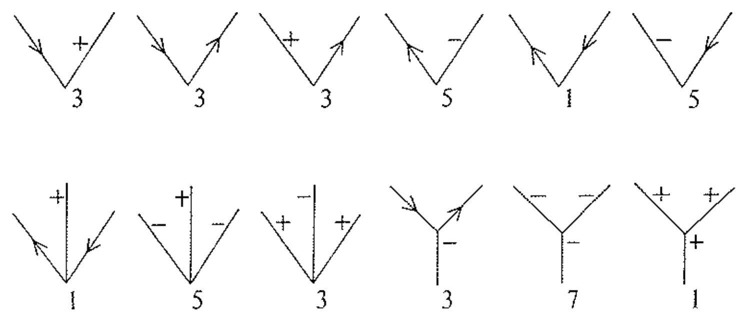


Figure 6. Possible labelled-line configurations around a picture node

THE PROBLEM OF PICTURE ANALYSIS

The mapping from scene to picture

Imagine a (3-dimensional) scene containing one or more polyhedra. In the process of producing a corresponding (2-dimensional) picture-graph we can imagine several different stages, at each of which some information is lost. For example, consider a rectangular parallelopiped through which a hole with a cross-section which is a parallelogram is completely drilled. A close approximation to a true representation of this object is given in the *labelled 'X-ray' picture* shown in figure 7(a). (For completeness a depth-index for all hidden lines and labels for the nodes could also be given but they would only unnecessarily clutter the picture here.) Some information about the actual dimensions of the object has already been lost, unavoidably, particularly information about the distances in the direction of the rays which pass through the camera position. This lost information is associated with the now-missing 'third' dimension.

If no indication is given of hidden lines we obtain the *labelled picture-graph* in figure 7(b). (The labels for the nodes of a picture-graph of a polyhedron with degree-3 vertices are uniquely determined by the line-labels; *see* figure 6. Thus these labels would be redundant.) Here we have irretrievably lost even more information about the real object since there is an infinite

variety of hidden protrusions, cavities and obscured objects which would not be apparent in this type of picture. At this stage there is an inherent ambiguity about the hidden parts of the object. For example, in our example we have no evidence about how far the hole penetrates into the object. Finally, by dropping the line-labels we obtain the *unlabelled picture-graph* itself [figure (7c)].

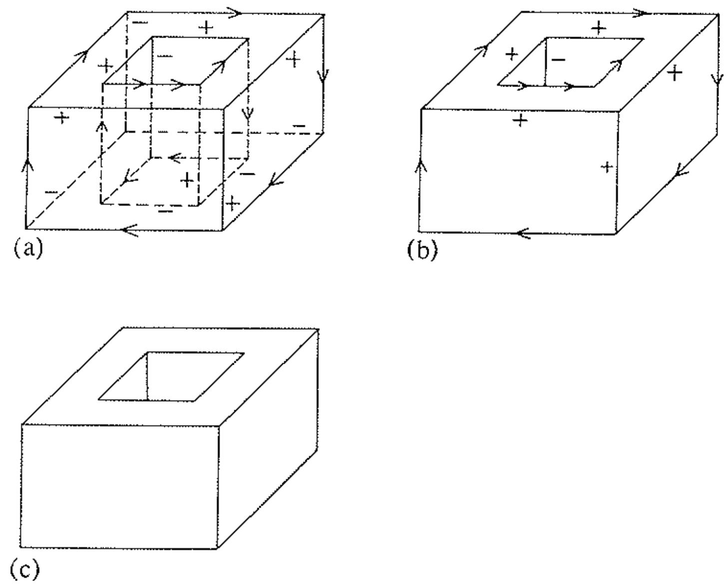


Figure 7. Illustrating three kinds of pictures: (a) labelled 'X-ray' picture-graph; (b) labelled picture-graph; (c) unlabelled picture-graph

Reconstruction of a scene from a picture

Certainly a necessary condition for the realizability of an unlabelled picture-graph as a scene containing polyhedra is that there exists some labelling which places exactly one of the four possible line-labels on each line so that around each node of the graph there exists just one of the allowable configurations given in figure 6. (The 'T-nodes' are exceptions which have already been noted.) The successful labelling of the lines around each node in a picture uniquely corresponds to an identification of the types of vertices and edges in the associated scene. It is also clear that the two picture areas on either side of a line which has been labelled with a '+' or a '−' must belong to the same object.

The labelling process for the lines and nodes of an unlabelled picture is analogous to the parsing of sentences in other languages. In our language the 'alphabet' consists of four types ('T', 'V', 'W', and 'Y') of nodes. In order to be 'syntactically' correct a picture containing these symbols must have them interconnected only in certain allowed ways. As we have noted above, even a properly-labelled picture can have a wide variety of interpretations. A totally-satisfactory analysis procedure would allow us to specify in detail at least one labelled 'X-ray' picture from which a corresponding object or set of objects could be constructed.

Forbidden picture-subgraphs

The syntactic rules for any language impose constraints on the manner in which the symbols of its alphabet may be related to each other. This is certainly true of our pictures of polyhedra. Some forbidden combinations are shown in figure 8. The lines of none of them can be consistently labelled. (The reader should verify this fact for himself.) The configuration consisting of two 'T-nodes' [figure 8(a)] is worthy of special notice since it is forbidden

