The Art Gallery Problem: An Overview and Extension to Chromatic Coloring and Mobile Guards

Nicole Chesnokov

May 16, 2018

Contents

Abstract

In this paper we explore the art gallery problem, presented by Chvátal in 1975, which aims to find the smallest number of guards required to cover an art gallery shaped as a polygon of n vertices. Chvátal found that it is occasionally necessary and always sufficient to have $\lfloor \frac{n}{3} \rfloor$ guards to cover an n-vertex polygon. First we provide a well-known proof of this bound for the art gallery problem. Then we provide two extensions of this problem. The first extension enhances the original problem by determining the minimum number of colors required to color a guard set S, such that no two guards with conflicting areas of visibility are assigned the same color. The second extension allows guards to be mobile and aims to determine the shortest route for each guard to guard the space.

1 Introduction

One of the main goals in robotics is to create robots that are able to mimic human behavior, a large component of which is motion sensors and navigation. Robots with these capabilities would be able to localize themselves in a new environment, which would solve a myriad of problems, one of the simplest being guarding an art gallery. The art gallery problem was one of the earliest and most influential problems in sensor placement. [S]

The problem was first posed to Václav Chvátal by Victor Klee in 1973 and was stated as: Consider an art gallery, what is the minimum number of stationary guards needed to protect the room? In geometric terms, the problem was stated as: given a n-vertex simple polygon, what is the minimum number of guards to see every point of the interior of the polygon? Chvátal was able to prove that for simple polygons $\lfloor \frac{n}{3} \rfloor$ guards is necessary and sufficient to guard the gallery when there are n vertices in the polygon. However, this proof was very complex and used the method of induction. In 1978 Steve Fisk constructed a much simpler proof via triangulation, which is a method of decomposing a polygon into triangles, and coloring of vertices. Fisk's method will be presented in Section 2. The many real life applications of this problem not only inspired the mathematics community to find a tighter bound due to real-life restrictions, but also inspired many variations of the problem that modeled real-life situations. [S]

One of the many extensions of this problem is the chromatic art gallery problem, which aims to determine the minimum number of colors required to color a guard set, a set of vertices in an n-vertex polygon. A guard set is colored such that no two conflicting guards have the same color, where two conflicting guards are those whose areas of visibility overlap. A lower bound for this problem was found by Erickson and LaValle in 2010, which stated that for any value k there exists a polygon with $3k^2 + 2$ vertices such that the minimum number of guards needed is k. Additionally, Erickson and LaValle determined that for a spiral polygon the chromatic guard number has an upper bound of 2. These theorems will be introduced further and proven in Section 3. [EL]

The second extension that will be explored in this paper, in Section 4, is the watchman route problem. The watchman route problem is an optimization problem in geometry where guards are now mobile and the aim is to determine the shortest route that a watchman should take such that all points are visible from this route. The first polynomial time algorithm for this problem was determined by Carlsson, Johnsson and Nilsson in 1999 and runs in worst case $O(n^6)$. The premise of the algorithm is to precompute the shortest fixed watchman routes via reflections at the fragment end points, such that the only case to consider is when the routes make perfect reflections. Thus having reduced the problem, a process called *sliding* simulates the "reflection of the watchman route as the reflection point moves between the two endpoints of an active fragment." Further details of the algorithm, the definitions of the procedures and the runtime are explained in Section 4. [CJN]

2 The Art Gallery Problem

The art gallery problem is formulated in geometry as the minimum number of guards that need to be placed in an n-vertex simple polygon such that all points of the interior are visible. A simple polygon is a connected closed region whose boundary is defined by a finite number of line segments. Visibility is defined such that two points u and v are mutually visible if the line segment joining them lies inside the polygon. Using these definitions, Steve Fisk was able to prove Chvátal's initial theorem using triangulation and vertex coloring.

Theorem 1 $\lfloor \frac{n}{3} \rfloor$ *guards are occasionally necessary and always sufficient to cover an n-vertex polygon.*

2.1 Proof

We begin the proof of **Theorem 1** by considering some example polygons and the respective number of guards necessary to ensure that the whole area is guarded. Then we introduce triangulation and vertex coloring in order to find the generalized bound of $\lfloor \frac{n}{3} \rfloor$ guards.

