

Construction and optimization of CSG representations

Vadim Shapiro and Donald L Vossler*

Boundary representations (B-reps) and constructive solid geometry (CSG) are widely used representation schemes for solids. While the problem of computing a B-rep from a CSG representation is relatively well understood, the inverse problem of B-rep to CSG conversion has not been addressed in general. The ability to perform B-rep to CSG conversion has important implications for the architecture of solid modelling systems and, in addition, is of considerable theoretical interest.

The paper presents a general approach to B-rep to CSG conversion based on a partition of Euclidean space by surfaces induced from a B-rep, and on the well known fact that closed regular sets and regularized set operations form a Boolean algebra. It is shown that the conversion problem is well defined, and that the solution results in a CSG representation that is unique for a fixed set of halfspaces that serve as a 'basis' for the representation. The 'basis' set contains halfspaces induced from a B-rep plus additional non-unique separating halfspaces.

An important characteristic of B-rep to CSG conversion is the size of a resulting CSG representation. We consider minimization of CSG representations in some detail and suggest new minimization techniques.

While many important geometric and combinatorial issues remain open, a companion paper shows that the proposed approach to B-rep to CSG conversion and minimization is effective in E^2 . In E^3 , an experimental system currently converts natural-quadric B-reps in PARASOLID to efficient CSG representations in PADL-2.

solid modelling, boundary representation, constructive solid geometry, Boolean operations

Rigid homogeneous solids may be modelled by sets of points in E^3 that are compact, regular, and semi-analytic; such sets are called 'r-sets'^{1,2}. Six families of unambiguous (informationally complete) representation schemes for r-sets are known³. The two representation schemes that are most widely used today, constructive solid geometry (CSG) and boundary representation (B-rep), are illustrated in Figure 1. While the problem

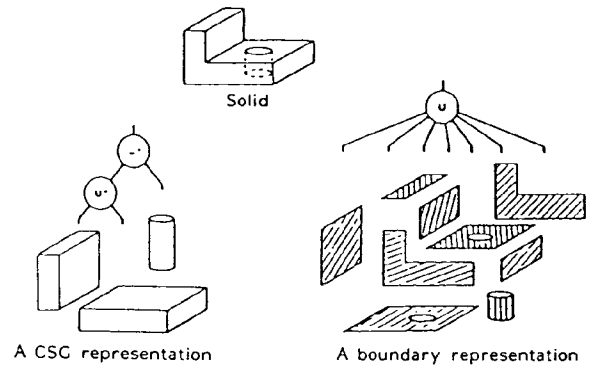


Figure 1. A simple solid and its representations: constructive solid geometry (CSG) and boundary (B-rep)

of computing a B-rep from a CSG representation is relatively well understood⁴, the inverse problem of B-rep to CSG conversion has not been addressed in general. The importance of B-rep to CSG conversion can be seen from the following considerations:

- The inability to perform B-rep to CSG conversion has put significant constraints on the design of modern solid modelling systems³, as shown in Figure 2(a). Figure 2(b) shows an 'ideal' architecture of a solid modelling system that allows representation-specific technology to be used on alternative representations through a bilateral B-rep-CSG conversion⁵.
- B-rep to CSG conversion for planar solids is performed in user-interfaces of solid modellers like PADL-2⁶ and Unisolds⁷. Many common mechanical parts can be easily constructed by extruding or revolving a CSG representation of a planar cross-section in the direction perpendicular to the plane (Figure 3(a)). 2D B-rep to CSG conversion has received wide attention in the literature (for a survey see Shapiro and Vossler⁸).
- The size of computed CSG representations is crucial to the performance of subsequent application programs. Figure 3(b) shows all the primitives used in a CSG representation of the solid in Figure 3(a). This representation is verbose because many primitives have coincident surfaces while some primitives may not be necessary. Thus we must deal with CSG minimization, which subsumes the

Sibley School of Mechanical & Aerospace Engineering, Cornell University, Ithaca, NY 14850, USA (also with the Computer Science Department, GM Research Laboratories, Warren, MI)
*McDonnell Douglas Corporation, Cypress, CA, USA
Paper received: 4 May 1990. Revised: 29 October 1990

problems of redundancy elimination and null object detection, as addressed for example by Tilove⁹, Rossignac and Voelcker¹⁰ and Woodward¹¹.

- Finally, the issues surrounding B-rep to CSG conversion are of significant theoretical interest. We shall see that conversion relies on tools from algebraic and computational geometry, and raises unexplored issues.

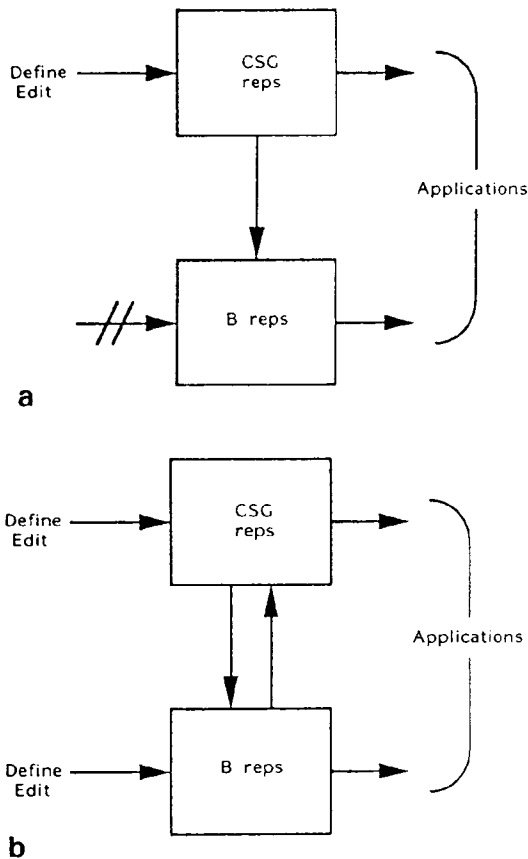


Figure 2. Architectures of dual-representation solid modelling systems: (a) classical unilateral, (b) 'ideal' symmetric

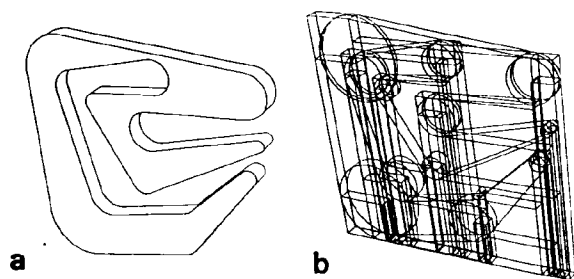


Figure 3. (a) Solid obtained by extruding a planar cross-section (courtesy of Richard Marisa). (b) The primitives used in an inefficient CSG representation of the solid in (a)

B-rep and CSG representations of closed regular sets

In this paper, it is assumed that a solid object S is given by a valid B-rep \hat{S} . The particular type of B-rep is not important, as long as it specifies exactly all surfaces containing the faces of S . Each surface is used to induce a halfspace. B-reps whose faces do not associate naturally with halfspaces, e.g. B-reps whose faces are parametric 'patches', pose problems that lie outside the scope of this paper.

Halfspaces are denoted by lower-case literals, such as g and h , and define regular semi-analytic subsets of W^* . For example, a halfspace is often given by $\{p \in W \mid f(p) \geq 0\}$, where f is a (semi-)analytic real function, or it could be defined by any other unambiguous representation. The important property of a halfspace is that any point $p \in W$ can be classified with respect to h as being in, on, or out of h , using the language of set membership classification¹².

Properties of CSG representations are summarized in Requicha and Voelcker¹³, following research on mathematical foundations by Requicha¹ and Requicha and Tilove². Central to CSG are notions of closed-set regularity and regularization relative to the topology of the universal set W . Regularization of a set X is defined by

$$\text{reg}X = \text{closure}(\text{interior}(X))$$

and a set X is called closed regular (or simply regular) if $X = \text{reg}X$. Similarly, regularized set operations are defined as follows

$$h_i \cup^* h_j = \text{reg}(h_i \cup h_j)$$

$$h_i \cap^* h_j = \text{reg}(h_i \cap h_j)$$

$$h_i -^* h_j = \text{reg}(h_i - h_j)$$

$$\bar{h}_i = \text{reg}(W - h_i)$$

The properties of regular sets and regularized set operations were studied by Kuratowski and Mostowski¹⁴ and Requicha and Tilove². In particular, it is well known that regular sets form a Boolean algebra under operations of \cup^* , \cap^* , $-^*$. The class of CSG-representable objects is determined by the set of primitive regular halfspaces.

It is important to distinguish between a regular semi-analytic set S and its CSG representations. A CSG expression Φ is a Boolean form, i.e. an expression composed from halfspace literals h_1, \dots, h_n , and symbols denoting regularized set operations, defining a Boolean function F . When Φ is applied to a specific set of halfspaces H , it becomes a CSG representation $\Phi(H)$ whose value $F(H) = |\Phi(H)|$ is a semi-analytic set

† In this paper W is a d -dimensional Euclidean space E^d , $d \leq 3$.

‡ Note that r -sets do not form a Boolean algebra, but rather a ring with operations \cup^* , \cap^* , $-^*$, due to the requirement that r -sets be bounded. Without loss of generality, we deal with CSG representations of regular (perhaps unbounded) sets, with the understanding that such representations are valid only when the boundedness of the represented sets can be established

of points S . The set $S = |\Phi(H)|$ is unique because CSG representations are unambiguous, while its representation $\Phi(H)$ is not. Two representations $\Phi_1(H)$ and $\Phi_2(H)$ represent the same set if and only if $|\Phi_1(H)| = |\Phi_2(H)|$. Note that, technically, a halfspace literal h represents a regular set $|h|$. For example, we should properly speak of classification with respect to $|h|$, and not h . However, when the meaning of a single halfspace is obvious as above, we will use h to denote either the literal or the set it represents.

We will use capital Roman symbols, such as F , G , and S , for functions defining regular sets and Greek symbols, such as Φ , Ψ , and Π , to denote CSG expressions. We say that Φ represents set S if and only if $S = |\Phi|$. Henceforth we will use Boolean addition ($+$) and multiplication (\cdot) to denote regularized union (\cup^*) and regularized intersection (\cap^*) respectively.

B-rep to CSG conversion problem

The fact that CSG representations are not unique¹³ is commonly perceived as a major difficulty in performing B-rep to CSG conversion, leading to many *ad hoc* solutions that are often restricted to narrow geometric domains. On the other hand, every Boolean function on n variables can be written in a canonical form that is unique for a fixed choice of variables (see for example Kuratowski and Mostowski¹⁴ and Miller¹⁵). Thus, such a form must exist for every regular set S that can be represented by some CSG expression.

We shall show that the existence of the CSG canonical form is the key to constructing a solution to the B-rep to CSG problem. Furthermore, this form can be used as a basis for a deterministic CSG minimization procedure. As a result, we obtain a solution to the conversion problem that is both exact and practical.

The general problem of B-rep to CSG conversion for a solid S can be subdivided in the following three subproblems, addressed in this paper:

- Construct a set of halfspaces H that is sufficient for a CSG representation of S .
- Find a (smallest) subset of H that is both necessary and sufficient to represent S .
- Given a sufficient (and perhaps necessary) set of halfspaces H , construct a CSG representation of S that is minimal in some sense.