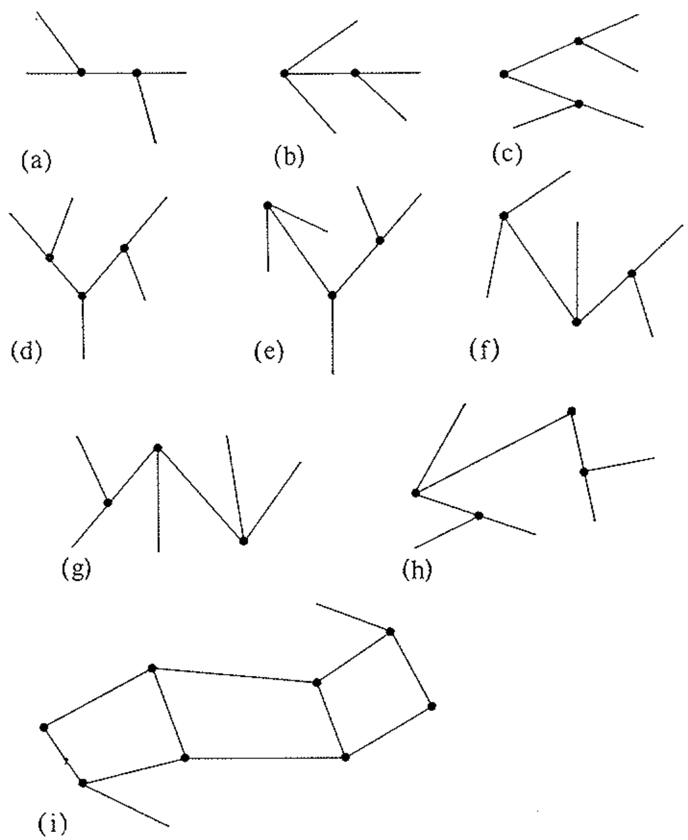


Figure 8. Examples of forbidden subpictures

even in pictures of polyhedra which may have an arbitrary number of edges associated with a vertex.

Pictures which cannot be labelled

In figure 9 eight examples of pictures of objects are given which are impossible to realize as degree-3 polyhedra. In these examples each of the lines can be labelled so as to satisfy the necessary constraints at all but one node. The identity of the node (in general, nodes) at which the labelling is not an allowed one may depend upon the order in which the labels are derived. Our primary concern is with whether or not a labelling is possible for the picture as a whole rather than with where the labelling fails. Often, however, the

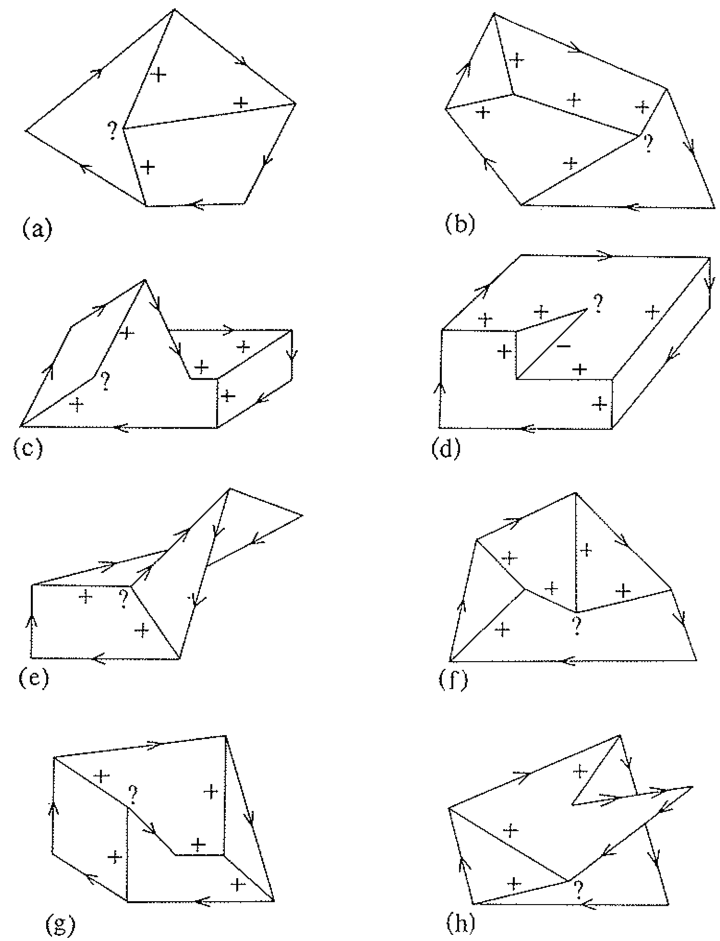


Figure 9. Pictures of objects for which there is no labelling possible

location of a picture region where labelling is impossible will be where a human subject reports that the picture 'doesn't look right' to him.

We have assumed that each of the eight pictures in figure 9 is of a complete object rather than, for example, a recess in the surface of a larger object. This assumption allows us to label each of the lines on the periphery of the picture with an arrow pointing clockwise. Other lines incident on nodes on the periphery can be labelled '+' or '−' depending on whether the arrows on the periphery take a right or left turn at those nodes (see the appropriate W and Y configurations in figure 6). Additional lines which can be labelled immediately with an arrow are those associated with a T-node. The remaining lines are labelled so that, if possible, an allowable configuration exists at each node. In this labelling procedure it may be necessary to assume a given label for a line in order to make further progress and to revise this assumption if it leads to a stage at which further labelling is impossible.

Pictures of 'unlikely' objects

The pictures in figure 10 can all be labelled consistently (in a unique way) and each describes an object which would probably be 'unlikely' to be seen by most of us. The objects in figures 10(a) and (b) can be visualized more easily if one imagines that a cavity in the form of a skew parallelopiped is removed from the more familiar part of each object pictured. In these two pictures we have violated the condition of a general position for the camera. If the camera position were moved somewhat certain edges of the cavities would no longer appear parallel to other edges of the object.

The object in figure 10(c) is in the general position. However, if the camera were moved a significant amount to the right of the object it would record the hidden construction on the back side of the object which was necessary to make the object 'possible'. The object in figure 10(d) can be realized by combining two tetrahedra.

The 'skew cube' of figure 10(e) looks impossible to many human subjects but it is realizable regardless of the exact orientation of the lines in the picture. In order to prove this we can imagine the construction of the object itself. First, orient three plane surfaces in any of the (infinite number of) ways in which they meet so that the edges associated with their pairwise intersections appear to the camera to be the three lines at the middle of the picture. Next, locate two more edges in each of these three surfaces so that the set of six appear to the camera to be in the positions of the six peripheral lines of the picture. Each of the vertices corresponding to the three 'W' configurations has associated with it a pair of edges determining uniquely the location of one of the three hidden surfaces of the object and the point common to these three plane surfaces is also uniquely determined.

