Vision Flash 4
December, 1970
Revised
February, 1971

The Object Partition Problem

Eugene C. Freuder

Vision Group
A.I. Laboratory
M.I.T.

I.

We provide a theoretical framework for the "object partition problem".

We begin by considering this problem in a rather abstract context. Consider the semantic domain of physical objects, and the syntactic domain of picture graphs. The object partition problem can be viewed in terms of specifying syntactic operations that yield partitions of the picture graph, which when interpreted in the semantic domain correspond to possible partitions of the scene into physical objects. We term such syntactic partitions "physically realizable partitions" or simply "realizable partitions". The problem has several aspects. We may seek all possible partitions, the one "best" partition, several plausible partitions, judge or rank proposed partitions, and so forth.

The most important aspect of the problem, in some sense, is determining the "best" or most likely realizable partition. Guzman's SEE attempts to deal with this problem. Guzman uses his understanding of the semantic physical world of objects to make local choices on the best partition of a given picture graph. These choices are somewhat interrelated. They are combined to inform global decisions which may also be interrelated.

Guzman's SEE is a remarkable heuristic programming achievement. However, as might be expected of the germinal achievement in its field, it lacks a satisfactory theoretical framework. There is some difficulty in determining the motivations and implications of Guzman's heuristic decisions. It is not always obvious just what semantic observations inform the selection of the syntactic partitioning choices. As it is not clear which possible interpretations are being discarded at each stage of the procedure, the process cannot produce alternative partitions, even where several plausible interpretations are present in the physical scene. To judge or extend the work, one must essentially repeat Guzman's experience with individual scenes.

The program does not function as a good element for a heterarchical vision system. There is no proper framework from which to launch dialogue with other knowledge structures relevant to the object partition problem. There is not the flexibility to provide alternative analysis on the basis of higher level dissatisfaction.

It may be argued that in order to establish a satisfactory theoretical base for a property of physical scenes, such as object partition, one must present a system capable of dealing with all possible physical interpretations of a given picture graph, in terms of the relevant property. This would require our theory of object partition to be "complete" in the sense that it could deal with any physically realizable partition. One criteria of completeness would be that the theory could produce all realizable partitions, though the system would not necessarily be "generative" oriented.

A "complete" characterization of realizable object partitions would hopefully provide an organized framework in which choices could be made among possible local partitionings in determining the "best" possible global partition. Decisions involving these choices would be semantically and heterarchically informed, and their implications and motivations would be clearly understood. Alternative choices could be made, in some plausible order.

Such a characterization would cut down our decision space by eliminating inconsistent or impossible partioning choices. The nature and range of the remaining heuristic choices would be clarified.

Recently Huffman has approached another problem of scene analysis, the "configuration" problem, in somewhat this fashion by attempting to identify physically unrealizable configurations. The results of his venture recommend this type of approach.

However, Huffman's success has also encouraged speculation that his theory may contain as well the desired complete characterization approach to the "object partition" problem. The "physically realizable configuration" and "physically realizable partition" problems are interrelated and interdependent. However, they are not identical, and it would be a mistake to base an approach to one upon the conceptual units that characterize the other. The treatment of the former problem has been begun by Huffman. Much remains to be done, however; some relevant observations will appear in future papers.

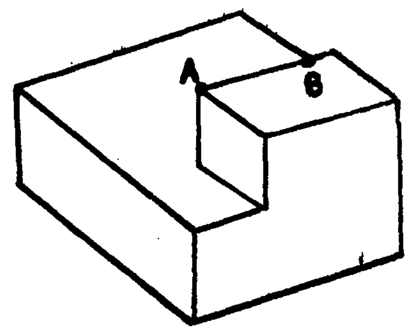
The "characterization" approach to the "object partition problem" will be outlined below.

II.

We take for our basic units of analysis a line predicate and its negation. The predicates are "belong to the same body" and "do not belong to the same body", as applied to the (two) regions bordering a line.

This choice of our unit of analysis is neither as trivial or as circular as it may appear. To begin with the basic predicate defined is NOT equivalent to the predicate "links". The success of the link predicate is probably a major reason that the more fundamental predicate we described was not used earlier. However, deep problems in partition analysis reveal that "link", while a useful concept, is actually a handicap to optimal thinking when used as the basic unit of analysis.