Consider the shaded triangle $S \subset E^2$ in Figure 4. There are many CSG representations of S using primitive halfspaces h_1 , h_2 , h_3 , and h_4 . For example

$$\begin{aligned} S &= |\Phi_1| = |h_1 h_2 h_3| \\ &= |\Phi_2| = |h_1 h_2 h_3 h_4| \\ &= |\Phi_3| = |h_1 h_2 h_3 + h_1 h_2 h_3 \bar{h}_4| \end{aligned}$$

and others. It is intuitively clear that Φ_1 is the best CSG representation because it uses the smallest number of halfspace literals, while Φ_3 is the worst of the three. These observations are made precise by noting that halfspaces h_1 , h_2 , and h_3 are both necessary and sufficient for any CSG representation of S , while

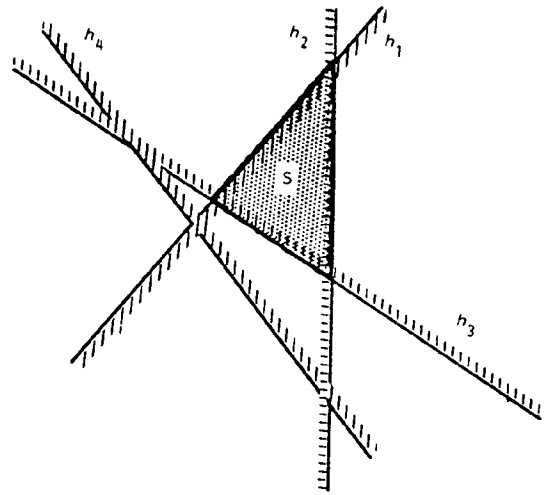


Figure 4. Planar partition by four halfplanes allows many different CSG representations of shaded triangle S

halfspace h_4 is not necessary. In addition, the CSG representation $\Phi_1 = h_1 h_2 h_3$ is the absolutely minimal representation for the triangle S .

Related work on B-rep to CSG conversion

The 2D B-rep to CSG conversion problem has been studied extensively for polygons. The most popular approach is to represent a polygon as a difference of its convex hull and a finite number of 'concavities' (e.g. Rvachev¹⁶, Tor and Middleditch¹⁷ and others). Each of the concavities is processed recursively using the same convex hull property. It is also known that all new edges used in the construction of convex hulls can be discarded, and that the resulting CSG representation has exactly one linear halfspace for every edge of the polygon¹⁸. A short and elegant proof of this fact can also be found in Dobkin et al.¹⁹, where an efficient $O(n \log n)$ convex-hull based algorithm is presented. It is worth noting that such representations generally are not minimal. Smaller representations usually can be found when some of the polygon's sides are collinear, when the solid has periodic indentation, and in other generic cases (see examples in Shapiro and Vossler⁸).

Different classes of CSG representations can be obtained using polygon decomposition techniques (recent surveys can be found in O'Rourke²⁰ and Chazelle²¹). Broadly, these techniques can be divided into partitioning methods and covering methods. Partitioning methods are used to represent a polygon as a union of non-overlapping convex pieces, but lead to unnecessarily verbose CSG representations. On the other hand, covering of a polygon with possibly overlapping pieces^{22,23} may produce relatively good CSG representations that take advantage of the polygon's collinear edges. Problems of computing various minimum covers tend to be in NP²⁴. Polygon covering methods are closely related to our CSG minimization approach, as described in a later section.

Significantly less is known about the B-rep to CSG conversion problem for curved planar solids, e.g. those bounded by circular arcs. A companion paper⁸, which uses results from the work reported here to solve the B-rep to CSG conversion problem* for curved planar solids, provides a bibliography on various restricted approaches.

We know of no significant results pertinent to 3D B-rep to CSG conversion for solids bounded by curved surfaces. We observe that, in principle, cylindrical algebraic decomposition algorithms²⁵ can be used to perform the B-rep to CSG conversion in E^d . At present, however, the practical consequences of this observation are not clear.

Even for linear polyhedra in E^3 , only limited, mainly negative, and non-constructive results are available (e.g. in Dobkin et al.¹⁹). Early attempts to extend a 2D convex-hull based algorithm to 3D polyhedra have failed²⁶. More recently such an algorithm has been found²⁷, but it relies on fairly brute-force decomposition strategies that tend to produce verbose CSG representations. It should be noted that a CSG representation of a polyhedron can be easily constructed from its binary space partition (BSP) tree representation²⁸. Procedures for computing a BSP tree representation of a polyhedron from its B-rep have been developed²⁹ and were used to show that every polyhedron has a CSG representation of size $O(n^2)$. No lower bounds on the size of a minimal CSG representation for polyhedra are known, except for certain special covers²⁰.

Finally we wish to point out an important connection between the problems considered in this paper and research on arrangements of halfspaces. The complexity of arrangements in E^d has direct bearing on both the development of practical and efficient algorithms for B-rep to CSG conversion and determining the complexity of the constructed CSG representations. Arrangements of hyperplanes have been studied extensively in Edelsbrunner³⁰. Results on arrangements of curved halfspaces are also beginning to appear³¹⁻³².

Related work on CSG minimization

Returning to the triangle in Figure 4, notice that if h_4 in Φ_2 is replaced by the universal set W , Φ_2 in fact reduces to Φ_1 . Similarly, if h_1 in the second term of Φ_3 is replaced by the empty set \emptyset , Φ_3 can be transformed into Φ_1 using the identities of Boolean algebra.

This example demonstrates the notions of W - and \emptyset -redundant primitives in CSG representations that were defined and studied by Tilove⁹. Tilove observed that a CSG representation often can be reduced in size by eliminating redundant primitives and subsequently 'pruning' the resulting CSG tree. Furthermore, redundancy can be used to formulate some very useful problems. For example, a null object detection problem of determining whether a CSG expression represents an empty set \emptyset reduces to determining whether every primitive halfspace is redundant. Since null object detection is a fundamental task useful in formulating such problems as interference detection, solids comparison, and boundary evaluation⁹, redundancy elimination has become a universal tool.

Determining whether a primitive is redundant in a CSG representation requires expensive geometric computation. Sophisticated techniques for redundancy detection relying on partial boundary evaluation have been developed and an extensive bibliography on the subject is given by Rossignac and Voelcker¹⁰. A somewhat different approach to redundancy elimination based on the notion of constituents is described in Woodward¹¹, which implicitly uses the existence of the CSG canonical form. Unfortunately, the suggested algorithms are either approximate or impractical. The redundancy of a halfspace is defined with respect to a particular CSG representation Φ , and not the set S that Φ represents. Furthermore, if $S \neq \emptyset$, the redundancy of a halfspace may be order-dependent on the redundancy of other halfspaces.

In the work reported here, we rely heavily on Boolean minimization techniques³³ for CSG reduction. We also use the fact that CSG minimization can be reduced to switching function minimization³⁴, and thus many results from switching theory (e.g. Miller¹⁵) apply. We shall see that the concepts of "halfspace necessity" and Boolean minimization lead to provably good results which subsume the notion of redundancy elimination.

Paper organization

The next section of this paper establishes that B-rep to CSG conversion is well defined by a disjunctive decomposition of W . It culminates in a 'Describability Theorem' which states the necessary and sufficient conditions for the existence of canonical CSG representations. The theorem is used in the section that follows to construct a set of primitive halfspaces that are sufficient (and perhaps necessary) to represent a regular set by a CSG expression. This step is equivalent to choosing a set of n independent Boolean variables that assure existence of canonical CSG representations.

The fourth section is devoted to the CSG minimization problem. We show that the problem is an instance of Boolean optimization and can be reduced to an essentially 'syntactic' procedure. We discuss both exact and heuristic minimization algorithms and consider geometric phenomena that are useful in understanding and improving algorithms. The required geometric utilities are discussed in the last section, which also summarizes our results and concludes with a number of open issues.

Throughout the paper we use simple 2D examples, but all results apply in E^3 unless specifically stated otherwise.

EXISTENCE OF CANONICAL CSG REPRESENTATIONS

B-rep to CSG conversion defined

We have assumed that the B-rep of a solid S is valid, i.e. represents unambiguously a boundary ∂S which determines a unique set S that is a compact, regular, and semi-analytic subset of W^1 . From the semi-analyticity of S we know that there exists a CSG representation

of S using a set of primitive analytic halfspaces $H = \{h_1, \dots, h_n\}$. Not only are there many different CSG representations of S on a fixed set H of primitives, but H itself, in general, is not unique. A set of halfspaces H can be viewed as a basis in a space of constructible CSG representations; in many ways a choice of H defines the B-rep to CSG conversion problem and, by and large, determines the success and practicality of the conversion.

For example: since S is semi-analytic, so is its boundary ∂S . Then there exists a triangulation of ∂S that can be extended to the whole of $W^{1,35}$. The union of all triangles in S constitutes a valid CSG representation of S . But, as noted in Requicha¹, such triangulations are seldom available in practice and can be very difficult to compute. While practical algorithms for triangulating linear 3D polyhedra exist, the triangulation problem is not completely understood^{20,36}. Exact triangulation of 3D curved solids is a more difficult problem.

In most practical situations S is not only semi-analytic, but *semi-algebraic*. Therefore there exists a set H of algebraic halfspaces $h_i = \{p \mid f_i(p) \geq 0\}$, where $f_i = 0$ is a polynomial surface in W , and S can be constructed using set operations on H . Furthermore, the B-rep to CSG conversion can be computed using a cylindrical algebraic decomposition algorithm²⁵.

Specifically, we call h_i a natural halfspace of a solid S if the corresponding surface $f_i = 0$ contains some face^{*} of the B-rep ∂S . Let $H_b = \{h_1, \dots, h_n\}$ be a set of all natural halfspaces, one halfspace for every face of ∂S . A set of polynomials $F = \{f_1, \dots, f_n\}$ corresponding to the natural halfspaces can be used to construct a cylindrical algebraic decomposition $K = \{c_j\}$ of W . Every cell $c_j \in K$ is a relatively open set with a property that for every $f_i \in F$, and for all $p \in c_j$, either $f_i(p) = 0$, or $f_i(p) > 0$, or $f_i(p) < 0$. Moreover, it follows from the results in Schwartz and Sharir³⁸ that the solid S is a regular cell complex, i.e. can be represented by a finite union of cells $c_j \in K$.

Cylindrical algebraic decomposition may become a practical tool for B-rep to CSG conversion. A recent survey and examples of algebraic methods for geometric algorithms can be found in Buchberger *et al.*³⁹. We will show an example of the cylindrical decomposition later. At present, this approach exhibits a number of undesirable properties:

- The cells $c_j \in K$ are represented by a set Q of polynomials computed from original natural halfspaces using the method of resultants. The degree of polynomials in Q is higher than that of the natural halfspaces, and their number is fairly large. For example, for n natural halfspaces of degree k in E^2 , Q includes $O(n^2)$ polynomials of degree $O(k^2)$. This may lead to very verbose CSG representations on relatively high degree polynomial halfspaces Q .
- For a fixed dimension d of E^d the complexity of the cylindrical decomposition algorithm is polynomial.

* Without loss of generality, we assume that a face of a solid is a "maximal face" (an m -face in the jargon of Silva³⁷) in the sense that no other face contains it. A maximal face contains all subsets of ∂S contained in the surface $f_i = 0$.

Nevertheless, the method relies on exact arithmetic and the total computational requirements of the algorithm are prohibitive.