If this point would appear to the camera to be inside the periphery of the object in the picture the object can be constructed with the six surfaces referred to above. If this point would appear to the camera to be outside

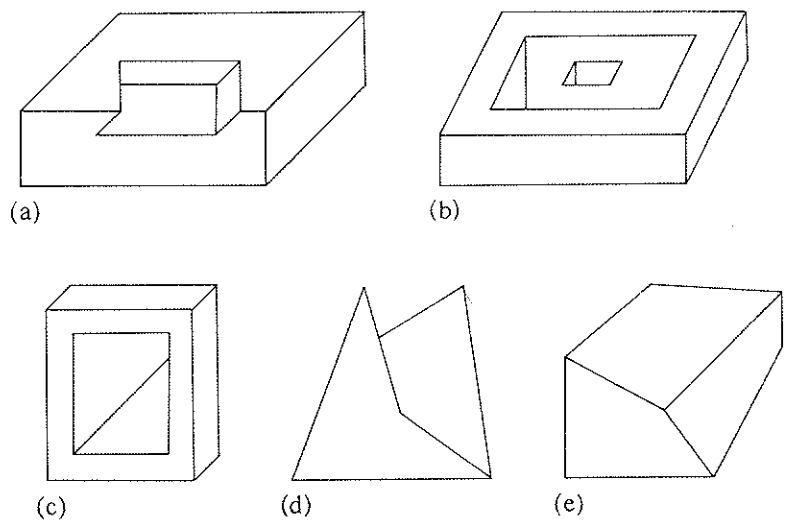


Figure 10. Examples of 'unlikely' objects

that periphery (and in front of one of the three visible surfaces) a seventh surface would be necessary in order to realize the object pictured. The orientation of the picture lines in the example given would require this construction.

In many applications the assumption that 'unlikely' objects do not exist in the scene can be as important a constraint as some of the others mentioned in this paper. The fewer the assumed number of types of objects the easier the picture analysis is.

ADDITIONAL PICTURE CONSTRAINTS

Introductory remarks

It will be seen in this section that even if a picture can be labelled consistently it may not be of a realizable object or set of objects. Additional constraints corresponding to more refined geometric tests must also be satisfied. Examples of pictures which can be labelled but which are not possible are given in figure 11. Figure 11(a) illustrates that the same object surface cannot be associated with two different sides of the same line (independent of which of the four labels the line has). The objects in figures 11(b), (c), and (d) are not possible since the picture gives evidence that two planes intersect along two different edges. In figures 11(e) and (f) the pairwise intersection of three planes gives three lines in the picture which do not intersect (when extended) at a common point, as they should. The same argument eliminates the picture in figure 11(g) as possible; this time the third plane is hidden from

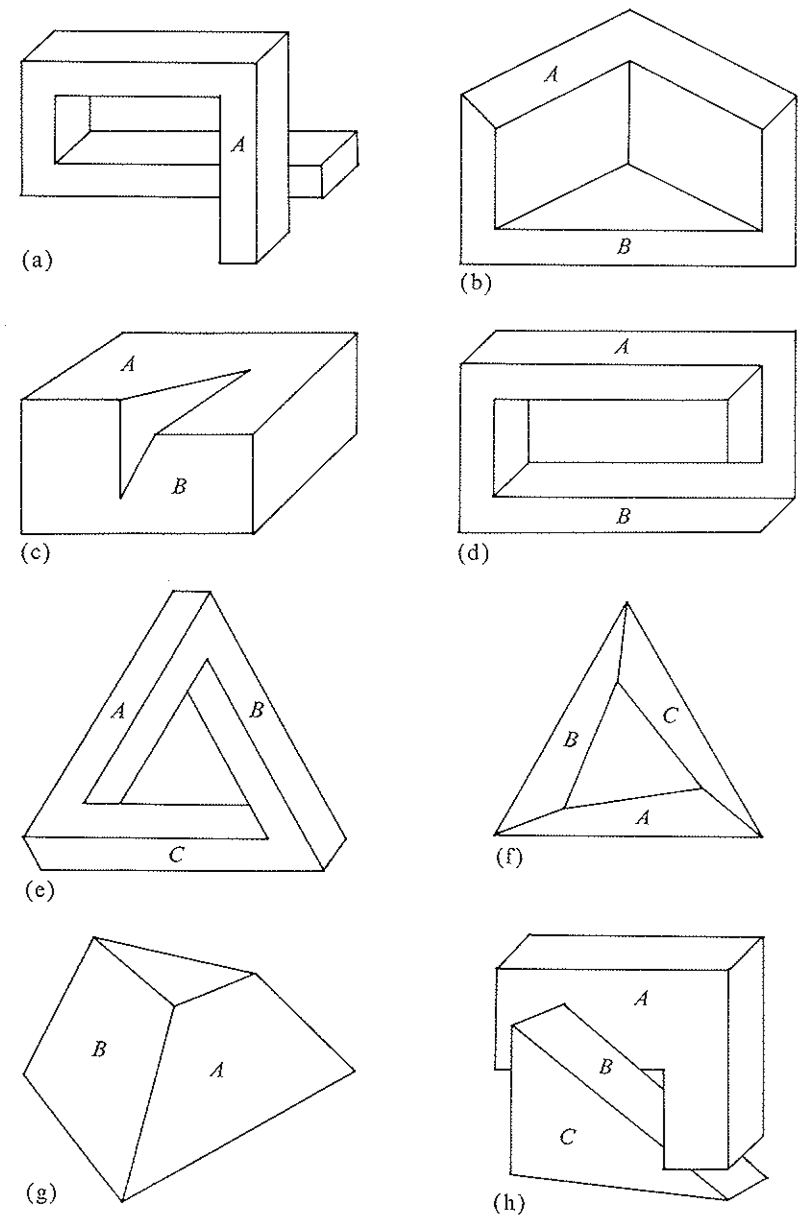


Figure 11. Examples illustrating conditions involving lines common to two faces

view even though the edges at which it intersects the other two planes is visible to the camera. The objects (or object: the two subobjects could be joined by a hidden coupling) in figure 11(h) are possible. This example shows pairs of areas in the picture which are associated with two different lines, but in this case the lines do not represent edges common to the two associated surfaces.

The geometric constraints which were violated in the examples above are simple illustrations of a more general set of necessary constraints. Let us define a *cyclically-ordered edge set* (or 'cyclic-set' for short) to be a cyclically ordered sequence of edges purporting in the picture to have the property that each pair of edges which are consecutive in the sequence determine a distinct plane. We know from the examples above that a cyclic-set with one member is not physically possible. A cyclic-set with two members is not possible unless the two edges are identical.

