# Defining and computing accessibility radius

Vishal Verma Jack Snoeyink
Department of Computer Science
University of North Carolina at Chapel Hill
[verma,snoeyink]@cs.unc.edu

## Abstract

In structural molecular biology software, the accessibility of an atom in a molecule to a ball representing solvent is either computed slowly by building a solvent-accessible surface, or estimated quickly by counting neighboring atoms within some radius. We believe that techniques from computational geometry, notably retraction motion planning in Voronoi diagrams, can give a more direct measure that is faster than surface computation.

We propose an $O(n \log n)$ time algorithm for computing the accessibility radii for sets of points in $\mathbb{R}^2$, and extend this algorithm to sets of circles in $\mathbb{R}^2$ and sets of spheres in $\mathbb{R}^3$. We provide results in $\mathbb{R}^3$ that approximates accessibility radii from the power diagram.

## 1 Introduction

Hydrogen bonds stabilize the most fundamental structures in protein molecules, so modeling their energies is important in software for protein structure prediction or validation [4]. When proteins fold into their three dimensional structure, atoms that need hydrogen bond partners are either found on the surface, where they can bond with water molecules in the solvent, or are buried in the core, where they must have a partner from the protein [7, §1–4]. Thus, in order to properly assign the energy for an atom in a given molecule forming or failing to form a hydrogen bond, we would like to quantify how accessible the atom is to solvent.

Two approaches to quantifying accessibility are common in structural molecular biology software. The accurate but slower approach is to compute the *solvent accessible surface* [5], which is the surface composed of patches of spheres and tori carved out by a probe sphere that is not allowed to penetrate the atom spheres. (1.4Å is the nominal radius for water as a probe sphere.) Atoms are considered accessible in proportion to their areas that appear on this surface. The faster estimate is to simply count whether the number of atoms within a given radius (e.g., 10Å) is below some threshold (e.g., 24, if counting backbone $C_\alpha$ atoms). This is often used in simulation and folding, where the energy computation is repeated many times.

The structural biology models are often based on surprisingly simple geometry: atoms are hard spheres, a molecule is some thousands of atoms joined by segments with fixed length and preferred angles that do not otherwise collide with their near neighbors. The fact that computational geometry has explored algorithms for many interesting notions of depth [8] encouraged us to explore measures of solvent accessibility that might be faster than building a surface, but more direct and accurate than neighbor counting. We therefore formalize the problem of computing the *accessibility radius*, starting with points.

Given a finite set of points or ball *sites* $S \subset \mathbb{R}^d$, we define the *accessibility radius* of a site $p \in S$ as the largest radius of a closed ball that starts completely outside the convex hull of $S$ and can travel to touch $p$ without its interior intersecting any site in $S$. We can use ideas of retraction-based motion planning to compute accessibility radius. In the next section we show these ideas in the plane, where for $n$ points or unit-radius circles we can compute accessibility radius for all sites in $O(n \log n)$ time. Section 4 describes an algorithm for computing the accessibility radii of such atoms (modeled as spheres in $\mathbb{R}^3$) from a Voronoi or power diagram.

## 2 Accessibility in $\mathbb{R}^2$ by retraction

Given a set of point sites $S$ in $\mathbb{R}^2$, Ó'Dúnlaing et al. [6] define the *clearance* of any point $p \in \mathbb{R}^2$ as the smallest distance from $p$ to any site of $S$. The clearance of a path in $\mathbb{R}^2$ is defined as the minimum of clearance among the points on this path.

They [6] also introduce the idea of *retraction* that maps a path in $\mathbb{R}^2$ to a path on the edge skeleton of the Voronoi diagram of $S$ that has at least as much clearance as the original path. The key idea is to move each point $p$ of the path in the direction of gradient of clearance until it reaches its image point $q$ on the edge of the Voronoi diagram. Ó'Dúnlaing et al. [6] show that this map is continuous over $\mathbb{R}^2 \setminus S$ and, when applied to a path in $\mathbb{R}^2$ that avoids $S$, gives a continuous path on the Voronoi diagram.

Thus, retraction reduces the search space for the path of maximum clearance to the set of paths that lie on

the edge-skeleton of the Voronoi diagram of $S$. The maximum clearance path can be found using a best-first graph search that labels each vertex with the maximum clearance among all paths that reach the vertex.