We note that "belong to the same body" does not even imply that the line referenced corresponds to a physical edge of both neighboring regions. Consider line AB in the following figure:



We begin by enumerating all physically realizable interpretations or "labellings" of the various types of vertices, in terms of our two predicates. All physically realizable partionings of a given picture graph are then obtained from all consistent combinations of local labellings. (Using the criteria that an interpretation applies to an entire line, i.e. line segment, and thus a line cannot receive opposite labels from its two vertices.)

We will describe this process in some further detail. we demonstrate first how this approach circumscribes the realizable interpretations, in achieving a solution to the "all realizable partitions" aspect of the object partition problem. We then go on to indicate how this approach provides a basis for the decision making procedure which deals with "best" realizable partition and related problems. We will gradually shift our focus from an abstract theory to a theoretical model embodied in a "partition system" embedded in a heterarchical vision system.

We will find it easier perhaps to talk in terms of labellings and deal with picture graph elements so we introduce the notation "ı" for "belong to same body" and "W" for "do not belong to same body". (These notations are, if you like, syntactic elements in the picture graph language which are interpreted as the indicated physical relations.) We will refer to "ı" labellings as "ties" and "W" labellings as "breaks". If we simply consider any possible syntactic labelling of an n-line vertex we obviously have 2 to the n possible labellings. However, we make the restriction that "tie" is an equivalence relation, using an observation of the properties of the "belong" predicate. For an n-element vertex, transitivity thus eliminates all labellings with n-1 tie labellings and one break labelling.

Consider three line vertices. The number of physically realizable labellings now corresponds to the number of labellings with no "ties", all "breaks", with one tie, two breaks, and with three ties. I.e.

$$3!/0!3! + 3!/1!2! + 3!/3!0!$$

$$= 1 + 3 + 1$$

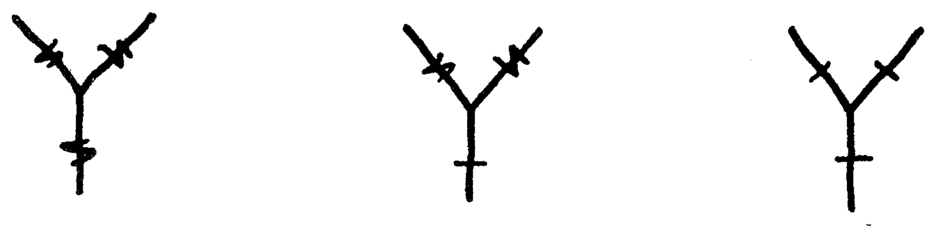
$$= 5$$

A similar analysis can be carried out for n-line vertices.

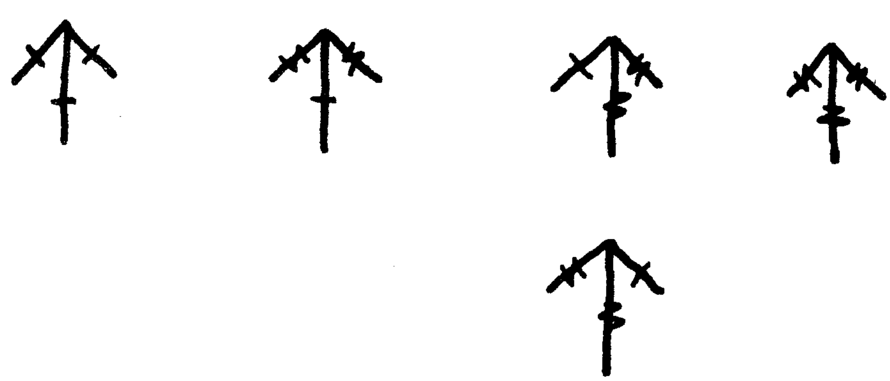
In considering vertices of specific forms, however, we may find that different symmetries further reduce the number of different labellings or group them into classes. We will call the vertex labellings that have a physically realizable interpretation simply "realizable labellings".

Examine the realizable labellings for the different three line vertex types, forks, arrows, and T's.