A cyclic-set with three members is not possible unless the three edges (extended if necessary) would meet at a common vertex. Since we have assumed that our pictures were taken from a general position this implies that the associated triple of lines in the picture would also have to meet at a common node. The three lines incident at a picture node representing a vertex automatically meet the realizability condition. When a cyclic-set has four or more edges the realizability condition is more complicated. For example, it is not necessary that either they or the corresponding picture lines meet at a common point. [See, for example, the four lines extending from the 'base' of the object in figure 13(a).]

The 'gain' concept

Consider the drawing of figure 12(a). The heavy lines (a, b, c, d and e) are intended to represent the edges of a single plane and the lines L_i to represent the directions taken by the 'third' edges at the points P_i. These lines (L_i) represent the cyclic set in question. The lines L_1 and L_2 (and a) are in one plane, the lines L_2 and L_3 (and b) in the next, and so forth.

There is no loss of generality in assuming that the given plane is parallel to the plane of the picture itself. Now imagine that a second plane, parallel to the first, is constructed with edges a', b', c', d', and e'. The lines representing these edges would have to be parallel to the corresponding lines constituting the boundary of the first plane. If the lines L_i have orientations corresponding to a physically-realizable scene the lines representing the boundary of the second plane would have to intersect on the lines L_i as shown in the example.

If the lines L_i have arbitrary orientations the construction illustrated in figure 12(a) will, in general, not be possible. Assume, for instance, that the orientation of L_1 is changed and that those of the other lines associated with the cyclic-set are not. Choose a point X on L_1 and construct a' parallel to a. From the intersection of a with L_2 construct b' parallel to b, and so forth. In general the procedure will determine a point Y (the intersection of e' and

L_1) which is different from X, as has been illustrated in figures 12(b), 12(c), and (12d). The ratio between the distance $\overline{P_1Y}$ and the distance $\overline{P_1X}$ will be defined as the (clockwise) *gain* associated with the cyclic-set. We could have defined a counterclockwise gain; it would be reciprocally related to the clockwise gain. The gain is independent of how far the initial point X chosen for the construction procedure is from P_1.

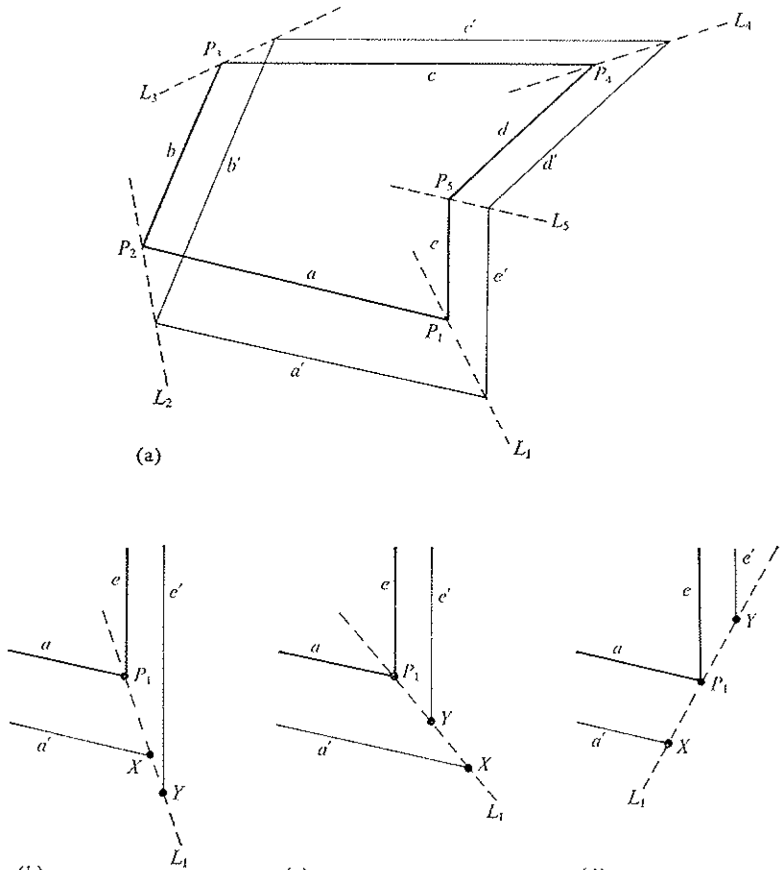


Figure 12. Illustrating the 'gain' concept: (a) illustrating unity gain condition; (b) illustrating gain$\cong$3/2; (c) illustrating gain$\cong$1/2; (d) illustrating gain$\cong -1$

It is clear that a necessary condition for the realizability of an object depicted by a picture is that there be a *gain of unity associated with each cyclically-ordered edge-set*. When these gains differ from unity by only small fractions a picture will generally, but not always, look 'right' (*see* figure 13). Those in which the gains are appreciably different from unity will often look

'warped' (*see* figure 11(f), for example). In a sense the gain (or its reciprocal) is a measure of how 'ungrammatical' a picture is. It would seem possible to draw sets of pictures of objects in which the gain differed more and more from unity so as to test the sensitivity of human subjects to this type of picture distortion. In this application the advantage of having a precise numerical value for the distortion is obvious.

It is fairly easy to show that when the gain associated with a cyclic-set is negative the picture lines cannot even be labelled properly. This is beyond the scope of this paper and is left as an exercise to the reader.