Base Cases:

1. Only one guard necessary:

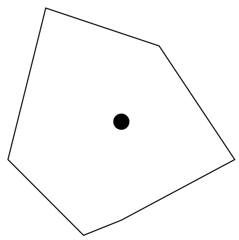

Figure 1:
Convex Polygon

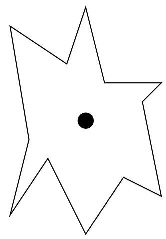

Figure 2: Star

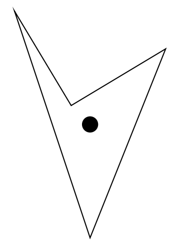

Figure 3:
4-sided polygon

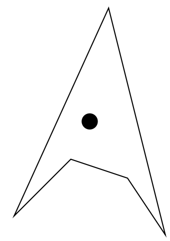

Figure 4:
5-sided polygon

2. Two guards necessary:

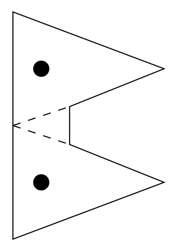

Figure 5: 6-sided polygon

Now we will generalize how we would determine the number of guards necessary for any n-vertex polygon. First we must triangulate the polygon. That is, we must draw $n-3$ non-crossing diagonals between corners of the walls until the whole area consists of triangles.

Claim 1 *The graph produced upon triangulation is 3-colorable.*

Proof For the base case $n = 3$ this is trivial. Suppose that $n > 3$: We draw a total of $n - 3$ non-crossing diagonals between corners of the polygon, until the whole polygon is triangulated. Select two vertices u and v that are connected via a diagonal. This diagonal splits the polygon into two sections, which both share the edge uv. We can continue connecting two vertices via diagonals within each of the sections, such that we end up with triangles covering the area of the polygon ie we have triangulated the polygon. Then, we can begin 3-coloring the polygon by assigning one color to vertex u and the another color to vertex v. We then color the third vertex adjacent to these two with the third color. Then, since the whole polygon is triangulated, by induction we can color each triangle with one of the three colors, such that the initial colorings of vertex u and vertex v hold. Another way to conceptualize this is that since the colors of the vertices making up the first edge is set two the first two colors, the colors of the vertices that create a triangle with this edge is predetermined (ie they will all be colored the third color). Now these three edges serve as bases of triangles, such that the color of the vertices that make up triangles with each of those edges is also predetermined (ie they get colored the remaining color not used in the vertices of the edge they form a triangle with).

Since there are n vertices, we know that one of the colors contains at most $\lfloor \frac{n}{3} \rfloor$ vertices. Thus, select one color and place the guards at the vertices of that color, such that the maximum number of guards is $\lfloor \frac{n}{3} \rfloor$. Since each of the triangles contains at least one vertex of this color, this means that every triangle is guarded and thus the whole polygon is guarded.[AZ]

Above we assumed that a triangulation exists for all graphs and just used the technique, however we must prove that a triangulation always exists. In general, a triangulation does not always exist, particularly in three-dimensions as detailed in *Proofs from THE BOOK*. However, in the art gallery problem we are only dealing with planar non-convex polygons and thus will show that triangulation does indeed always exist for such geometric shapes. [AZ]

Claim 2 *Triangulation always exists for planar non-convex polygons.*

Proof We prove this theorem via induction. The base case is $n = 3$, in which case the polygon is a triangle and it clearly possible to triangulate it, that is it is already triangulated. Suppose now that $n \geq 4$. In order to use induction, we must show that there exists a diagonal that divides the polygon into two parts, each of which can be triangulated.

Recall that the sum of the interior angles of a polygon is $(n-2) * 180°$ and that a convex angle has measure less than $180°$. Combining these two properties there must exist a convex vertex in the polygon via the pigeonhole principle. We can use the pigeonhole principle to find a tighter bound on the number of convex vertices. If there are less than two convex vertices, then there would be at least $n-1$ non-convex vertices, in which case the interior angles would sum to at least $(n-1) * 180°$, which is greater than $(n-2) * 180°$ and thus is a contradiction. If there were exactly two convex vertices, there would be exactly $n-2$ non-convex vertices, in which case the interior angles would sum to at least $(n-2) * 180°$, leaving two vertices that must have degree greater than 0. Thus, this too would be greater than $(n-2) * 180°$, which is a contradiction. However, with three convex vertices, there are $n-3$ non-convex vertices that sum to at least $(n-3) * 180°$, leaving at least $180°$ for the remaining 3 vertices, which is valid. Thus, there must exist at least three convex vertices.