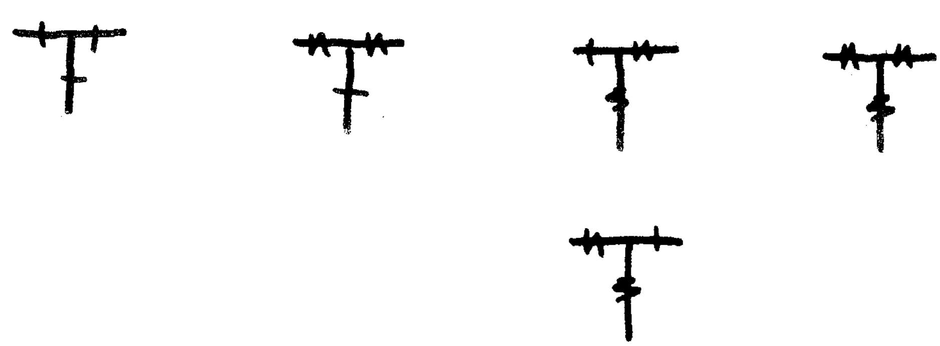
Forks



Arrows

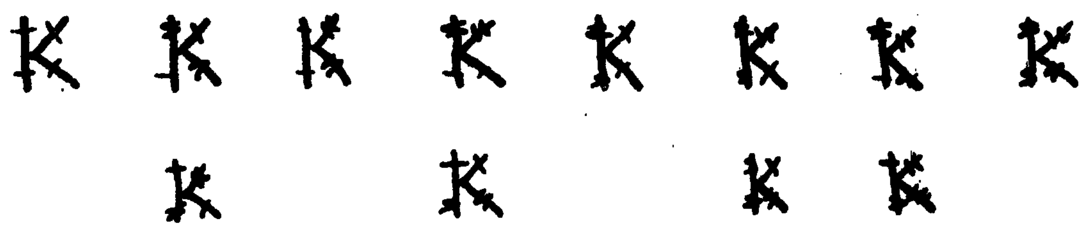


T's



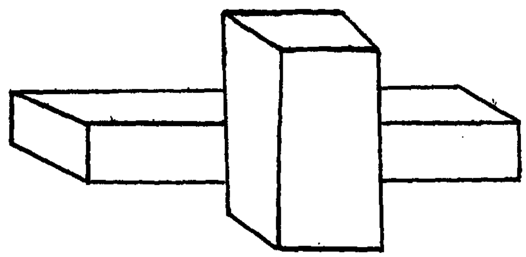
We can easily dispose of the two line vertex "L".

Consider also the interpretable labellings of "K" type vertices.



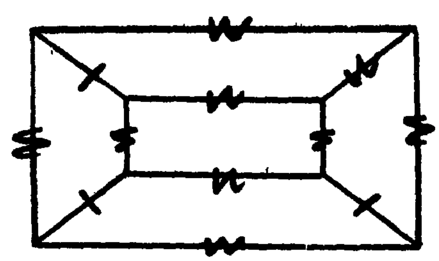
We find that the number of different realizable labelling classes even for the K is surprisingly manageable. Recall that we have placed no restrictions on the physical domain (beyond arbitrary planar polyhedra), or on possible interpretations of picture elements. If we wish to do so we may cut down the number of realizable labellings even further. For example, we may impose restrictions of physical objects to degree three polyhedra, and certain "general position" requirements.

(We might note that this analysis does not pertain to totally disconnected bodies, e.g. joining the two halves of the partially hidden body in the following figure:



This is a different class of problem.)

Applying these realizable labellings to a scene in all consistent combinations provides a solution to the "all physically realizable partitions" aspect of the partition problem. We note that finding consistent combinations also involves checking for "global transitivity". A region R cannot be asserted to belong to the same body as region S by transitivity of the "tie" predicate while at the same time a "break" predicate asserts that R does not belong to the same body as S. We have restricted our labellings so that this inconsistency cannot occur around a single vertex; however, we still have to guard against its occurence across several line boundaries, as in the following figure:



III.

A complete characterization of the partition problem has been achieved.

Beyond this a foundation has also been laid for dealing with the "best possible realizable partion" and related aspects of the partition problem. The objective here is to produce the most plausible partition, with the ability to retrench and produce alternatives if required.

Any realizable partition of a picture graph may be produced by an appropriate labelling of the vertices. At any vertex we have a number of choices of realizable labellings depending on the type of vertex. We consider factors which may prohibit or dictate choices, or rate them on a plausibility scale. These factors may be global, in the sense that they affect our choices of arrow labellings, say, regardless of where the arrows appear in the picture graph, or they may be local, in the sense that they affect a decision at a particular vertex. (Or their effect may fall somewhere in between these extremes.)