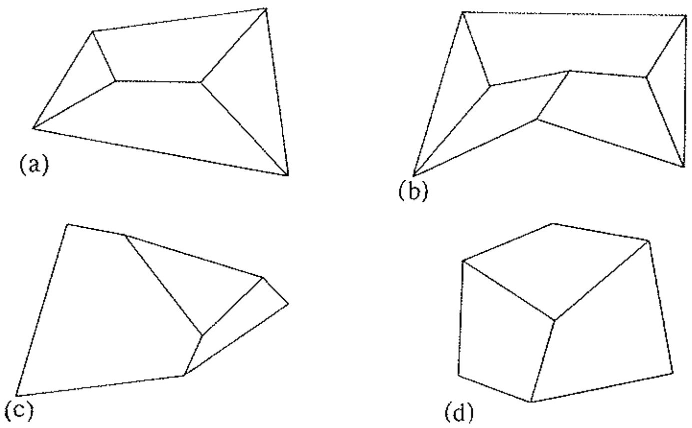
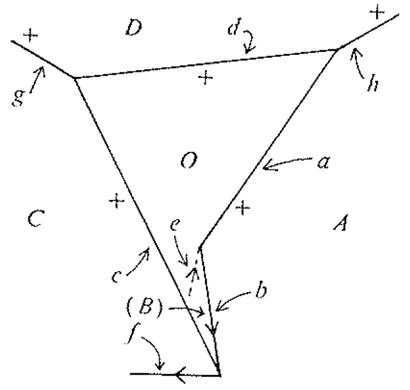


Figure 13. Pictures satisfying the unity-gain condition

The dual representation

A representation which is dual to the picture-graph will be summarized here briefly. Such a representation can be useful in visualizing the orientation of the surfaces of an object and in ascertaining whether or not they can be made mutually compatible.

A *dual picture-graph* is a graph in which there is a node corresponding to every surface represented in the picture, a (triangular) area corresponding to every (degree-3) node represented in the picture and a line corresponding to every line in the picture (*see* figure 14). If the picture is of a possible object it will be possible to construct the dual-graph and orient it so that every line in the dual-graph is at right angles to the corresponding line in the picture itself. The lines a, b, c and d in figure 14(a) represent the boundary of a plane surface in a picture. The directions of the lines e, f, g and h have been chosen so that the unit-gain requirement has been met. Therefore it is possible to construct surfaces of an associated object so that the surfaces A (defined by lines a, e and h), B (defined by the lines b, e and f), C (defined by the lines c, f and g) and D (defined by the lines d, g and h) are plane surfaces.



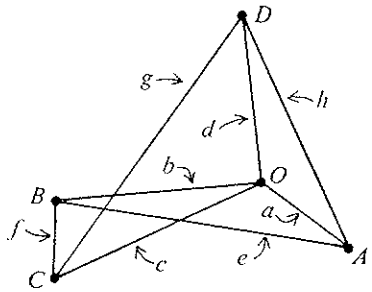


Figure 14. Illustrating duality: (a) picture-graph; (b) dual picture-graph.

We note that the directions of the three lines associated with a node in the picture uniquely determine the proportions of a corresponding triangle drawn in the dual-graph. By progressing around the boundary of a plane represented in the picture we generate in the dual a sequence of triangles the locations of which are uniquely determined once the first triangle in the sequence has been drawn. (Note, for example, in figure 14 the sequence of triangles *abe*, *bcf*, *cdg* and *dah*.) In this process a sequence of lines corresponding to a cyclic-set is generated in the dual graph (for example, the sequence *e*, *f*, *g*, *h*). This sequence will close on itself if and only if the unit-gain requirement is met.

With the dual-graph constructed in this manner its directed line-segments correspond to *gradients* and differences between gradients of planes in the picture. By 'gradient' in this context we shall mean a (two-dimensional) vector associated with a plane surface represented in the picture which, by its direction and magnitude, shows the direction and magnitude of the greatest positive rate of change of distance from the camera with respect to motion in the picture.

Assume in figure 14(a) that the object surface bounded by *a*, *b*, *c* and *d* is parallel to the picture plane. (The gradient associated with that plane is the zero-vector.) The gradient associated with *D* would have the direction of the directed line segment $\vec{OD}$ in the dual-graph. The gradient associated with *B* would have the direction of $\vec{OB}$ in the dual-graph. (Note that the hidden surface *B* in the picture is under the reference plane and is tilted down toward the left.) The lines *a* and *c* have similar interpretations.

The vectors associated with the other lines can be interpreted as (vector)

differences between pairs of the other gradients. With the reference plane assumed to have a zero gradient the node 0 in the dual can be considered to be the origin. Moving the origin away from this point corresponds to transforming the object pictured in such a way that its projection on the picture plane remains constant. Reducing the lengths of all lines in the dual-graph by a given factor corresponds to 'flattening' the object pictured by the same factor.

The (perpendicular) distance from the origin to an arbitrary line in the dual-graph is a measure of the slope of the corresponding object edge with respect to the picture-plane. Thus lines passing through the origin correspond to those edges which are parallel to the picture-plane. Pairs of surfaces which are parallel in the scene would be associated with the same point in the dual-graph.

It is possible to consider the dual-graph itself as a picture of a dual-scene in which information about the 'third dimension' has been lost. In that case the directed line segments in the original picture correspond to gradients of the surface in a 3-dimensional dual-scene. The other obvious dual relationships also hold. Special comment should, nevertheless, be made about the way pairs of lines are oriented with respect to each other in the scene and the dual-scene. If at the apparent point of intersection of two lines, L_i and L_j, in the picture L_i is in the scene actually closer to the camera than L_j is, then in the dual-scene L_j will be closer to the camera than L_i is. It follows that lines which intersect in the scene will also intersect in the dual-scene.

Summary remarks

We considered first a comparatively simple set of labelling constraints which we later found were not a sufficient test for the realizability of an object represented in a picture-graph. A more refined set of necessary conditions resulted from the unit-gain requirement and the alternate viewpoint available from the dual picture-graph. These constitute a productive way of demonstrating whether or not all of the surfaces in the scene can be realized as plane surfaces and whether or not they can have compatible orientations. In spite of this there are still examples of pictures of objects which cannot be analyzed by using only these tests in a straightforward way.

Consider, for example, the object(s) in figure 15. The edges can all be labelled consistently. The unit-gain requirement is met for all cyclic-sets of edges. And yet it is obvious that the scene is impossible to realize. One way of demonstrating this is to note that a properly-oriented piercing line would intersect the *A*-plane in two different points, *p* and *q*. Alternately, one could conclude from the picture that there must be at least two different lines in which the surfaces *A* and *B* must intersect: one in the left half of the picture and one in the right half. Other relatively simple tests give equivalent results.