However, the existence of a disjunctive cellular decomposition of W suggests a paradigm for the B-rep to CSG conversion, which we will customize to provide a relatively efficient conversion procedure. In particular, we observe that the natural halfspaces of a solid S define unambiguously a cellular decomposition of W and the solid S is a regularized union of d -dimensional cells in that decomposition. Therefore, B-rep to CSG conversion can be done as follows:

- Induce a decomposition of W into (open) d -dimensional cells using H_b , the natural halfspaces of S .
- Refine the decomposition, if necessary, to make all cells contained in S describable in CSG. This refinement is done by using additional halfspaces, which are called 'separating halfspaces' for reasons given later.
- Classify the cells in the refined composition with respect to the given B-rep of S to find those contained in S . The regularized union of all cells in S is a canonic CSG representation of S .

We consider these tasks in detail in the remainder of this section and in the next section. While the resulting CSG representation is still excessively verbose, it possesses important uniqueness properties necessary for the Boolean optimization algorithms described in the fourth section.

Disjunctive decompositions of W and canonical CSG representations

Given a set of halfspaces $H = \{h_1, \dots, h_n\}$, we define a *canonical intersection term* as $\Pi_k = g_1 g_2 \dots g_n$, $g_i \in \{h_i, \bar{h}_i\}$ ⁸. Since each canonical intersection term contains every halfspace h_i once there are exactly 2^n distinct Π_k s. For example, $H = \{h_1, h_2, h_3, h_4\}$ in Figure 4. Then there are exactly 16 canonical intersection terms:

$$h_1 h_2 h_3 h_4, h_1 h_2 h_3 \bar{h}_4, h_1 h_2 \bar{h}_3 h_4, \dots, \bar{h}_1 \bar{h}_2 \bar{h}_3 \bar{h}_4$$

The 2^n canonical intersection terms Π_k form a partition (i.e. an exhaustive, disjunctive decomposition) of W since

$$|\Pi_i| \cap |\Pi_j| = \emptyset, i \neq j \quad \text{and} \quad \sum_{k=1}^{2^n} |\Pi_k| = W \quad (1)$$

However, it is well known^{14,15} that a Boolean function of n variables can be written in the disjunctive canonical form (DCF). Specifically, if a set S can be represented by some CSG expression Φ on halfspaces H , it can also be represented by a unique DCF on halfspaces H

§ In set theory such a product term is often called a constituent^{11,14}

computed as follows

$$S = |\Phi(h_1, \dots, h_n)|$$

$$= \left| \sum_{e=(e_1, \dots, e_n)} \Phi(e_1, \dots, e_n) h_1^{e_1} \dots h_n^{e_n} \right| \quad (2)$$

where $e_i = 0$ or 1 , $h_i^0 = \bar{h}_i$, $h_i^1 = h_i$, $e = (e_1, e_2, \dots, e_n)$ is an n -tuple of 0s and 1s, and the summation extends over all combinations of n 0s and 1s for the e s. For sets, 0 corresponds to an \emptyset , and 1 corresponds to W . Thus, $|\Phi(e_1, \dots, e_n)|$ is always either \emptyset , or W .

Rewriting Eq. (2) in terms of canonical intersection terms, we get

$$S = W \cdot \sum_{i=1}^{n_{in}} |\Pi_i^{in}| + 0 \cdot \sum_{i=1}^{n_{out}} |\Pi_i^{out}| = \sum_{i=1}^{n_{in}} |\Pi_i^{in}|, \quad (3)$$

where n_{in} is the number of canonical terms in S , n_{out} is the number of canonical terms in \bar{S} , and $n_{in} + n_{out} = N \leq 2^n$.

Thus canonical intersection terms can be viewed as the smallest spatial building blocks that can be 'glued' together to describe any solid in W representable by a CSG on H . This is particularly natural and appealing when all halfspaces h_i are hyperplanes. A partition of W by n hyperplanes is usually called an arrangement; every canonical intersection term Π_k represents either a convex connected cell in W , or an empty set \emptyset . It is well known that there are at most $O(n^d)$ nonempty d -dimensional cells in any given arrangement³⁰. Thus for a large n , most of the 2^n canonical intersection terms represent empty sets. In Figure 4 only eleven of the sixteen Π_k s are not empty.

Consider now a situation when H includes halfspaces whose boundaries are not hyperplanes. Figure 5 shows a simple $n = 2$ example in E^2 . It is clear that properties (1) hold. However, the nature of the canonical intersection terms has changed: obviously they may now represent disconnected sets (as Π_2 and Π_3 do).

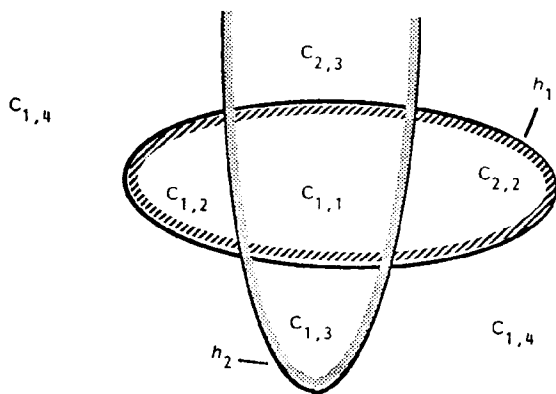


Figure 5. Planar partition by two curved halfspaces. Canonical intersection terms $\Pi_2 = h_1 \bar{h}_2$ and $\Pi_3 = \bar{h}_1 h_2$ represent disconnected sets

Thus

$$|\Pi_k| = \sum_{m=1}^{M_k} C_{m,k} \quad (4)$$

where $C_{m,k}$ is m th connected set, or component, of the $|\Pi_k|$, and M_k is the total number of components in $|\Pi_k|$. It is always true that the number of non-empty canonical intersection terms $N \leq M = \sum_{k=1}^N M_k$.

We shall assume that halfspaces in H are such that $M \ll 2^n$, and therefore we can effectively inspect all components in any partition of W . This assumption rules out general analytic halfspaces that may result in an infinite number of components. On the other hand, for n algebraic halfspaces of degree k the assumption is reasonable because there can be at most $(kn)^{O(d)}$ components in the CCD of E^d (e.g. see Renegar⁴⁰). Tighter bounds on complexity of arrangements of curves and surfaces are only beginning to appear, for example in Clarkson *et al.*³² and Edelsbrunner *et al.*³¹.

Substituting equation (4) into (1), we get

$$W = \sum_{k=1}^N |\Pi_k| = \sum_{k=1}^N \left(\sum_{m=1}^{M_k} C_{m,k} \right) \quad (5)$$

an exhaustive disjunctive connected component decomposition (CCD) of W . Geometrically, each $C_{m,k}$ is the largest (semi-) bounded set of points in the partition of W whose interior does not contain any halfspace boundaries. Similarly we can rewrite equation (3) as

$$S = \sum_{i=1}^{n_{in}} \left(\sum_{m=1}^{M_i} C_{m,i} \right) \quad (6)$$

which states that any regular set S can be always represented by a union of components in the partition of W by H . (It is worth noting that the validity of the B-rep guarantees that every component $C_{m,i}$ is bounded.) Furthermore, as for a fixed set of halfspaces H equation (6) defines a unique cellular decomposition, the B-rep to CSG conversion is now well defined. It reduces to finding a (non-unique) CSG representation for each $C_{m,i}$ in equation (6) in the partition of W by natural halfspaces H_b .

Describability theorem

Definition: a set S is describable by halfspaces $H = \{h_1, \dots, h_n\}$ if there exists a CSG expression $\Phi(h_1, \dots, h_n)$ that represents it.

We want to determine the conditions under which S is describable by H , based on the fact that any CSG representation of S can be written in a DCF given by Equation (3). We assume that

$$\partial S \subseteq (\partial h_1 \cup \partial h_2 \dots \cup \partial h_n), \quad (7)$$

since it is easy to prove that S is not describable by H if this condition is not satisfied (also see Requicha and Voelcker¹³ and Requicha and Tilove²).

Any two points that are in the same component C must have the same classification with respect to every halfspace h_i . (But it is not always true that any two points that have the same classification with respect to every halfspace belong to the same component.) This defines an equivalence class consisting of all the points in the interior of the same component C . Therefore we can talk about a specialized form of component classification with respect to a halfspace h_i , or with respect to any set S satisfying (7). Namely, C in S can have only two values: if $C \subseteq S$, C in $S = C$, in which case we say that the result of the classification is true, or if $C \subseteq \bar{S}$, C in $S = \emptyset$, and we say that the result of the classification is false.

Before we state and prove our main result on descriptibility of solids, let us consider an illustrative example. Figure 6 shows a planar solid S bounded by linear edges and circular arcs. By extending the solid's edges we obtain the planar CCD partition of E^2 . The components A and B have identical classifications with respect to every halfplane; therefore they belong to the same canonical intersection term and are not describable separately. Yet $A \subset S$ while $B \subset \bar{S}$. It is intuitively clear that the solid S is not describable by the halfspaces shown in Figure 6. Thus we have the following theorem.

Describability Theorem: given a set of halfspaces $H = \{h_1, \dots, h_n\}$ and a regular set S satisfying (7), S is describable by H if and only if all components $C_{m,k}$ of every canonical intersection term $|\Pi_k|$ have the same classification with respect to S .

Proof: The boundary ∂S partitions W into two disjoint sets S and \bar{S} .

'If' part. Assume that components $C_{m,k}$ of every term $|\Pi_k|$ are either all in S or all in \bar{S} . Then $C_{m,k} \subseteq |\Pi_k| \subseteq S$, for all $C_{m,k} \subseteq S$. Therefore, union of all canonical intersection terms $|\Pi_k| \subseteq S$ is a valid CSG representation for S , given by equation (3).

'Only if' part. Proof by contradiction. Assume that S is describable by H . Without loss of generality, suppose there are two components $A, B \subset |\Pi_k|$ with $A \subset S$ and $B \subset \bar{S}$. By assumption and equation (3), S can be represented as a union of intersection terms $|\Pi_k| \subseteq S$. But $A \subset |\Pi_k| \not\subseteq S$. Thus we have a contradiction. And so every pair of components A and B of the same canonical intersection term $|\Pi_k|$ must have the same classification with respect to S . \square

While every canonical intersection term $|\Pi_k|$ represents some d -dimensional (possibly empty, or disconnected) set, an individual component $C_{m,k}$ may not be describable. This is unfortunate since it seems natural to use $C_{m,k}$ (and not $|\Pi_k|$) as a basic 'building block'. Apparently, we have a representational deficiency: we may not be able to represent each individual component using CSG on natural halfspaces in H_b . The representational variety of the CCD (equation (6)) exceeds the representational capacity given by DCF of equation (3). Specifically, CCD defines 2^M distinct nonempty sets $S \subset W$, but there are only 2^N sets that

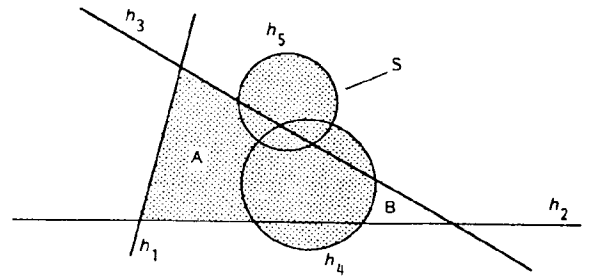


Figure 6. Shaded solid S is not describable by $H = \{h_1, h_2, h_3, h_4, h_5\}$ because components A and B are represented by the same canonical intersection term

can be represented by a finite union of canonical intersection terms.