Now select one of the convex vertices of the polygon and create a diagonal by connecting its two neighboring vertices. If this diagonal is fully in the interior of the polygon then this will serve as the diagonal we are looking for. However, if this diagonal does not fully lie within the polygon, then we slide it down towards the convex vertex until the first reached vertex. Then we create a diagonal between the convex vertex and this new vertex. This is further detailed in the example, where vertex A is the convex vertex (Figure 6). Now that we have found the diagonal we can apply induction to triangulate the rest of the polygon and thus a triangulation always exists. [AZ]

2.2 Example

Let us now apply the techniques of triangulation and vertex coloring detailed above on an example gallery from *Proofs from THE BOOK* on page 267. We begin with the graph shown below:

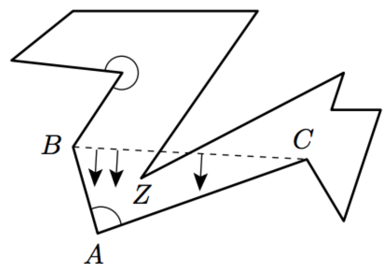

Figure 6: 12-sided polygon: step 1 of triangulation [CJN]

where A is the convex vertex and a diagonal is created between its two neighbors B and C. However, the diagonal is not fully within the polygon, thus we move the diagonal until it reaches the first vertex which is Z. Thus, we create a diagonal from A to Z.

Using this first diagonal we can then further triangulate the polygon. Then using the triangulation we can 3-color the vertices of the polygon, yielding the following:

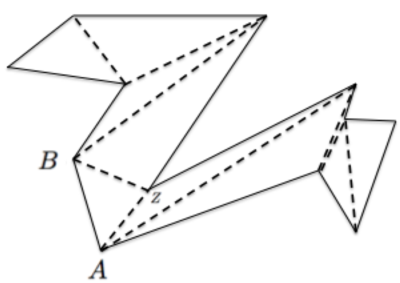

Figure 7: Triangulated 12-sided polygon

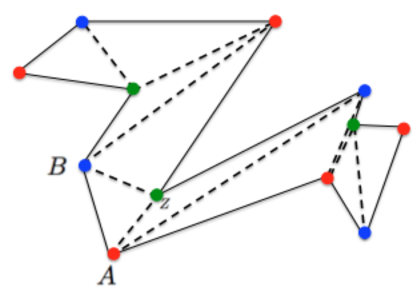

Figure 8: 3-colored 12-sided polygon

Looking at the colored figure above, we can select the green vertices as the guards, thus determining the minimum number of guards necessary to guard this art gallery. There is a green vertex in each of the triangles, thus each of the triangles is guarded and thus the whole polygon is guarded.

3 The Chromatic Art Gallery Problem

Stemming from the chromatic coloring procedure detailed in Section 2 is a variation of the art gallery problem called the chromatic art gallery problem. Given a set of vertices S, of a convex polygon P and visible regions for each vertex (guard), determine the minimum number of colors necessary such that no two guards whose visible regions overlap share the same color. The minimum number of colors necessary to color this set is denoted as $C(S)$. Visibility of a vertex v ($V(v)$) is all points p in the polygon, such that p is visible from v. Then, two vertices u and v are considered to be conflicting if $V(u) \cap V(v) \neq 0$. The minimum number of colors necessary to color a guard set, across all sets of vertices of the polygon, $T(P)$, is defined as $\chi_G(P) = min_{S \in T(P)} C(S)$. This value is called the chromatic guard number of the polygon P. [EL]

Theorem 2 *For every positive integer k, there exists a polygon P with $3k^2 + 2$ vertices such that $\chi_G(P) \geq k$.*