It is not difficult to imagine other more complex sets of objects which can be proved to be possible or impossible *if* an appropriate set of auxiliary points, lines, or planes are added to the picture. Once these are added a proof of impossibility may be quite short and based on easily-understood principles of geometry. But to know *where* to apply these principles will probably require, in general, an analysis, the details of which are specific to the given picture.

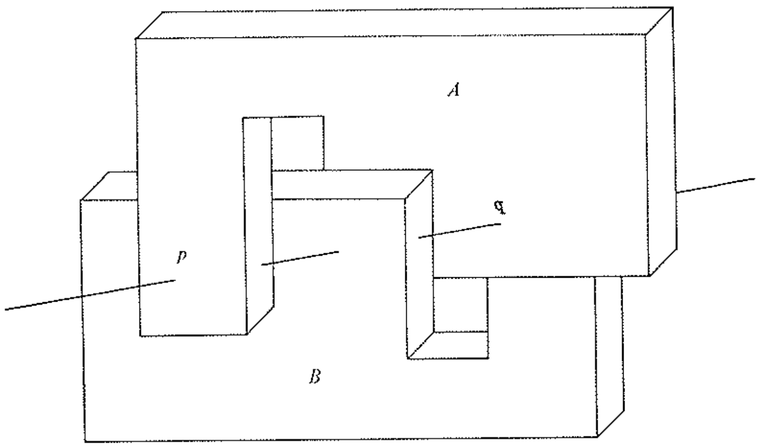


Figure 15. An impossible pair of objects

We have assumed for the purposes of this paper that the polyhedra which are pictured have only three edges associated with each vertex. This assumption is not so restrictive as it might at first seem. Vertices having four or more associated edges can be converted to sets of vertices having only three by a procedure of 'grinding' and/or 'filling' in the vicinity of the vertices. The vertex of the original polyhedron can then be considered to be the limit reached by the inverse procedure (in which the amount of ground or filled material is reduced to zero) and the configuration of lines at the corresponding picture node to be the conjunction of several of the basic configurations shown in figure 6. If we may assume that the polyhedra in our pictures can have arbitrarily large numbers of edges associated with their vertices the primary condition for the realizability of the objects pictured is that no two 'T' configurations exist which would imply arrow labels in different directions on the same line-segment [*see*, for example, figure 8(a)]. The results relating to the unit-gain requirement and the dual formulation apply without significant modification to pictures of these more general polyhedra.

A LANGUAGE FOR PICTURES OF SMOOTH OBJECTS

Types of picture lines and nodes

In this section I will present a brief summary report of work involving pictures of 'smooth' objects. A more detailed paper will be forthcoming at a later date.

'Smooth' here will be given an informal definition. By a *smooth surface* we shall mean one having, for each surface point, a well-defined directed line orthogonal to, and pointed away from the surface such that the orientation of this line is a continuous function of the position of the associated point on the surface. For the sake of simplicity we shall allow in our pictures only smooth surfaces bound by *smooth edges* (defined in a manner analogous to the one in which smooth surfaces were defined) and solid objects bound by smooth surfaces having no edges. Smooth objects may, of course, be quite complex. For instance, it is possible to associate either orientable ('two-sided') or nonorientable (' one-sided') smooth surfaces with arbitrarily large sets of edges which are arbitrarily knotted together. Consequently, smooth objects furnish another rich source of pictures for analysis.

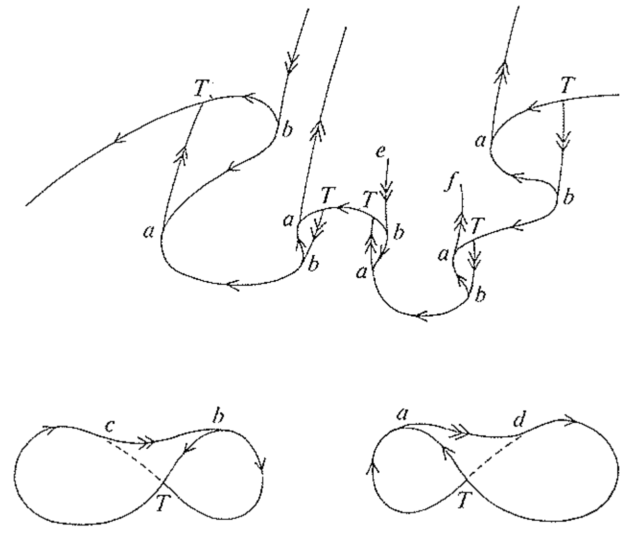


Figure 16. Cloth and saddles

In figure 16 there are some examples of pictures of smooth objects. Note that picture lines can represent either *edges* of a surface or *folds* in the surface. We distinguish between these two types of lines by labelling the first type with *single arrows* and the second type with *double arrows*. In each case we orient the arrow so that, as we travel in the direction of the arrow, the associated surface (one for a line representing an edge) or surfaces (two for a line representing a fold) are to the right in the picture. Thus the number of arrows in the label for a line is also a measure of the difference between the numbers of surfaces represented in the two areas on the two sides of the line.

It will be noted that at each indicated node in the picture (except the T-nodes) two of the lines meet at a *cusp*. This is true even for the nodes identified by '*c*', '*d*', '*e*' and '*f*' (at these nodes one of the lines represents a hidden edge or fold). This is a consequence of the fact that the surfaces and edges represented are smooth. For nodes of types '*e*' and '*f*' only one visible line is incident and this line always represents a fold. For pictures of *solid* smooth objects only these two latter types of nodes (and T-nodes) are necessary.

Derived constraints

The possible configurations of lines which can exist at picture nodes is summarized in figure 17. The last four configurations are possible in pictures of *creases* and are added to show how this style of picture language could be expanded. Surfaces which are creased are, however, not smooth in the sense that that term has been defined here and consequently in this paper we shall not consider further these last four types of picture nodes.