CONSTRUCTION OF NECESSARY AND SUFFICIENT HALFSACES

Sufficient set of halfspaces

If the boundary of every natural halfspace $h_i \in H_b$ is a hyperplane, every non-empty canonical intersection term $|\Pi_k|$ represents a convex (connected) polyhedral set. Thus $|\Pi_k| = C_{1,k}$ for all k , and every $C_{1,k}$ is describable by H_b . This leads to the well-known conclusion that any polyhedral solid is describable by planar halfspaces associated with its faces.

In a more general setting with curved halfspaces, the Describability Theorem establishes the necessary and sufficient conditions for a solid S to be describable by a set of natural halfspaces H_b . More importantly it gives a practical test (practical because of our assumption about the number of components in CCD of E^d) to determine if halfspaces in H_b are sufficient to represent S . If at least two components $A, B \subset |\Pi_k|$ are separated by a solid's boundary, H_b is not sufficient to describe the solid S , and it is only natural to try to augment H_b with an additional halfspace g , chosen so that $A \subset g$ and $B \subset \bar{g}$. Addition of such a halfspace g , if possible, effectively separates the canonical intersection term $|\Pi_k| = A + B$ into two new canonical intersection terms $|\Pi_k g| = A$ and $|\Pi_k \bar{g}| = B$. We will call such a halfspace g a separating halfspace. Hence the following problem:

Given a set S and a set of natural halfspaces H_b induced from ∂S , construct a set of separating halfspaces H_s such that S is describable by $H = H_b \cup H_s$.

Consider a simple example in Figure 7. The three components A, B and C are all subsets of the canonical intersection term $\Pi = h_1 \bar{h}_2 \bar{h}_3 h_4 h_5 h_6$, with $A, B \subset \bar{S}$ and $C \subset S$ (Figure 7(a)). Figure 7(b) shows that adding four separating halfspaces is sufficient for a CSG representation of the planar solid S . It should be clear that the four halfspaces are not unique. Figure 7(c) demonstrates that adding two (properly chosen) separating halfspaces to H_b is also sufficient to represent S . Finally, in

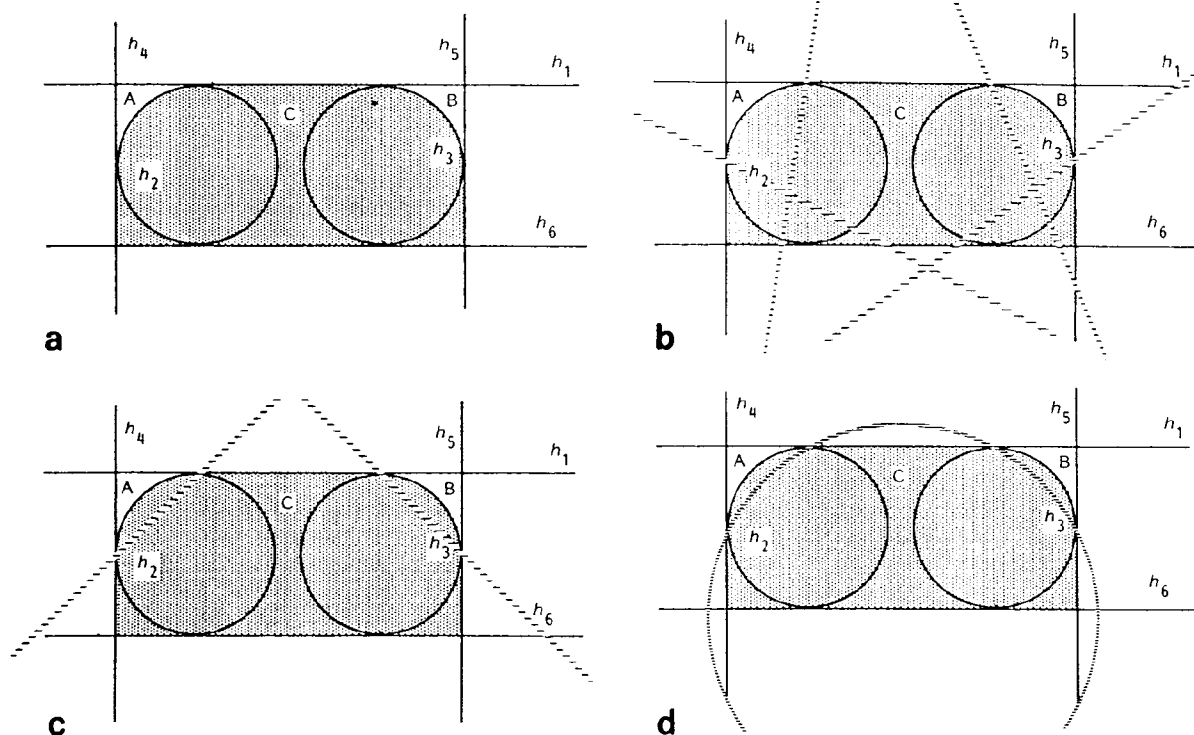


Figure 7. (a) Halfspaces h_1, \dots, h_6 are not sufficient to represent the shaded figure S ; (b) addition of four linear separating halfspaces is enough to represent S ; (c) addition of two linear separating halfspaces is sufficient to represent S ; (d) addition of one circular separating halfspace is sufficient to represent S

Figure 7(d) it is seen that a sufficient set of halfspaces can be obtained by addition of a single circular halfspace.

This simple example demonstrates a number of important properties of separating halfspaces:

- the set H_s of separating halfspaces is not unique
- separation of several pairs of components may be achieved simultaneously (as is the case in Figure 7(d))
- using higher order separating halfspaces may reduce the total number of required separating halfspaces
- boundaries of separating halfspaces may intersect components that are being separated (e.g. in Figure 7(b) component C is intersected by all four separating halfspaces)
- boundaries of separating halfspaces do not 'contribute' to ∂S

In what follows we compare several alternative strategies for constructing a sufficient set H_s of separating halfspaces.

Strategies for separation of components

Before we consider possible approaches to construction of H_s , let us point out some properties of components that are useful in the construction of separating halfspaces. Consider two components A and B separated by some linear halfspace $g \in H$ so that $A \subseteq g$ while $B \subseteq \bar{g}$. Clearly, A and B cannot be components

of the same intersection term Π_k . The following property is trivial to prove.

Hyperplane Separation Property: If a line segment ab connecting two points $a \in A$ and $b \in B$ intersects some hyperplane $\hat{c}g$, $g \in H$, then A and B are not components of the same canonical intersection term Π_k . \square

More generally, consider a case when every halfspace $h_i \in H$ is convex. Then a canonical intersection term Π_k can be written as

$$\Pi_k = h_1 \dots h_i \bar{h}_{i+1} \dots \bar{h}_n$$

Intersection of the first i halfspaces is a convex set which is connected by definition. Thus disconnected components arise only due to intersections of halfspace complements that are concave.

Suppose that the set S is not describable by natural halfspaces H_b , because there is a single canonical intersection term $|\Pi| = A + B$, with $A \subseteq S$ and $B \subseteq \bar{S}$. Let us assume that A and B can be separated by a hyperplane, or by a polyhedral surface $P \subseteq (\hat{c}g_1 \cup \dots \cup \hat{c}g_k)$, where g_i , $1 \leq i \leq k$, are linear halfspaces. Then, by the hyperplane separation property, all points of A are separated from all points of B by some halfspace g_i . By the descriptibility theorem, S is describable by $H_b \cup \{g_1, \dots, g_k\}$.

Unfortunately, the construction of a polyhedral separating surface is not always possible, because

components of the same intersection term $|\Pi_k|$ may share common edges (but not faces). If the shared edge is not planar, the separating halfspace cannot be linear because its boundary surface must contain the shared edge. On the other hand, in many practical situations natural halfspaces H_b are such that separating hyperplanes are always sufficient, while addition of curved separating halfspaces may reduce the number of needed halfspaces⁹.

If S is a semi-algebraic set, we know that a sufficient H_s exists and can be computed using the cylindrical algebraic decomposition algorithm²⁵. Because the algorithm itself is not acceptable for reasons discussed in the earlier subsection on 'B-rep to CSG conversion defined', we must consider alternative strategies for separation of components. Consider a planar solid S in Figure 8(a). The set of natural halfspaces H_b contains three halfspaces: a circular disk h_1 , h_2 which consists of the two disconnected regions bounded by a hyperbola, and a linear halfspace h_3 . S is not describable by $\{h_1, h_2, h_3\}$ because $|\Pi_1| = |\bar{h}_1 \bar{h}_2 h_3| = A + B$, with

$A \subset S$ and $B \subset \bar{S}$. We consider three alternative separation schemes:

- *Test-based separation*: a sufficient set of halfspaces can be constructed by adding to H_b one separating halfspace at a time, until the conditions of the Describability Theorem are satisfied. In our example this is fairly simple (e.g. Figure 8(b)). However, to use this strategy, we usually must determine the boundaries of components A and B and then compute a set of halfspaces that separates them. This problem can be more difficult than a motion planning problem.
- *Global separation*: the idea is simple: construct a set of separating halfspaces H_s such that every canonical intersection term Π_k in the decomposition of W by $H_s \cup H_b$ represents a single component. The cylindrical algebraic decomposition can achieve this. Figure 8(c) shows a cylindrical decomposition for h_1, h_2, h_3 modified from Buchberger et al.³⁹. It is not clear whether better general algorithms can be found for