Our first approach should be to return to the semantic context in which our picture graphs are to be interpreted. On the most general level this means combinations of planar polyhedra. By studying such combinations we can make judgments as to the relative plausibility of the various realizable labellings for the different vertex types. Even in a gross statistical sense, we can observe that certain labelling interpretations are more common or likely than others. These observations may be extended to certain combinations of vertices.

Beyond this stage we may recognize that certain restricted physical contexts limit the range of interpretable labellings or otherwise affect the relative likelihood of certain interpretations. We observed above, for example, that limiting the physical domain to degree three polyhedra would have this effect. Limitation to convex objects, or likelihood of concave objects, would also affect labelling choices.

The fork possibility, for example, that contains two breaks and one tie, is highly unlikely except in a concave object. With the "K" possibilities clearly in hand we may overcome the intimidation of K analysis and quickly observe that many possibilities are highly unlikely or impossible in many common contexts. In fact, as the unlikelihood of alternative choices rises much faster for K's than for forks say, we may find K's at times more helpful to our analysis than some "simpler" vertices. And we can afford to rate certain choices as implausible since that does not mean we have not dealt with them; they are still available as alternatives in our complete analysis.

The point to be made here is not so much that a limited system could be designed that would function well in a specific restricted context. Rather a complete partition decision structure, imbedded in a heterarchical vision system, could employ information from a context decision structure to advise its labelling decisions or alter its plausibility ratings.

Of course the context decision structure, in turn, could benefit from an understanding of the implications of which labelling choices were being made. In particular, if the partition process experienced difficulty in providing a plausible partition within the constraints of the context analysis provided it, the context analyser could be prompted to reconsider its findings. In a similar fashion the partition analyser may complain that an implausible choice is perhaps due to improper input data from the preprocessor; or it may be informed to expect certain unlikely configurations as a result of missing lines, shadows, whatever.

Aside from what we might call "context" information, labelling choices may be informed by knowledge and knowledge structures of many different kinds within a heterarchical structure. Background information, for example will immediately determine many labellings and circumscribe or suggest others. There are a variety of line predicates aside from the basic partitioning predicate which are relevant to partitioning decisions (concave and convex fall into this category). And there are more global criteria involved. Future papers will deal with these factors, their relation to partition analysis and to scene analysis in general.

Relevant information may restrict or advise our partitioning choices similarly at other than the vertex labelling level, of course. Decisions may be indicated at the individual line level, at higher levels involving types of vertex combinations, whatever.

The beauty of our characterization approach is that all this information and all these decisions can be done within a systematic, complete framework. All the realizable interpretations are available for analysis and comparison. The system can make decisions with a clear understanding of just what the choices are and what the implications of these choices are. We know what possibilities are discarded at each decision point.

These features of the characterization theory also indicate why it is a great aid in organizing, stimulating, and clarifying our thinking, in determining precisely how the relevant factors outlined above determine the specific decision procedures used by a partition system.

One of the difficulties in dealing with partition analysis is that the problem, like many scene analysis problems, is "potentially global". For any judgment one makes about a certain local configuration implying a certain partioning, one can usually find an exception by embedding the configuration in a sufficiently complex environment. Another advantage of our theoretical approach is that the basic aspect of this global determination is built in to the structure of the consistent labelling approach, on a network level.

Local labelling decisions affect others in a potentially global relationship. Once some (or all) of the lines of a vertex have been labelled by neighboring labelling decisions, the labelling of that vertex may be determined, or the possibilities cut down. The labelling of one or more lines of a vertex may direct our attention to the most plausible labelling, for the vertex type, which agrees with the already labelled lines.

The interrelationship built into the labelling system guarantees that potentially global determining relationships will be considered. When this necessitates resolution of conflicts and discarding decisions our system will know just what decisions were involved and be able to evaluate them. The appropriate alternatives will be available. And our theoretical base insures that our options provide a complete set of possible solutions.

Our system, in other words, has the "freedom to fail". It has long been an educational cliche that this is a prerequisite for accomplishment. The very notion of "heuristic" programming implies, not algorithms, but sets of principles, some of which may fail at any given application. The most successful concepts of heuristic programming have dealt with systems that could recover from or use these failures in some fashion. Without this capability one is often forced to tightly restrict the problem domain, or to deal with problems of hopeless complexity. In scene analysis the "potentially global" problem, in particular, makes it difficult to make decisions that "have to work". Our approach to the "best partition" problem gives us the flexibility to fail, with the added confidence that, since we have characterized the set of realizable partitions, we must, at least in some theoretical sense, be able to achieve the desired solution.