3.1 Proof

We begin this proof by considering a polygon called the standard "comb" which is frequently used to show the occasional necessity of $\lfloor \frac{n}{3} \rfloor$ guards in the original art gallery problem. The polygon consists of $3k^2 + 2$ vertices, such that there is a rectangular region and k^2 "teeth" attached at the bottom of this rectangle. Using this general construction, we can construct a polygon P with the following coordinates: $[(0,1), (1,0), (2,1), (4,2), (5,0), (6,1) ... (4k^2 - 4, 1), (4k^2 - 3, 0), (4k^2 - 2, 1), (4k^2 - 2, 2k - 2), (0, 2k - 2)]$. The rectangular region has corners $(0,1), (4k^2 - 2, 1), (4k^2 - 2, 2k - 2), (0, 2k - 2)$.

Now we must determine where we can place guards in this polygon such that the whole polygon is guarded. Let us define guards whose coordinate $y < 1$ *apex* guards and those whose $y \geq 1$ *body* guards. Intuitively, apex guards can be placed above an individual "tooth" of the comb, such that they guard only one "tooth". However, that would result in a lot of interfering guards and thus would increase the chromatic guard number although we are trying to minimize it. The aim is to place apex guards such that they do not interfere and thus can guard more than one "tooth", thus minimizing the chromatic guard number. Suppose that an apex guard is placed at the center of the "tooth" $(1,0)$, then its visibility range will be bounded by the ray extending from the rightmost wall of that tooth. This ray, using basic trigonometry, will intersect the top of the rectangular region a distance of $2k - 2$ from the center of the "tooth" (the angle from the center of the tooth is 45 degrees and the height is $2k - 2$). That also means that the nearest conflicting guard will also be able to guard $2k - 2$ in horizontal length and will thus be centered at a "tooth" $4k - 4$ away. Thus, two guards will conflict if the distance between their corresponding "teeth" is at most $4k$. We then define consecutive "teeth" as a set of "teeth" where the maximum distance between any two apex points corresponding to any two "teeth" is $4k$. Then define m_{apex} as the maximum number of apex guards in any consecutive set of k "teeth" (s_2 in Figure 8).

Body guards, on the other hand, are able to guard up to k distinct "teeth" since two non-conflicting guards must be placed at least $4k$ distance apart and teeth are a distance of 4 apart, as defined earlier in the construction of the "comb". The visibility of a body guard includes the entire rectangular region of the polygon and every apex guard's visibility region includes the rectangular region, thus all body guards will conflict with all other guards in the polygon. We define m_{body} as the number of body guards used in a guard set of the polygon (s_1 in Figure 8).

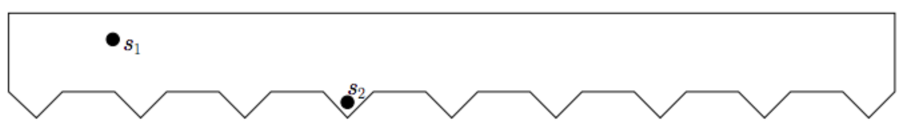

Figure 9: The point s_1 is an apex guard and the point s_2 is a body guard, for a polygon with 9 "teeth" ($k = 3$).

Suppose now that polygon P has a guard set S that only requires $\chi_G(P)$ colors. If we consider k consecutive "teeth" in P with m_{apex} apex guards, then all of the apex guards will conflict with each other, and each of these will also conflict with all of the body guards, m_{body}. Thus, this means that $\chi_G(P) \geq m_{apex} + m_{body}$. As defined earlier, each body guard can guard at most k "teeth" and there are a total of k^2 "teeth", thus by the pigeonhole principle apex guards can guard at most km_{apex} "teeth". Additionally, each "tooth" must be guarded so $km_{apex} + km_{body} \geq k^2 \rightarrow m_{apex} + m_{body} \geq k$, so therefore $\chi_G(P) \geq m_{apex} + m_{body} \geq k$. [EL] An example of this coloring is shown below in Figure 9.

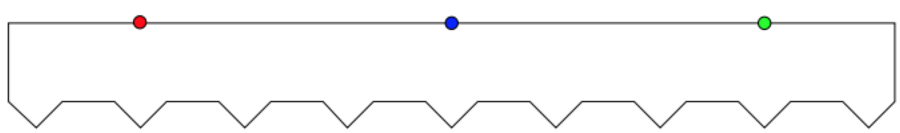

Figure 10: This a polygon with 9 "teeth" ($k = 3$), with a guard placement that requires three colors.