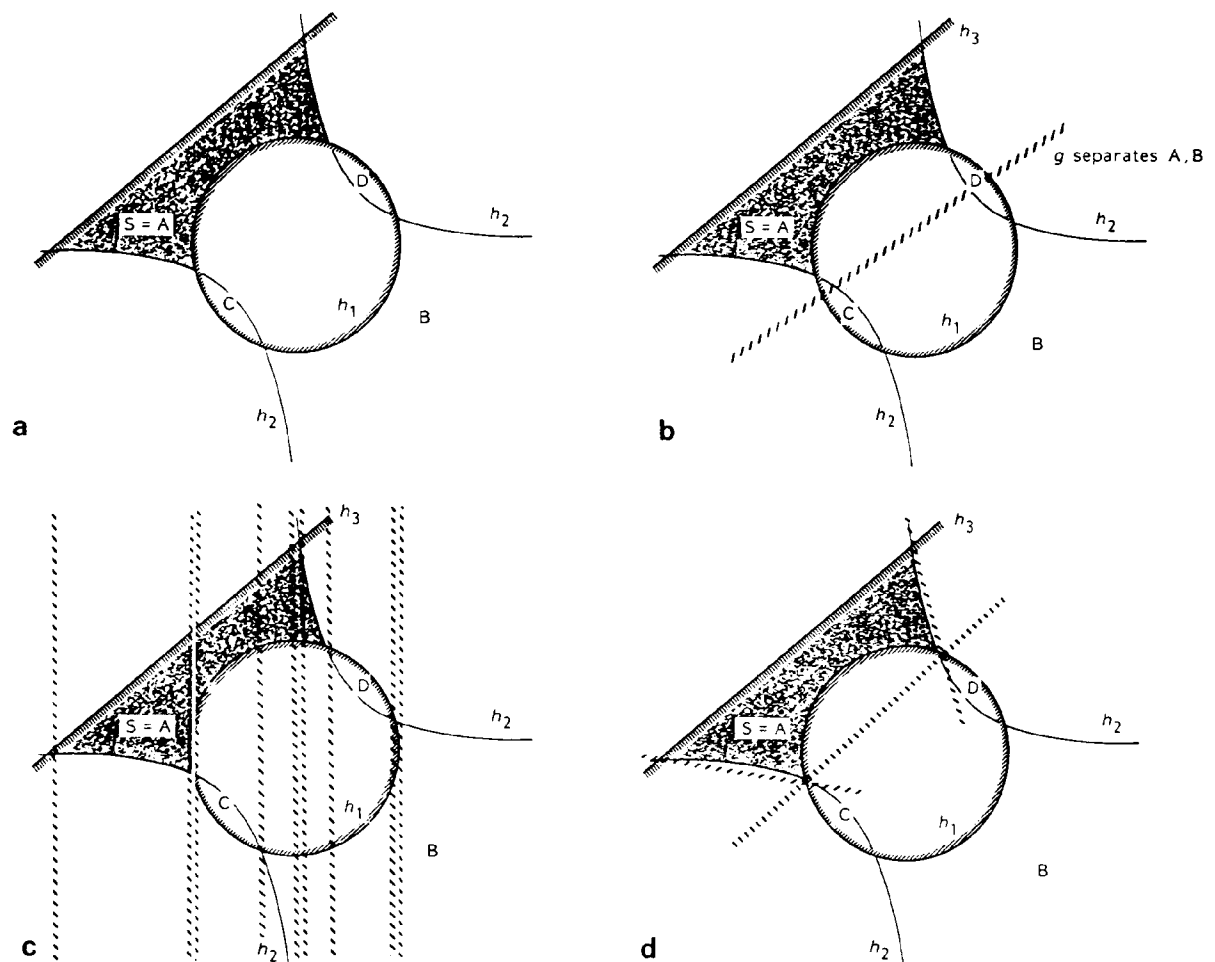


Figure 8. Different separation strategies: (a) solid $S = A$ is not describable by $\{h_1, h_2, h_3\}$ because $|\bar{h}_1 \bar{h}_2 h_3| = A + B$; (b) separating A and B by a set of halfspaces can be viewed as a generalization of a motion planning problem; (c) cylindrical algebraic decomposition for h_1, h_2, h_3 (additional halfspaces are required for global separation); (d) boundary-based separation in E^2 can be achieved using linear chordal halfspaces

particular types of natural halfspaces, e.g. second-degree curves and surfaces. It is also worth noting that only bounded d -dimensional cells need be represented for B-rep to CSG conversion, because regularized set operations assure the compactness of the resulting object.

- **Boundary-based separation:** observe that $|\Pi_2| = |h_1 h_2 h_3| = C + D$. Yet we do not need to separate components C and D because both are outside of S . Thus, we only need to separate components inside S from components outside of S . This observation suggests that the boundary of a solid can be used to construct a sufficient H_s which is potentially much smaller than the one resulting from the global separation. Figure 8(d) shows such a construction. We have proven⁸ that the set of all 'chordal' halfspaces is a sufficient set H_s of separating halfspaces for a large class of planar objects.

The test-based separation strategy is likely to produce a much smaller set of separating halfspaces than either global or boundary-based separation. However, as we already pointed out, this approach may be computationally prohibitive. The advantages of the global and the boundary-based separation strategies are that a set H_s is computed *a priori*, and that they do not require any 'planning' algorithms depending on knowledge of the boundaries of components. Furthermore, the boundary-based strategy is likely to produce a much smaller set H_s of separating halfspaces than the global strategy. None of these strategies are guaranteed to produce a H_s which is minimal, or indeed even necessary.

Necessary and minimal sets of separating halfspaces

Given a set of halfspaces H_s such that set S is describable by $H = H_b \cup H_s$, it is not difficult to compute a set $H^* \subseteq H$ that is necessary and sufficient for a CSG representation of S , if we have a representation of every $|\Pi_k|$.

Definition: A halfspace g is necessary in H if S is not describable by $H - \{g\}$.

Every natural halfspace $h_i \in H_b$ is necessary in H by construction. To test whether a separating halfspace $g_k \in H_s$ is necessary in H we compute all pairs of canonical intersection terms (Π_i, Π_j) such that

- $|\Pi_i|$ and $|\Pi_j|$ have different classifications with respect to S
- $|\Pi_i|$ and $|\Pi_j|$ have different classifications with respect to g_k
- $|\Pi_i|$ and $|\Pi_j|$ have identical classifications with respect to all other halfspaces $h \in H - \{g_k\}$

If there are no such pairs (Π_i, Π_j) , halfspace g_k is not necessary in H ; we can remove it from H_s , and then consider the remaining halfspaces. When all halfspaces in H_s are tested, the set of remaining separating halfspaces H_s^* is necessary in $H^* = H_b \cup H_s^*$. Therefore we conclude that H^* is both necessary and sufficient

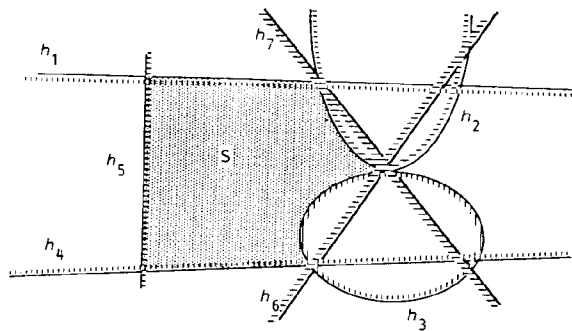


Figure 9. Only one of $\{h_6, h_7\}$ is necessary to represent S . Prime implicants $h_1 \bar{h}_2 \bar{h}_3 h_4 h_5 h_6$ and $h_1 \bar{h}_2 \bar{h}_3 h_4 h_5 h_7$ represent the same shaded set S

for a CSG representation of S . This approach has been used successfully by Shapiro and Vossler⁸.

With such a definition, the 'necessity' of a halfspace is not absolute but is dependent on the presence of other separating halfspaces. For example, in Figure 9 neither h_6 nor h_7 is necessary in $H = \{h_1, \dots, h_7\}$ for a CSG representation of the shaded solid S , but h_6 is necessary in $\{h_1, \dots, h_6\}$ while h_7 is necessary in $\{h_1, \dots, h_5, h_7\}$. Thus the necessary set of halfspaces is not unique, and we may look for a *minimal* set H_s^* . The complexity of this problem is not well understood, though it is easy to map it into a set cover problem which is known to be NP-complete. A 'greedy' algorithm to minimize the size of H_s^* is described in Shapiro and Vossler⁸.

MINIMIZATION OF CSG REPRESENTATIONS

Types of minimization problems

Let us assume $H = \{h_1, \dots, h_n\}$ is a set of halfspaces that is sufficient for CSG representation of a set S . A CSG representation on H is not unique for S , and this section is devoted to the problem of computing a minimal CSG representation. A CSG representation of S is minimal if any other CSG representation of S involves an equal or larger number of regularized union and intersection operations. Since k binary operations operate on exactly $k+1$ halfspace literals, we may alternatively state the problem in terms of number of halfspace literals.

Given a set of halfspaces H that is sufficient to represent a set S , construct $\Phi(H)$ with the smallest possible number of halfspace literals, such that $S = |\Phi|$.

Such a formulation of the problem precludes comparison of CSG representations on distinct sets of halfspaces. For example, we would not compare CSG representations of the same solid S on different (but all sufficient) sets of halfspaces shown in Figures 7(b)–(d). The relationship between the *number* and *type* of halfspaces in H and the size of the minimal CSG representation is not known. O'Rourke gives an example where addition of halfspaces that are not necessary to represent a

polygon leads to a smaller size cover of that polygon³⁴. In other words, suppose H_1, \dots, H_m are different sets of halfspaces, each of which is sufficient to describe S , and let H_i be the smallest among them. (For instance, the set of halfspaces in Figure 7(d) is the smallest possible and contains only seven halfspaces.) It is possible that a minimal CSG representation on H_i is larger than a minimal CSG representation on H_j , $j \neq i$.

For convenience we shall assume that H is a set of halfspaces that is both sufficient and necessary to represent S . Knowing that every $h_i \in H$ must appear in every CSG representation of a set S allows us to reduce the CSG minimization problem to a problem of minimizing a number of occurrences of every halfspace h_i . Ideally, we would like to compute a CSG representation with every halfspace literal used exactly once, but this may not be possible even for very simple objects⁸.

Broadly speaking, the defined CSG minimization problem can be viewed as an instance of multilevel Boolean function minimization. The general problem is discussed in Lawler³³, where it is shown that a two-level[†] Boolean minimization is an important special case of the general problem. Both problems are known to be NP. We consider both minimization problems for CSG representations, after establishing an analogy between CSG representations and switching functions.

Inclusion tests for CSG representations

Minimization of switching functions has been studied extensively in the literature^{35,41}. The pioneering work of Pavlidis²² and O'Rourke²³ suggest that two-level minimization techniques from switching theory may be applicable to minimization of CSG representations. Table 1 summarizes the correspondence between general Boolean algebra, regular sets and regularized set operations, and switching functions.

Inevitably, any minimization approach relies on the ability to compare different Boolean forms using one or more inclusion tests. By definition, for any two elements a and b , $a \subseteq b$ if and only if $a \cdot b = a$.

A very special property of switching functions is that there are only two possible elements: 0 (false) and 1 (true). Thus, inclusion becomes formal implication, i.e. $|\Phi_1|$ implies $|\Phi_2|$ if and only if every truth assignment that makes $|\Phi_1| = 1$ also makes $|\Phi_2| = 1$. In particular, when both Φ_1 and Φ_2 are product terms, the implication test reduces to a pure syntactic test: $|\Phi_1|$ implies $|\Phi_2|$ if and only if every literal in Φ_1 also appears in Φ_2 .

Broadly speaking, there is no analogous syntactic test for CSG representations (but see O'Rourke²³). Given two representations Φ_1 and Φ_2 , the only way to determine whether $|\Phi_1| \subseteq |\Phi_2|$ is by a geometric test that may be computationally expensive. We show now that in order to perform inclusion tests on any number of CSG representations Φ_1, \dots, Φ_m , defined on the same fixed set of halfspaces H , the necessary geometric computations need be done only once.

Table 1. Boolean algebras

Boolean algebra	Regular sets (CSG)	Switching functions
Element	Set of points in W	Element $e \in \{0, 1\}$
Addition $+$	Regularized union \cup^*	Logical OR \vee
Multiplication (\cdot)	Regularized Intersection \cap^*	Logical AND \wedge
Complement \bar{x}	Regularized complement \bar{x}	Logical NOT \neg
Inclusion	Subset \subseteq	Implication
0-element	Empty set \emptyset	0
1-element	W	1

Suppose a set S is represented by a CSG representation $\Phi(H)$. There is a total of N non-empty canonical intersection terms Π_k in the partition of W by H . Since S is describable by H , only two types of such terms are possible.

- with all components in S , i.e. $|\Pi_k^{\text{in}}| \subseteq S$
- with all components out of S , i.e. $|\Pi_k^{\text{out}}| \subseteq \bar{S}$

Then the DCF for Φ given by equation (3) represents the unique decomposition for a set S defined by a CSG representation on H . It provides a basis for comparing any two CSG representations $\Phi_1(H)$ and $\Phi_2(H)$.