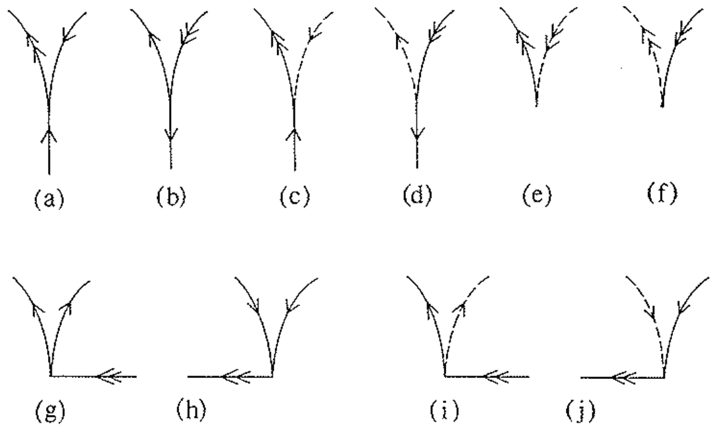


Figure 17. Configurations possible in a picture of smooth objects

By examining the first six configurations of figure 17 we can conclude that they occur in pairs which are related by being mirror images of each other. The mirror image of a configuration has its arrows pointing in directions opposite to the corresponding arrows in the original.

The hidden lines in figures 17(c), 17(d), 17(e), and 17(f) are all hidden under exactly one surface. When each of these configurations is placed under one or more surfaces none of the lines will be visible but the one will have a depth-index which is one more than those of the other lines.

In each of these configurations the number of arrows into a node equals the number of arrows out of that node (when hidden lines are taken into account). This is a consequence of the fact that the number of surfaces associated with a given area of the picture (bounded by lines which may be visible or hidden) is a constant. We can conclude that the net number of arrows into *each* closed picture region is also zero.

The surfaces, edges, and folds in a scene map into three basic unlabelled

configurations (other than the 'T') in a picture: the 'gamma' node [figures 17(a), 17(b)], the 'lineal' node [figures 17(c), 17(d)] and the 'terminal' node [figures 17(e), 17 (f)]. In making decisions about the possibility or impossibility of smooth objects the lineal nodes cause special concern since one must consider the possibility that one or more lineal nodes may be concealed along any curved line segment. Nevertheless, even though the *number* of arrows associated with a line label changes as a lineal node is passed, the *direction* of the arrow(s) is unchanged. Consequently we know that the direction of the arrows appropriate for labelling a given segment of line is fixed over that line segment (until a gamma node or a terminal node is reached).

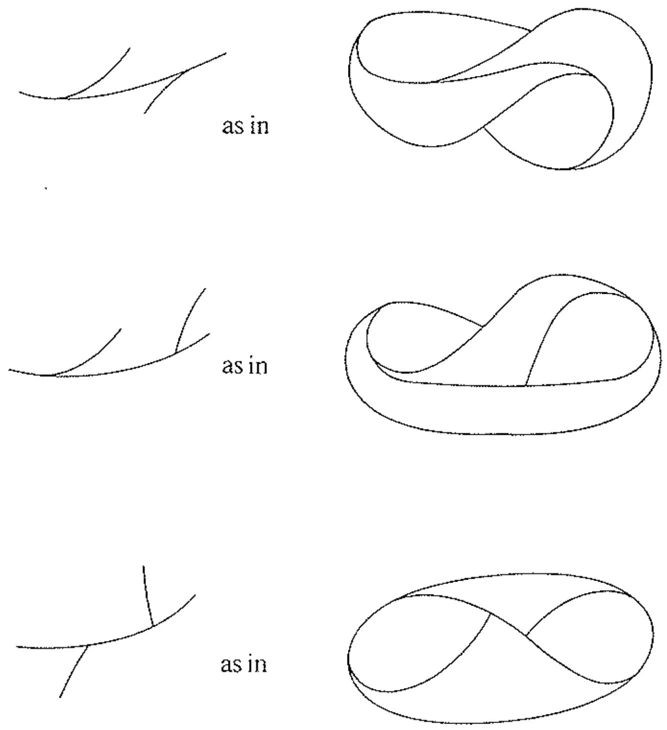


Figure 18. Forbidden configurations in pictures of smooth objects

An equally important conclusion follows from the observation of the direction of the arrows in the 'arms' of a gamma configuration: the arrow on the right arm is always into the node and the arrow on the left arm is always out of the node. This fact and the constancy of arrow direction in the vicinity of lineal nodes allow us to conclude that the three configurations in figure 18 (and their mirror images) cannot occur in pictures of possible objects. A slightly subtle version of the configuration of figure 18(b) is imbedded in the picture of figure 3(d) and prevents that object from being possible.

An example for study

The pictures in figure 19 show four different ways of completing the hidden line structure for the example of figure 3(a). The first of these shows the simplest and perhaps the most obvious way of continuing the visible lines as hidden lines. The labellings given on the hidden lines are compatible with those on the visible lines. Nevertheless the object is *not* possible.

In order to prove the impossibility of this object we examine the line which goes from p past q to r. This line indicates an edge which disappears between p and q and reappears between q and r. The line which *crosses* the path between p and q has a double arrow; this gives evidence that the edge represented at q is under *two* surfaces. On the other hand the line which crosses between q and r has a single arrow and this gives evidence that the edge represented at q is under only *one* surface. Since the line at q can have only one depth-index the object with the hidden line structure given in figure 19(a) is clearly impossible.

Other ways of completing the hidden lines are given in figures 19(b), (c), and (d). For each of these there is a way not only of labelling the lines in a consistent way but also of assigning a depth-index (shown by integers in the figure) to each line segment. Each configuration at a hidden picture node is basically the same as one of those in figure 17 but with the depth-index of all lines of the configuration increased uniformly by a number equal to the number of surfaces existing between the basic configuration and the camera. The picture in figure 3(b) is of a realizable object. One lineal node and one hidden gamma node suffice. The picture in figure 3(c) is also of a realizable object. Two lineal nodes and one hidden gamma node suffice. These examples are left as exercises for the reader.

Conditions for realizability

Any one of the three successful ways of adding hidden lines to the picture example above demonstrates the possibility of the object. The question of which is 'best' cannot, of course, be answered on an absolute basis. A reasonable preference is for the one (or ones) which require the fewest additional picture nodes. A reasonable constraint to impose is a minimum allowable radius of curvature for the hidden lines. Such a constraint would be somewhat analogous to an upper frequency limit or bandwidth restriction for continuous signals.