3.2 Upper Bounds for the Chromatic Guard Problem

In addition to determining the lower bound for the the chromatic guard number, detailed above, Erickson and LaValle also determined upper bounds on the chromatic guard number. The trivial upper bound is $\lfloor \frac{n}{3} \rfloor$ for an n-vertex polygon because $\lfloor \frac{n}{3} \rfloor$ guards is always sufficient to guard the polygon (as proven in Section 2.1), in which case each guard is colored a unique color. However, there are many polygons with a large number of vertices whose chromatic guard number is 2, so this bound is not satisfactory. Thus, Erickson and LaValle prove better bounds for spiral polygons.

A spiral polygon is a polygon with exactly one maximal reflex subchain, which is "a chain that forms part of the boundary of a polygon and where all internal vertices have an internal angle of greater than π radians." [EL]

Theorem 3 *For any spiral polygon P, $\chi_G(P) \leq 2$.*

The proof of this theorem is fairly complex, so we will highlight the general ideas.[1] By definition of a spiral polygon, there is exactly one reflex subchain and one convex subchain. Additionally, guards are placed along the edges of the convex subchain. The first guard gets placed at the start of the convex subchain. The visibility of this guard is determined by the most clockwise point along the convex subchain, p_n, and the most counterclockwise point along the reflex subchain, b_n. Then let g_n be the first vertex clockwise from b_n and let r_n be the point colinear with g_n and b_n on the convex subchain. Thus, the next guard is placed on the convex subchain in the interval between p_n and r_n. An example of this procedure, as detailed in the paper by Erickson and LaValle [EL], is shown below.

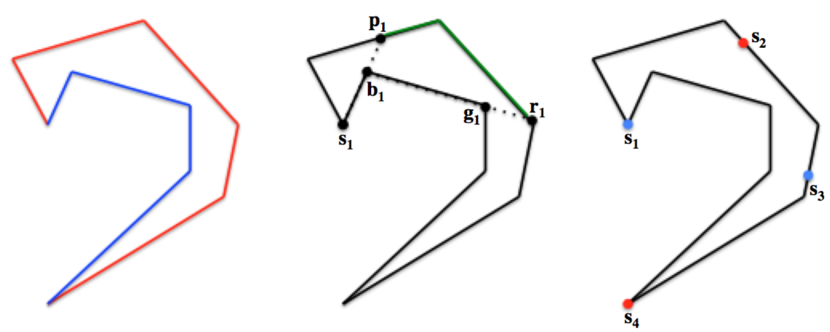

Figure 11: LEFT: The red segment is the convex subchain and the blue segment is the reflex subchain.
MIDDLE: The first guard is placed at s_1 and it guards the region up to the dotted line connecting b_1 and p_1. The points b_1, p_1, g_1 and r_1 are marked as determined by the procedure detailed above. The segment where s_2 can be placed is marked in green.
RIGHT: Continuing this procedure, yields the placement of four guards indicated by $s_i \forall i \in [1, 4]$. The guards are then colored blue or red, odd or even index, respectively.

Upon completing this procedure for placement of guards, Erickson and LaValle prove that this yields a triangulation of the polygon. Then, since all guards are

[1] A detailed proof can be found in the paper by Erickson and LaValle [EL].

placed along the convex subchain, if two guards conflict then their guarded areas must intersect somewhere along the convex subchain. However, s_{n+1} and s_{n-1} cannot conflict, else s_n would not have been placed on the convex subchain. The guards can then be colored based on their index, such that all even indexed guards are colored one color and all odd indexed guards are colored another color. [EL]

4 The Watchman Route Problem

Another extension of the original art gallery problem is to relax the constraints such that the guards are now mobile, which was first introduced by Toussaint and Avis in 1981 as the watchman route problem. A watchman route W of a polygon P is defined as a curve inside of P such that W guards P. This problems aims to find the optimal watchman route for which a polygon is covered, such that the optimal route is the shortest route. Thus, this problem not only focuses on the geometric concept of visibility, but also focuses on determining an optimization algorithm. The first polynomial time algorithm for finding the shortest watchman route was determined by Carlsson, Jonsson and Nilsson [CJN].