Given an arbitrary CSG representation $\Phi(H)$ and the disjunction canonical decomposition of W by H , it is easy to compute the DCF of equation (3) for $|\Phi|$. This is achieved by performing a test $|\Pi_k| \subseteq |\Phi|$ for every nonempty canonical intersection term Π_k , $k = 1, \dots, N$. It is shown in O'Rourke^{23,34} that the latter test is quite simple. Each Π_k corresponds to a 'truth assignment' in the following sense. Pick an arbitrary point $p \in |\Pi_k|$. If $p \in h_i$, then assign $h_i = 1$; otherwise assign $h_i = 0$. Thus, a single point simultaneously and uniquely assigns truth to all n halfspaces $h_i \in H$. Then, the inclusion test reduces to a formal implication test

$$|\Pi_k| \subseteq |\Phi| \text{ if and only if } \Pi_k \text{ implies } \Phi$$

In particular, when Φ is a product of halfspaces, the inclusion (implication) test is a syntactic procedure identical to the one used in switching theory.

Consider now two CSG representations $\Phi_1(H)$ and $\Phi_2(H)$. Using the above procedure we can compute sets $|\Phi_1|$ and $|\Phi_2|$ in the DCF of equation (3). Then the inclusion relation can be redefined in the following fashion

$$|\Phi_1| \subseteq |\Phi_2| \text{ if and only if}$$

$$|\Pi_k| \subseteq |\Phi_1| \text{ implies}$$

$$|\Pi_k| \subseteq |\Phi_2|, \text{ for all } 1 \leq k \leq N$$

So if all non-empty canonical intersection terms Π_k are known, any inclusion test (as well as the intersection and the equality tests) reduces to a number of formal implication tests.

Two-level CSG representations

We first consider minimization of sum-of-products, or disjunctive normal form (DNF), CSG representations.

[†] By 'two-level' we mean either sum-of-products or product-of-sums expressions

By duality, similar results apply to the conjunctive normal form minimization. This problem has been partially addressed for polygons in Pavlidis²² and O'Rourke²³.

If a CSG representation is written as a sum of product terms, but not every product term contains every halfspace literal h_i , we say that the representation is a disjunctive normal form (DNF). A DNF is not unique in the sense that many different DNFs can have the same DCF. We say that a DNF is minimal for a set S if there is no other DNF that represents S with a smaller number of halfspace literals.

One DNF representation is immediately given by the DCF of equation (3). For example, Figure 10 shows a polygon whose CSG representation can be obtained as $S = \Pi_1 + \dots + \Pi_6$ or, using halfspaces $H = \{a, b, c, d, e, f\}$,

$$S = (abc\bar{d}ef) + (abc\bar{d}ef) + (a\bar{b}c\bar{d}ef) + (a\bar{b}cdef) + (a\bar{b}c\bar{d}ef) + (a\bar{b}cdef)$$

While the above representation is easy to construct, it results in a very inefficient CSG representation for S . One of the reasons for this inefficiency is that a canonical intersection term Π_k is the most verbose intersection representation for a set of components given by $|\Pi_k|$. The same set can be described by the intersection of only those halfspaces that form the boundaries of $|\Pi_k|$, leading to its minimal, or most terse, CSG representation²⁴. For example, $\Pi_1 = abc\bar{d}$, $\Pi_2 = abcf$, and so on. Even more importantly, a number of terms $|\Pi_k|$ could be subsets of a larger set $P \subseteq S$ formed by an intersection of some halfspaces in H . Thus, for the polygon in Figure 10, Π_2 , Π_3 , and Π_4 are all subsets of $P = |acef|$.

DNF minimization is based on the notion of *prime implicants* that are defined, for a CSG-represented set S , as minimal CSG representations of the largest intersection subsets of S . More precisely, Ψ is a prime implicant of S if:

- Ψ is an implicant of S , i.e. $\Psi = x_1 \dots x_k$, $x_i \in \{h_i, \bar{h}_i\}$, and $|\Psi| \subseteq S$
- $|\Psi_i| = |x_1 \dots x_{i-1} x_{i+1} \dots x_k| \not\subseteq S$, $\forall x_i$, $A \leq i \leq k$, i.e. when any halfplane literal x_i is deleted, the remaining product term is no longer an implicant of S

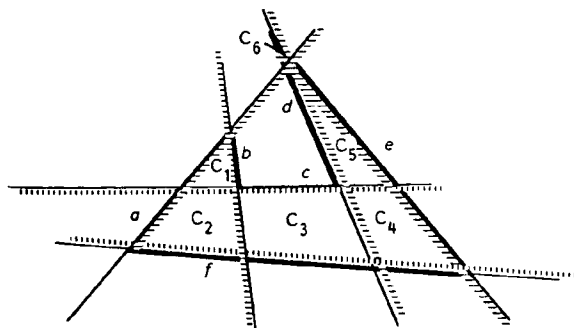


Figure 10. Polygon $S = C_1 + \dots + C_6$ has a minimal DNF using three prime implicants: $S = |abf + acef + def|$

The set of all prime implicants is unique for a given S and a set of halfspaces H , and their union is sufficient (but not always necessary) to represent S . It is well known that a minimal DNF is a sum of some implicants. For the polygon in Figure 10, the set of all prime implicants is: $\{abf, acef, def, ab\bar{c}, de\bar{c}, ac\bar{d}, \bar{b}cef, \bar{b}cd\bar{f}, \bar{a}de\}$.

It was long believed that for polygons the minimal DNF contains prime implicants formed by only positive (uncomplemented) halfspace literals^{22,23}, which could significantly improve the computation. Indeed the minimal DNF for the polygon in Figure 10 is given by

$$S = abf + acef + def$$

It was recently shown that this is not always true even for simple polygons.

While procedures for computing prime implicants may be adapted from switching theory^{43,44}, and then used for computation of a minimal DNF form²³, they may be inefficient, because they require many inclusion tests, and because the number of prime implicants can be very large. Figure 11 shows that even a simple 2D polygon can have an exponential number of prime implicants⁴⁵. (See also Aggarwal et al.⁴² for related examples and discussion.)

For planar polygons, every prime implicant is a unique representation of the polygon's subset. However, this is not true when H contains separating halfspaces or additional halfspaces that are not necessary. Figure 9 shows an example where two separating halfspaces h_6 and h_7 have been added to represent the shaded figure (only one of them is necessary). Observe that there are two distinct representations for the shaded figure: $h_1\bar{h}_2\bar{h}_3h_4h_5h_6$ and $h_1\bar{h}_2\bar{h}_3h_4h_5h_7$. Both are minimal, therefore both are prime implicants.

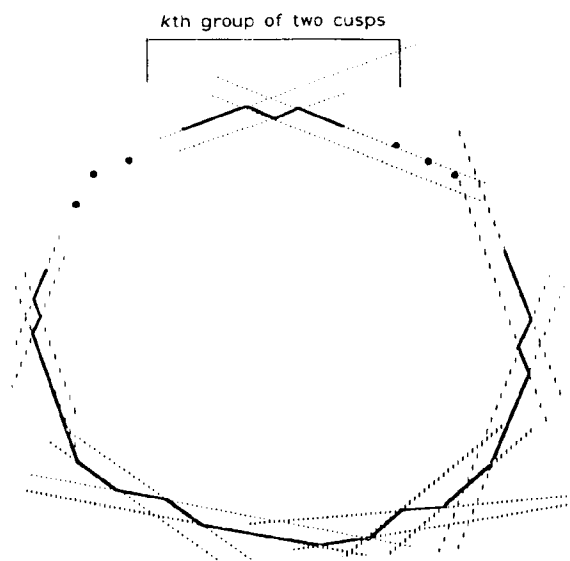


Figure 11. Each group of four edges of the polygon form two 'cusps'. A prime implicant is formed by choosing one of the two cusps from every group. Therefore, there is a total of $2^{(n/4)}$ prime implicants of this polygon

Even if all prime implicants are computed, choosing the smallest subset of them that covers the set S requires an exhaustive search, and it is not likely that more efficient algorithms can be found²⁴. Many algorithms that compute approximate minimal covers are known, e.g. Miller¹⁵. One such algorithm has been described and implemented by Shapiro and Vossler⁸. Additional improvements in both the quality of the computed cover and in the performance of the algorithm can be achieved by taking advantage of special prime implicants called 'dominating halfspaces' that commonly occur in solid models and are described in the next section.

Dominating halfspaces

Suppose that a regular set S is describable by a set of halfspaces H and for some halfspace $g \in H$ the following condition holds

$$S = g + S \quad (8)$$

We then say that a halfspace g dominates set S , or that g is a dominating halfspace. The following statements are easily verified to be equivalent:

- halfspace g dominates set S
- g is a prime implicant of S
- \bar{g} appears in every canonical intersection term Π_k such that $|\Pi_k| \subset \bar{S}$
- \bar{g} is a supporting halfspace of \bar{S} , i.e. $\bar{S} \subseteq \bar{g}$

The second property guarantees that there is a CSG representation of S in which g appears only once, the third property gives an efficient way to compute all the dominating halfspaces of S , while the fourth property gives an intuitive geometric interpretation of dominating halfspaces and explains why they are a frequent phenomenon in solid models of mechanical objects.

For example, consider a planar solid in Figure 12(a). Halfspace e dominates S , while halfspace \bar{a} dominates \bar{S} .

We now show that if equation (8) holds, the CSG minimization for S reduces to the CSG minimization for $S_1 \subseteq S$ on a smaller set of halfspaces $H - \{g\}$. Furthermore, the DCF for S_1 is easily computed from the canonical forms for S . Let us rewrite DCF of equation (3) as

$$S = \left| \sum_{i=1}^p \Pi_i^g + \sum_{j=1}^q \Pi_j^{\bar{g}} \right| \quad (9)$$

where Π_i^g are all canonical terms with g , and $\Pi_j^{\bar{g}}$ are all remaining terms with \bar{g} . Since g dominates S , $S = g + S$. But $g + \Pi_j^{\bar{g}} = g$, and so

$$S = g + S = g + \left| \sum_{j=1}^q \Pi_j^{\bar{g}} \right| \quad (10)$$

Using De Morgan's laws

$$\begin{aligned} g + \Pi_j^{\bar{g}} &= g + (h_1 \dots h_{n-1} \bar{g}) = g + h_1 \dots h_{n-1} \\ &= g + \Pi_j^* \end{aligned} \quad (11)$$

where Π_j^* is a canonical intersection term in the partition of W by $H - \{g\}$. Combining equations (10) and (11) gives

$$S = g + \left| \sum_{j=1}^q \Pi_j^{\bar{g}} \right| = g + \left| \sum_{j=1}^q \Pi_j^* \right| = g + S_1 \quad (12)$$

Thus DCF for S_1 is obtained from DCF for S by deleting from equation (9) all canonical intersections terms Π_j^g and dropping literal \bar{g} from all remaining terms $\Pi_j^{\bar{g}}$. This is consistent with the fact that S_1 is describable by the