The author conjectures that a necessary and sufficient condition for the realizability of a picture of a smooth object (or objects) is that there be no conflict among the arrow *directions* imposed on a picture line by its association with an arm of a gamma node or the bar of a T-node (recall figure 18) or because of its identification as a part of the boundary of an object. The

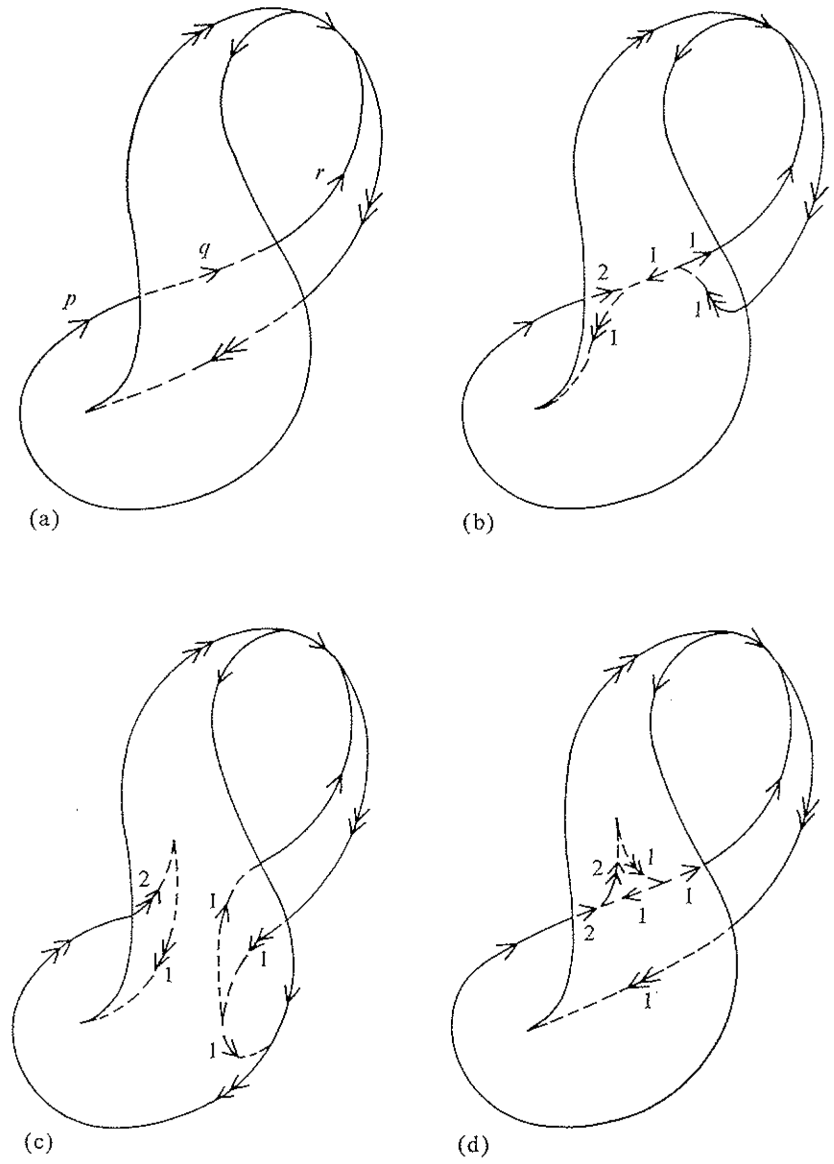


Figure 19. Possible hidden line configurations for a smooth object.

necessity of the condition is obvious. Its sufficiency would follow from the existence of a procedure for bringing together in a compatible way the hidden lines associated with T-nodes, lineal nodes, and terminal nodes. A key construction in this procedure was illustrated in figure 19(d). It allowed a single-arrowed line to go from one depth to another in the picture. Such techniques allow us to bring together hidden lines at the appropriate depth-indices.

In the proposed procedure (or in others which achieve the desired result) we must take into account the following issues. When two hidden lines cross in the picture the depth-index of the lower line must be changed by the number of arrows on the upper line. (This is a generalization of the rules appropriate at a T-node.) The depth-index of a line cannot exceed the number of surfaces associated with the areas on either side of the line, nor can the number of surfaces associated with a picture area be negative.

SUMMARY REMARKS

Very few real-world scenes contain only degree-3 polyhedra or simple smooth objects. Nevertheless the study of the picture languages associated with such simple scenes serves its purpose by giving us some perspective about what we might and might not expect to accomplish from the study of more complicated picture languages. As a minimum we should at least be better able to decide when sentences from a 'polyhedral language' or 'smooth language' are or are not present in a picture containing a babble of several languages. Pictures of objects with combinations of smooth and plane surfaces and having additional types of 'vertices' (for example, cones and cylinders) may be analyzed using combinations of the techniques illustrated in this paper.

As we have mentioned earlier additional picture configurations are possible if we allow creases in the surfaces. Other special types of surface features correspond to their own element types. For instance, the configurations around picture points which are associated with 'crumpled paper' surfaces are constrained by the fact that in the scene the angles which are represented at the picture point must sum to 360°. Wrinkles, puckers, scoring, and indentations all cause their own distinctive configurations in a picture and add to the list of symbols possible in a picture language.

Stereo pairs of pictures, sequences of pictures taken at slightly different times, and pictures containing shadow information (they are all closely related) offer other examples worthy of study.

Acknowledgement

I plead guilty to the charge that I deal with pictures of impossible objects because it is fun. It is, and that is reason enough. However, in addition to this I believe that much can be learned in the study of any language by asking '*Is* that a nonsense sentence?' and '*Why* is that a nonsense sentence?'.

Much of the material presented here was stimulated by contact with and tested on the members of the Artificial Intelligence Group at the Stanford Research Institute. I am especially grateful to the group manager, Charles A. Rosen, for his encouragement, enthusiasm, confidence, and patience.

REFERENCES

Guzman, A. (1968) Decomposition of a visual scene into three-dimensional bodies. *Proc. of the Fall Joint Computer Conference*, pp. 291–304.

Huffman, D.A. (1968) Decision criteria for a class of 'impossible objects'. *Proceedings of the First Hawaii International Conference on System Sciences*. Honolulu.

Penrose, L.S. & Penrose, R. (1958) Impossible objects: a special type of illusion. *Brit. J. Psych.*, **49**, 31.