**Input:** Set $S$ of sites in $\mathbb{R}^2$.
**Output:** Accessibility radii for all sites in $S$.
**Algorithm:**

1. Compute the Voronoi diagram $V(S)$ for sites $S$
2. Compute the minimum clearance of each edge of $V(S)$, which occurs at either the endpoints of the edge or the point of intersection of the edge with its dual edge from the Delaunay triangulation.
3. For each vertex $v \in V(S)$, create variable $v.c_{max}$ initialized to 0, except for the infinite vertex, which has $c_{max} = \infty$. Place these in a priority queue by decreasing $c_{max}$.
4. Repeatedly pop the vertex $v$ with greatest $v.c_{max}$, and update max clearance for every neighbor $u$ as $u.c_{max} \leftarrow \max(u.c_{max}, \min(v.c_{max}, uv.clearance))$

If the input is equal radius circles in $\mathbb{R}^2$, the problem reduces to computing the accessibility radius of their center points and subtracting the common radius. If the input is a set of circles of unequal radii, the only change is that the graph $V(S)$ computed in step 1 should be the additively weighted Voronoi diagram of the centers.

## 3 Points and equal-radius spheres in $\mathbb{R}^3$

Given a set $S$ of points in $\mathbb{R}^3$, the accessibility radius of a point in $S$ is defined as the largest radius of a ball that lies completely outside the convex hull of $S$ and can travel to $p$ without intersecting any point in $S$. To compute the accessibility radii of points in $S$, we can map paths in $\mathbb{R}^3 \setminus S$ to paths with higher clearance on the edge skeleton of the Voronoi diagram of $S$ in two steps. Then accessibility radii of points in $S$ can then be computed using best-first search. We are fortunate that the edge skeleton here is connected.

The first retraction step maps each point in the interior of a Voronoi cell to a point on a Voronoi face. If the resulting path passes through any local minimum of clearance inside a face, then it should be perturbed to avoid this minimum, which will increase clearance. Then the second retraction can map each point that is not a local minimum in the interior of a Voronoi face to a point on a Voronoi edge. The retraction is done such that at each step a point is mapped to another point with higher clearance. To see that it maps paths to paths, we can cover the path with a finite number of balls that avoid sites in each cell or discs that avoid local minima in each face and observe that they map to overlapping neighborhoods.

When $S$ is a set of equal radius spheres, we compute the accessibility radii of the set of centers of the spheres in $S$. Then, the accessibility radius of each sphere in $S$ is the accessibility radius of its center minus its radius.

## 4 Unequal radius spheres

When $S$ is a set of spheres with unequal radius we perform the first retraction step as in the previous section. However, we can not perform the second retraction step since the edge skeleton of the Voronoi diagram of $S$ may not be connected. We add new vertices and edges to the edge skeleton of the Voronoi diagram and obtain a connected graph that we call the *augmented edge skeleton*. We show that paths in faces can be mapped to paths with higher clearance on the augmented edge skeleton. Thus, accessibility radii of points in $S$ can be computed using a best-first search on this augmented edge skeleton.

### 4.1 Augmented edge skeleton

A bounded Voronoi face $F$ may have a disconnected boundary consisting of an outer connected component the we call the *outer boundary*, and possibly other connected components that we call the *inner boundaries*. Each inner boundary has a Voronoi edge that has a local maximum of clearance in its interior. Because the faces of an additively weighted Voronoi diagram are quadratic, the edge formed by their intersection either has constant clearance, or at most two maxima and minima (since any fixed clearance gives two quadratic curves that have at most four intersection points). We add a vertex at each local maximum (or one if all points are maxima) and split the Voronoi edge into two at the newly added vertex. Since there are $O(n)$ number of edges, these vertices can be constructed in $O(n)$ time.

**Lemma 1** *The gradient of clearance at a newly added vertex points inside $F$.*

Next, we connect these new vertices. Let $v$ be a vertex that is a local maximum of clearance on an inner boundary of $F$. From $v$ we move along the gradient of clearance on $F$ until we reach a (inner or outer) boundary point of $F$, then move along the boundary in the increasing direction of clearance until we reach a vertex $u$. The minimum of clearance of the path traveled from $v$ to $u$ occurs at $v$. We add an edge between $u$ and $v$ and call $u$ the parent of $v$. Note that this edge corresponds to an actual path on $F$. The parent of $v$ can be computed by intersecting the boundary edges of $F$ with the plane through $v$ and the two sites closest to points on $F$. There are $O(n)$ edges, and for each vertex its parent can be found in $O(n)$ time. Thus the new edges can be constructed in $O(n^2)$ time.

The graph formed by the addition of the above vertices and edges to the edge-skeleton of the Voronoi diagram of $S$ is called the augmented edge-skeleton. It has no cycles, so is connected.

### 4.2 Retraction from face to augmented skeleton

In this section we prove that paths on faces can be mapped to paths on the augmented edge-skeleton that have at least as much clearance as the original path.