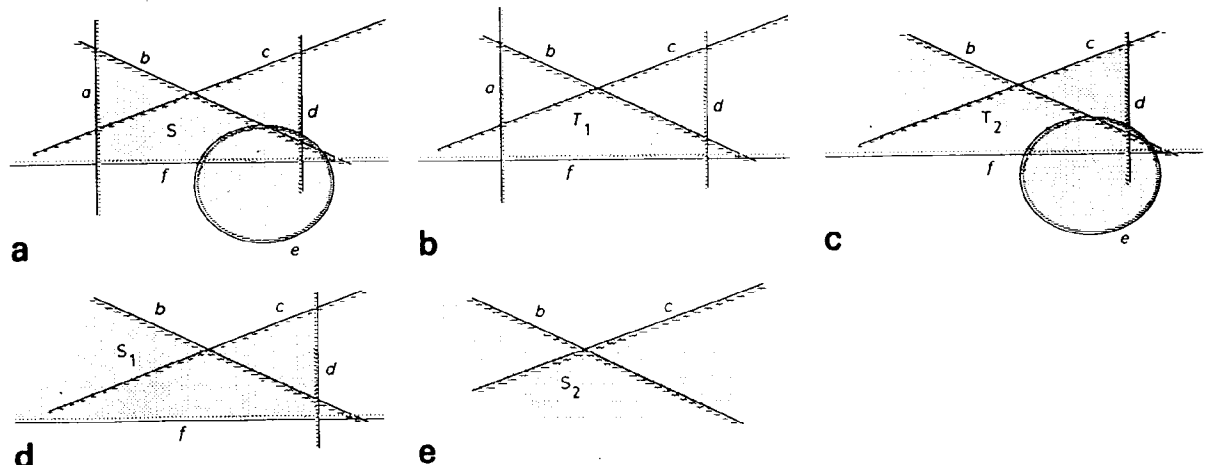


Figure 12. Solid S is completely represented by dominating expressions: $S = |a(e + df(b + c))| = |e + adf(b + c)|$: (a) halfspace \bar{a} dominates S , and halfspace e dominates S ; (b) Decomposition using e : $S = e + T_1$; (c) Decomposition using a : $S = a + T_2$; (d) Decomposition using a and e simultaneously: $S = a(e + S_1) = e + aS_1$; (e) Decomposition of S_1 with d and f : $S_1 = df S_2$. Decomposition of S_2 with b and c : $S_2 = b + c$

set of halfspaces $H - \{g\}$.[‡] Figure 12(b) shows the result of decomposition using dominating halfspace e geometrically.

Similarly, if a halfspace h dominates \bar{S} we can write S as

$$S = \bar{S} = \overline{h + \bar{S}_1} = \bar{h} \bar{S}_1 \quad (13)$$

and the minimization needs to be carried out only for $S_1 \supseteq S$, again with a smaller set of halfspaces $H - \{h\}$. Such a decomposition is shown in Figure 12(c), where halfspace \bar{a} dominates \bar{S} .

Absolutely minimal CSG representations

A minimal two-level CSG representation may not be truly minimal as defined in the subsection, 'types of minimization problems'. For example, it is well known that a polygon can be represented by a CSG representation on halfspaces associated with its edge so that every halfspace literal appears exactly once^{18,19}. Yet, for the polygon in Figure 10, it is not possible to construct a DNF with every halfspace literal appearing only once.

A CSG representation Φ is absolutely minimal if there is no other CSG representation that represents a solid S with fewer halfspace literals. It is shown by Lawler³³ that every n -level minimal form is a sum or a product of some $(n - 1)$ -level minimal forms. Thus, in principle, we can use two-level minimal forms to compute three-level minimal forms, and so on, until an absolutely minimal n -level form is obtained. Clearly, the amount of required computation is prohibitive and we must rely on approximate algorithms.

An absolutely minimal CSG representation of solid S can be written in one of the following forms

$$\begin{aligned} S &= |\Phi| = |\Phi_1 + \dots + \Phi_m| = |\Phi_1| + \dots + |\Phi_m| \\ &= S_1 + \dots + S_m \end{aligned} \quad (14)$$

where Φ_1, \dots, Φ_m are absolutely minimal forms for some sets $S_1, \dots, S_m \subseteq S$ respectively, or a product

$$\begin{aligned} S &= |\Phi| = |\Phi_1 \dots \Phi_m| = |\Phi_1| \dots |\Phi_m| \\ &= S_1 \dots S_m \end{aligned} \quad (15)$$

where Φ_1, \dots, Φ_m are absolutely minimal forms for some sets $S_1, \dots, S_m \subseteq \bar{S}$. One approach to multilevel minimization would be to recursively decompose S and/or \bar{S} into a number of smaller, possibly overlapping, subsets S_1, \dots, S_m , using some (heuristic) criteria until absolutely minimal forms for every S_i can be computed.

We focus on the number of times a halfspace literal must appear in a CSG representation. If every halfspace literal appears in a CSG representation of exactly one set S_i , all sets S_i can be represented by disjoint sets of halfspaces. In this case we say that set S admits a

symbolically disjoint decomposition into subsets S_i given by equation (14). A symbolically disjoint decomposition is a locally optimal way to subdivide the CSG minimization problem for S . Suppose we could always find a symbolically disjoint decomposition, i.e. not only for S , but also for every one of $S_i \subseteq S$, and for every subset of S_i , and so on. Then we would compute a CSG representation of S with every halfspace literal appearing exactly once (recall that every halfspace h_i is also necessary by assumption). Such a representation is absolutely minimal by definition.

It is not difficult to show that a set S admits a symbolically disjoint decomposition $S = S_1 + S_2$ if and only if $S = |\Psi_1 + \Psi_2|$, where Ψ_1, Ψ_2 are sums of prime implicants of S and every halfspace $h_i \in H$ appears only in either Ψ_1 , or Ψ_2 ⁸ but not both. This result can be used to show that symbolically disjoint decompositions always exist for a large class of linear polygons, while they do not exist for many curved planar solids. It can be also used as a basis for a heuristic algorithm that performs a multilevel CSG minimization using recursive decomposition⁸. We conclude this section by showing that sometimes an absolutely minimal CSG representation can be constructed based solely on the presence of dominating halfspaces introduced above.

Dominating expressions

According to equation (12), identifying g as a dominating halfspace corresponds to performing a special case of symbolically disjoint decomposition. The following proposition shows that we can take advantage of halfspaces dominating S and \bar{S} simultaneously, and in any order.

Proposition: suppose a halfspace g dominates S and a halfspace h dominates \bar{S} . Then $S = g + (\bar{h} \cdot S_1) = \bar{h} \cdot (g + S_1)$, where S_1 is describable by $H - \{g, h\}$.

Proof: let us use the decomposition of \bar{S} , i.e. $S = \bar{h} \cdot A$. Then $S \subseteq A$. Thus a halfspace g (that dominates S) also dominates A . Therefore $A = g + S_1$. Substituting in expression for S , we get $S = \bar{h} \cdot (g + S_1) = \bar{h}g + \bar{h}S_1$. But $\bar{h}g = g$, because $g \subseteq \bar{h}$. (To see this, observe that $g \subseteq S$ and $S \subseteq \bar{h}$.) Thus the expression for S reduces to $S = g + \bar{h} \cdot S_1$. \square

The result is demonstrated in Figure 12(d), where we have used halfspaces e and a simultaneously. The above proposition can be generalized further. Suppose G_1 is a set of m halfspaces that dominate S , and G_2 is a set of l complements of halfspaces that dominate \bar{S} . Then

$$S = |g_1 \oplus (g_2 \oplus (\dots \oplus (g_n \oplus S_1) \dots))| \quad (16)$$

where $\oplus = (+)$ if it follows $g_i \in G_1$, or $\oplus = (\cdot)$ if it follows $g_i \in G_2$, $n_1 = m + l$, and g_i s are taken in an arbitrary order. The right hand side of equation (16) consisting of a string of the dominating halfspaces and \oplus operators is called a dominating expression.

[‡] Note that decomposition $S = g + S_1$ may not be unique, because terms Π_i covered by g specify 'don't care' conditions for minimization of S_1 .

Consider now the remaining minimization problem: compute a minimal CSG representation for S_1 using only halfspaces from $H - (C_1 \cup C_2)$. It is possible that there are halfspace that do not dominate S (or \bar{S}) but dominate S_1 (or \bar{S}_1). In this case, S_1 in equation (16) can be decomposed further using its dominating expression. In our example, halfspaces f and d dominate \bar{S}_1 (Figure 12(d)) but are not dominating in Figure 12(a). Now the CSG minimization problem for S_1 has been reduced to a yet simpler CSG minimization problem for S_2 on a smaller number of halfspaces, and so on. Figure 12(e) shows the further decomposition of S_1 into product of halfspaces d , f and a set S_2 .

In the sequence of the CSG minimization problems for S_0, S_1, \dots, S_i with $S = S_0$, S_{i-1} is always a simpler minimization problem than S_i is. The sequence can terminate in several ways. If $S_i = \emptyset$, or $S_i = W$, the dominating expression is a minimal representation of S with every halfspace appearing exactly once. This is the case in the example of Figure 12, where halfspaces b and c dominate S_2 (Figure 12(e)), and $S_3 = \emptyset$. At other times, a symbolically disjoint decomposition of S_i or \bar{S}_i into smaller subsets could reveal additional dominating halfspaces. However, in general, if neither of these special conditions apply, we still have to carry out the CSG minimization for S_i as described in the previous subsection.

CONCLUSIONS

Summary

We have considered issues in the construction and optimization of CSG representations of closed regular sets that are given by their boundaries. We have shown that the B-rep to CSG conversion problem is well defined in the partition of W induced from the natural halfspaces of a set S . The Describability Theorem establishes the necessary and sufficient conditions for the existence of CSG representations for a fixed set of halfspaces H and a set S . A new approach to B-rep to CSG conversion was proposed in the third section based on the construction of separating halfspaces. The resulting canonical CSG representation is unique for a fixed set of halfspaces. We also discussed the problem of minimizing the set of halfspaces that are necessary and sufficient for a CSG representation of a set S . The fourth section demonstrated that Boolean minimization techniques can be used effectively to compute minimal or at least efficient CSG representations. It is important that, once the disjunctive canonical decomposition of W is known, no additional geometric computations are necessary.

In principle, the described techniques collectively represent a complete solution to the general B-rep to CSG conversion problem, as well as the problem of CSG minimization. The proposed methods apply to simply- and multiply-connected, manifold and non-manifold solids. A companion paper⁸ describes, in detail, a fully implemented solution of the B-rep to CSG conversion based on this approach.

Geometric computations

Throughout the paper we have assumed that a partition of W by a set of halfspaces H can be computed and is available as needed. This task is straightforward if H is a set of linear halfspaces, and is doable if H is a set of quadratic halfspaces. The problem becomes more difficult if H contains higher order halfspaces because practical methods to compute surface-surface intersections are limited (for a recent survey see Hoffmann⁴⁶).

We observe that, except for the test-based separation described in subsection, 'strategies for separation of components', boundaries of components $C_{m,k}$ are not required for any of the algorithms described in this paper. As we pointed out earlier, points that are in the same component $C_{m,k}$ constitute an equivalence class. Thus for the purpose of testing the describability of a set, a component is completely represented by any single point in its interior. Similarly, for the purpose of the CSG minimization, a canonical intersection term Π_k is completely represented by a single point from some component $C_{m,k} \subseteq |\Pi_k|$. While in theory computing a point in every cell of a cellular decomposition of W may not be easier than computing the boundaries of every cell, alternative practical methods to do so can be developed under relatively mild assumptions about halfspace intersections. For example, an algorithm based on offset halfspaces has been implemented⁸ and can be generalized to higher dimensions.

The only other required geometric computations are classification of a component $C_{m,k}$ against the B-rep of a set S and classification of a canonical intersection term $|\Pi_k|$ against a CSG representation of a set S . Both tests reduce to a point membership classification (PMC) test. The PMC test may be difficult if the point in question lies on the boundary of one or more halfspace in H , because one must deal with various ambiguities if S is represented in CSG¹², or face numerical robustness issues if S is given by its B-rep⁴⁶. Here, however, we are only interested in points lying in the interior of components. Such points can only be in or out of any set S satisfying condition (7). Thus, the PMC test against a B-rep can be performed in a relatively straightforward manner by casting a semi-infinite line from the point in question and counting the number of times the line intersects the B-rep. The PMC test against a CSG representation reduces to a simple syntactic procedure described in the subsection, 'inclusion tests for CSG representations'.

Future work and open issues

The problem of constructing a sufficient set of separating halfspaces must be resolved in particular settings, i.e. for particular classes of natural halfspaces. For example, in Shapira and Vossler⁸ we solve the problem for planar solids bounded by edges that are subsets of convex or concave curves. We are currently working on construction in E^3 which is much more challenging and is likely to require results from algebraic geometry. Studies of the separation properties of the natural quadratic halfspaces (i.e. planar, conical,

cylindrical, spherical halfspaces) have yielded enough understanding to enable us to build an experimental system that converts natural-quadric B-reps in PARASOLID to efficient CSG representations in PADL-2.

We have not addressed any computational complexity issues in this paper. It appears that exact solutions of all minimization problems will require an exhaustive search leading to exponential time algorithms. Thus further development of heuristic and output-sensitive algorithms is important.

Except for linear polygons, little is known about the size of minimal CSG representations. An example in Dobkin *et al.*¹⁹ shows that a CSG representation for some linear polyhedra in E^3 must use a halfspace more than once, but no lower bound on the minimal size is known. Similar limited results are available for curved planar objects⁸. Establishing the relationship between the number and degree of halfspaces in H and the size of a minimal CSG representation for a set S is another challenging problem that is not well understood.

Finally, the existence of unique CSG representation for solids opens new opportunities for development of alternative algorithms to perform solid comparison, interference detection and boundary evaluation. It also establishes a fundamental link between solid modelling algorithms and arrangements in computational geometry.

ACKNOWLEDGEMENTS

We would like to thank Professor Herb Voelcker for his encouragement and numerous suggestions that resulted in a significantly improved paper. The link between cylindrical algebraic decomposition and our work was observed by Don Peterson and Professor Ari Requicha.

This work was supported by the McDonnell Douglas Corporation, the General Motors Corporation (through a Corporate Fellowship), the National Science Foundation under grant MIP-8719196, and Industrial Associates of CPA (a Cornell Research Project).

REFERENCES

- 1 Requicha, A A G 'Mathematical models of rigid solid objects' *Tech. Memo. 28* Production Automation Project, University of Rochester, Rochester, NY, USA (November 1977)
- 2 Requicha, A A G and Tilove, R B 'Mathematical foundations of constructive solid geometry: general topology of closed regular sets' *Tech. Memo. 27a* Production Automation Project, University of Rochester, NY, USA (June 1978)
- 3 Requicha, A A G 'Representations for rigid solids: theory methods and systems' *ACM Computing Surveys* Vol 12 No 4 (1980) pp 437-464
- 4 Requicha, A A G and Voelcker, H B 'Boolean operations in solid modeling: boundary evaluation and merging algorithms' *Proc IEEE* Vol 73 No 1 (1985) pp 30-44
- 5 Miller, J R 'Architectural issues in solid modelers' *IEEE Comput. Graph. Appl.* Vol 9 (September 1989)
- 6 Brown, C K 'PADL-2: a technical summary' *IEEE Comput. Graph. Appl.* Vol 2 No 2 (1982) pp 69-84
- 7 Vossler, D L 'Sweep-to-CSG conversion using pattern recognition techniques' *IEEE Comput. Graph. Appl.* Vol 5 No 8 (1985) pp 61-68
- 8 Shapiro, V and Vossler, D L 'B-rep to CSG conversion II: efficient CSG representations of planar solids' *Tech. Report CPA89-4a* Cornell Programmable Automation, Cornell University, Ithaca, NY (also to appear in *Trans. ASME, J. Mech. Des.*)
- 9 Tilove, R B 'Exploiting spatial and structural locality in geometric modeling' *Tech. Memo. 38* Production Automation Project, University of Rochester, Rochester, NY, USA (October 1981)
- 10 Rossignac, J R and Voelcker, H B 'Active zones in CSG for accelerating boundary evaluation, redundancy elimination, interference detection, and shading algorithms' *ACM Trans. Graph.* Vol 8 No 1 (1989) pp 51-87
- 11 Woodwark, J R 'Eliminating redundant primitives from set-theoretic solid models by a consideration of constituents' *IEEE Comput. Graph. Appl.* Vol 8 No 3 (May 1988) pp 38-47
- 12 Tilove, R B 'Set membership classification: a unified approach to geometric intersection problems' *IEEE Trans. Comput.* Vol C-29 No 10 (1980) pp 874-883
- 13 Requicha, A A G and Voelcker, H B 'Constructive solid geometry' *Tech. Memo. 25* Production Automation Project, University of Rochester, Rochester, NY, USA (November 1977)
- 14 Kuratowski, K and Mostowski, A *Set Theory* North-Holland (1976)
- 15 Miller, R E *Switching Theory, Volume 1: Combinatorial Circuits* John Wiley (1965)
- 16 Ryachev, V L, Kurpa, L V, Sklepus, N G and Uchishvili, L A *Method of R-functions in Problems on Bending and Vibrations of Plates of Complex Shape* (in Russian) Naukova Dumka (1973)
- 17 Tor, S B and Middleditch, A E 'Convex decomposition of simple polygons' *ACM Trans. Graph.* Vol 3 No 4 (1984) pp 244-265
- 18 Peterson, D P 'Halfspace representation of extrusions, solids of evolution and pyramids' *SANDIA Report SAND84-0572* Sandia National Laboratories, USA (1984)
- 19 Dobkin, D, Guibas, L, Hershberger, J and Snoeyink, J 'An efficient algorithm for finding the CSG representation of a simple polygon' *Comput. Graph.* Vol 22 No 4 (1988) pp 31-40
- 20 O'Rourke, J *Art Gallery Theorems and Algorithms* Oxford University Press (1987)

- 21 **Chazelle, B** 'Approximation and decomposition of shapes' in **Schwartz, J T and Yap, C-K (eds)** *Advances in Robotics, Vol 1, Algorithmic and Geometric Aspects of Robotics* Lawrence Elbaum (1987) pp 145–185
- 22 **Pavlidis, T** 'Analysis of set patterns' *Pattern Recognition* Vol 1 (1968) pp 165–178
- 23 **O'Rourke, J** 'Polygon decomposition and switching function minimization' *Comput. Graph. Image Process.* Vol 18 (1982) pp 382–391
- 24 **O'Rourke, J and Supowit, K J** 'Some NP-hard polygon decomposition problems' *IEEE Trans. Inf. Theory* Vol IT-29 No 2 pp 181–190 (1983)
- 25 **Collins, G E** 'Quantifier elimination for real closed fields by cylindrical algebraic decomposition' in *Proc. 2nd GI Conf. Automata Theory and Formal Languages, Lecture Notes in Computer Science 33* Springer (1975) pp 134–183
- 26 **Woo, T C** 'Feature extraction by volume decomposition' *Proc. Conf. CAD/CAM Tech. Mech. Eng.* MIT, Cambridge, MA, USA (March 24–26 1982)
- 27 **Kim, Y S and Wilde, D J** 'A convergent convex decomposition of polyhedral objects' to appear in *Trans. ASME, J. Mech. Des.*
- 28 **Thibault, W C and Naylor, B F** 'Set operations on polyhedra using binary space partitioning trees' *ACM Comput. Graph.* Vol 21 No 4 (1987)
- 29 **Paterson, M S and Yao, F F** 'Binary partitions with applications to hidden-surface removal and solid modelling' *Proc. 5th Annual ACM Symp. Computational Geometry* (1989) pp 23–32
- 30 **Edelsbrunner, H** *Algorithms in Combinatorial Geometry, EATCS Monographs on Theoretical Computer Science, Volume 10* Springer-Verlag (1987)
- 31 **Edelsbrunner, H, Guibas, L, Pach, J, Pollack, R, Seidel, R and Sharir, M** 'Arrangements of curves in the plane – topology, combinatorics, and algorithms' *Technical Report No. 366* Courant Institute of Mathematical Sciences, NY, USA (April 1988)
- 32 **Clarkson, K L, Edelsbrunner, H., Guibas, L J, Sharir, M and Welzl, E** 'Combinatorial complexity bounds for arrangements of curves and surfaces' *29th Annual Symp. Foundations of Computer Science* (October 1988)
- 33 **Lawler, E L** 'An approach to multilevel Boolean minimization' *J. ACM*, Vol 11 No 3 (1964) pp 283–295
- 34 **O'Rourke, J** 'The complexity of computing minimum convex covers for polygon' *Proc. 20th Allerton Conf. Communications, Control and Computing* (1982) pp 75–84
- 35 **Hironaka, H** 'Triangulation of algebraic sets', *Proc. Symp. Pure Mathematics, Algebraic Geometry, ARCATA 1974* Vol 29 (1975) pp 165–185
- 36 **Ruppert, J and Seidel, R** 'On the difficulty of tetrahedralizing 3-Dimensional non-convex polyhedra' *Proc. 5th Annual ACM Symp. Computational Geometry* (1989) pp 380–392
- 37 **Silva, C E** 'Alternative definitions of faces in boundary representations of solid objects' *Tech. Memo. 36* Production Automation Project, University of Rochester, NY, USA (November 1981)
- 38 **Schwartz, J and Sharir, M** 'On piano movers' II. General techniques for computing topological properties of real algebraic manifolds' *Advances in Applied Mathematics* Vol 4 (1983) pp 298–351
- 39 **Buchberger, B, Collins, G E and Kutzler, B** 'Algebraic methods for geometric reasoning' *Annual Rev. Comput. Sci.* Vol 3 (1988)
- 40 **Renegar, J** 'On the computational complexity and geometry of the first-order theory of the reals' *Tech. Report No. 853* School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, USA (July 1989)
- 41 **Birkoff, G and Bartee, T C** *Modern Applied Algebra* McGraw-Hill (1970)
- 42 **Aggarwal, A, Ghosh, S K and Shyamasundar, R K** 'Computational complexity of restricted polygon decompositions' in **Toussaint, G T (ed)** *Computational Morphology* Elsevier Science Publishers (1988) pp 1–11
- 43 **Quine, W V** 'The problem of simplifying truth functions' *Am. Math. Monthly* Vol 59 No 8 (1952) pp 521–531
- 44 **McCluskey, E J** 'Minimization of Boolean functions' *Bell System Tech. Journal* Vol 35 No 5 (1956) pp 1417–1444
- 45 **Mitchell, J S B** *Computational Geometry Workshop* Cornell University, Ithaca, NY, USA (May 1989)
- 46 **Hoffmann, C H** *Geometric and Solid Modeling* Morgan Kaufmann (1989)