**Lemma 2** *For every boundary point, $p$, of $F$ there is a path on the augmented edge skeleton from $p$ to the outer boundary of $F$ that has the minimum of clearance at $p$.*

**Proof.** If $p$ lies on the outer boundary of $F$ then the lemma is trivially true.

If $p$ is a vertex on the inner boundary of $F$ and is not a local maximum of clearance, then there is an edge from $p$ to a vertex with strictly higher clearance than $p$. This edge has a minimum of clearance at $p$. If $p$ is a vertex on the inner boundary of $F$ and is a local maximum of clearance then there is an edge from $p$ to its parent that has a minimum of clearance at $p$. The clearance of the parent of $p$ is strictly greater than that of $p$. Thus, there is an edge from every inner boundary vertex of $F$ to a vertex of $F$ with strictly higher clearance. Since $F$ has a finite number of vertices on its boundary, this sequence of increasing vertices has to terminate at the outer boundary of $F$. Thus every inner boundary vertex $p$ is connected to an outer boundary vertex through a path on the augmented edge skeleton that has a minimum of clearance at $p$.

If $p$ lies on an edge of an inner boundary, then we can move from $p$ along this edge in the direction of increasing clearance until we reach a vertex, $v$. Since $v$ is connected to an outer boundary vertex through a path on the augmented edge skeleton that has a minimum of clearance at $v$, the lemma holds for $p$ too.

□

**Lemma 3** *Let $p$ be an inner boundary point and $p'$ be an outer boundary point of $F$. Let $P$ be a path from $p'$ to $p$ on $F$. Then there is a path from $p'$ to $p$ on the augmented edge skeleton that has at least as much clearance as $P$*

**Proof.** From the previous lemma we know that there is a path, $PQ$, on the augmented edge skeleton from $p$ to some outer boundary point $q$ that has the minimum of clearance at $p$.

Let $S_1$ and $S_2$ be the spheres in $S$ that are associated with the face $F$. Let $S_V$ be the set of points in $\mathbb{R}^3$ that are equidistant from $S_1$ and $S_2$. $F$ is a subset of $S_V$. The line connecting the the centers of $S_1$ and $S_2$ passes through exactly one point, $p_{min}$ in $S_V$. The clearance of any point on the closure of F is the distance of that point from $S_1$ and $S_2$. The gradient of this function at any point $p_0 \in S_V$ is parallel to the intersection of $S_V$ with the plane defined by $p_0$ and centers of $S_1$ and $S_2$.

Extend the curve $P$ from $p'$ in the direction of gradient of clearance to infinity to get the semi-infinite curve $P_{ext}$. Similarly extend the curve $PQ$ from $q$ in the direction of gradient of clearance to infinity to get the semi-infinite curve $PQ_{ext}$.

Since $F$ is bounded there are two paths from $q$ to $p'$, along the outer boundary, one clockwise, $Q^+$, and the other anticlockwise, $Q^-$. Let $S_{V1}$ be the subset of $S_V$ that is bounded by $P_{ext}$, $PQ_{ext}$ and contains $Q^+$ and $S_{V2}$ be the subset of $S_V$ that is bounded by $P_{ext}$, $PQ_{ext}$ and contains $Q^-$. The interiors of $S_{V1}$ and $S_{V2}$ are disjoint.

If $p_{min}$ lies on $PQ_{ext}$ or $P_{ext}$ then the lemma is trivially true.

Otherwise $p_{min}$ either lies in the interior of $S_{V1}$ or in the interior of $S_{V2}$. Without loss of generality, we assume that $p_{min}$ lies in the interior of $S_{V1}$. Then we prove that the path $PQ$ followed by $Q^-$ has at least as much clearance as $P$.

Let $u$ be any point on $Q^-$. Let $C$ be the curve of intersection of $S_V$ with the plane defined by centers of $S_1$ and $S_2$ and $u$. We further restrict our attention to $C'$, the part of $C$ between $p_{min}$ and $u$. $C'$ is parallel to the gradient of clearance on $S_V$ and intersects $PQ_{ext}$ or $P_{ext}$. Thus there is a point on $PQ_{ext}$ or $P_{ext}$ that has no more clearance than $u$.

Thus $Q^-$ has at least as much clearance as $P$ and $PQ$ and $Q^-$ lie completely on the augmented edge-skeleton.

□

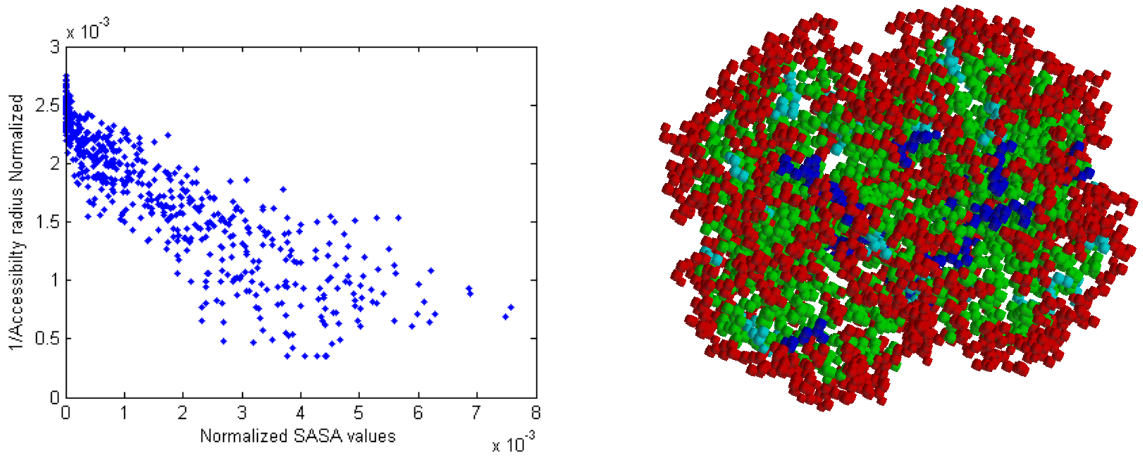
**Complexity of the Algorithm:** Additively weighted Voronoi diagram can be computed in $O(n^3)$ time [2]. From this diagram the augmented edge skeleton can be computed in $O(n^2)$ time. The best first search on the this edge skeleton is done in $O(n \log n)$ time. Since computing additively weighted Voronoi diagrams is expensive and uses high degree predicates, in the next section we test an approximate algorithm that uses the power diagram instead.

## 5 Results

In $\mathbb{R}^3$ we used the power diagram [1] instead of the additively weighted Voronoi diagram in the algorithm of Section 2. This is an approximation to accessibility radius because the radii of different atoms are not the same. Since the radii are similar and the atoms are well-packed, the edge skeleton of the power diagram is always connected and we do not need to add new edges and vertices.

Table 1 lists for three proteins, correlation coefficients between SASA values and three estimates: the number of $C_\alpha$ neighbors within 10Å, Dorr's neighbor vector measure [3], and accessibility radius.

Due to space constraints, we close with a single example of comparing the classification of individual atoms by these methods: Figure 1(b) show recombinant hemoglobin (PDB id 1c7d) colored by classifica-




(a) (b)

Figure 1: Comparing accessibility radius vs SASA for recombinant hemoglobin, PDB id 1C7D, 4,396 atoms. (a)Scatter plot of SASA vs 1/(Accessibility radius) (b) Colored by agreement (red=surface, green=buried) and disagreement (blue=AR buried, cyan=AR surface).

| PDB id | Number of Neighbors | Neighbor Vector | Accessibility Radius |
|---|---|---|---|
| 1C7D | -0.7745 | 0.8588 | -0.8664 |
| 1D0V | -0.7904 | 0.8726 | -0.8616 |
| 1MBN | -0.7923 | 0.8454 | -0.8618 |

Table 1: Correlation of SASA with number of neighbors, neighbor vector, and accessibility radius on three proteins: recombinant hemoglobin (1C7D), nicotinate mononucleotide (1D0V), and myoglobin (1MBN).

tion (surface/buried) by SASA and accessibility radius. Negative correlation can be seen in the scatter plot (Figure 1(a)) for SASA and the reciprocal of accessibility radius values.

## Acknowledgments

We thank Matt O'Meara for obtaining the SASA, neighbor, and neighbor vector values from Rosetta for our comparisons. This research was partially supported by the Darpa protein design project and NIH.

## References

[1] Franz Aurenhammer. Power diagrams: Properties, algorithms and applications. *SIAM J. Comput.*, 16(1):78–96, 1987.

[2] Jean-Daniel Boissonnat and Christophe Delage. Convex hull and voronoi diagram of additively weighted points. In G. S. Brodal and S. Leonardi, editors, *ESA*, volume 3669 of *LNCS*, pages 367–378. Springer, 2005.

[3] E. Durham, B. Dorr, N. Woetzel, R. Staritzbichler, and J. Meiler. Solvent accessible surface area approximations for rapid and accurate protein structure prediction. *J. Molec. Modeling*, February 2009. Epub ahead of print.

[4] G. A. Jeffrey and W. Saenger. *Hydrogen Bonding in Biology and Chemistry*. Springer-Verlag, Berlin, 1991.

[5] B Lee and FM Richards. The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol*, 55(3):379–400, 1971.

[6] Colm Ó'Dúnlaing, Micha Sharir, and Chee K. Yap. Retraction: A new approach to motion-planning. *ACM STOC*, pages 207–220, 1983.

[7] Gregory A Petsko and Dagmar Ringe. *Protein Structure and Function*. Primers in Biology. Sinauer Assoc, 2004.

[8] P. J. Rousseeuw and A. Struyf. Computation of robust statistics: Depth, median, and related measures. In J. E. Goodman and J. O'Rourke, editors, *Handbook of Discrete and Computational Geometry*, pages 1279–1292. Chapman & Hall/CRC, 2nd edition, 2004.