4.1 Key Terms and Techniques

Before presenting the algorithm we will introduce the idea of a fixed watchman route and key sub-procedures necessary for understanding the algorithm. Let us first define a fixed watchman route, which is a watchman route where the route passes through a fixed point d on the boundary of the polygon. Two techniques used to solve the fixed watchman route problem are *essential cuts* and *reflection principles.*

Define a *cut* to be a directed line segment in the polygon such that part of the cut's interior must lie within the polygon (polygon edges are not cuts) and the cut divides the polygon into two subpolygons. A specific type of cut, an *extension cut*, is defined by extending all two edges connected at a vertex until they reach a boundary point inside of the polygon. By looking at a route defined by such cuts, it is trivial to see that all guard sets must have a point on each extension cut or to the left, else all edges collinear to the cuts will not be seen by the guard set. Using this concept, we can define *essential cuts* which are extension cuts that are not dominated by any other extension cuts. An extension cuts dominates another extensions cut if all points in the polygon to the left of one cut are also to the left of the other cut. [CJN]

Essential cuts are used to define the *reflection principle*, which states that the shortest watchman route can only bend at essential cuts and at vertices of the polygon (shown by Chin and Ntafos [2]). By moving a bend point on an essential cut, while keeping other bend points constant, the route lengthens and

[2] W. Chin and S. C. Ntafos. Shortest watchman routes in simple polygons. *Discrete Computational Geometry*, 6:9?31, 1991.

if the bend point is in the interior of the cut then the outgoing and incoming edges are reflections of each other. If however the bend point is at the end of the cut, then the outgoing edge is bounded by the extension of the incoming edge and the reflection of the extension across the cut. [CJN]

Upon introducing key terms and sub-procedures, Carlsson, Jonsson and Nilsson [CJN] state the following lemma without proof:

Lemma 1 *A closed curve is a watchman route if and only if the curve has at least one point to the left of (or on) each essential cut.*

Using the lemma and terminology defined above, the watchman route problem can be formulated as: "Compute the shortest closed curve that intersects all essential cuts".

Theorem 4 *There is an algorithm that, given a boundary point d in a simple polygon of n edges, the backward essential cuts with respect to d, and their subdivision into fragments, computes the shortest fixed watchman route through d in $O(n|C|F)$ time and $O(n|C|)$ storage, where $|C|$ is the number of essential cuts and F is the number of fragments.*

4.2 The Algorithm

The algorithm that runs in $O(n|C|F)$ as stated by **Theorem 4** is complex and is the crux of the Carlsson, Jonsson and Nilsson paper, thus we will only present an overview of the algorithm. The general idea of the algorithm is to precompute the shortest fixed watchman routes making reflections at the fragment end points. However, this can result in a case where the routes make only perfect reflections in the interior of fragments, which are reflections where the incoming angle equals the outgoing angle of the reflection. In order to resolve this issue a technique called *sliding* is used which simulates the motion of a reflection point as it is moved between the two end points of an active fragment, which are fragments that contain the points of the intersection of the route and the cut. A more detailed description of the sliding procedure can be seen in Section 4.4 of the Carlsson, Jonsson and Nilsson paper [CJN].

5 Conclusion

The art gallery problem has served an integral part in developing algorithms for creating autonomous programs that will guide robots in closed spaces. It has inspired much research in this field by varying the constraints of the original problem to include chromatic coloring of guards as well as finding the shortest route that guards the art gallery. The chromatic art gallery problem can be applied in robotics by helping develop programs that prevent interference and collisions of moving robots in closed spaces they are surveilling, thus optimizing the space they cover and limiting likelihood of collisions. The algorithms developed for robotics that stem from the watchman route problem are pursuit and evasion algorithms. This would have several applications in capture or retrieving tasks, some of which could be applied in military and medical settings. Thus, although a fairly simply stated problem, the art gallery problem has evolved to become a strong foundation and inspiration for solving more complex problems that can one day be applied to autonomous systems.

References

[AZ] Aigner, M., Ziegler, G. M. (2004). *Proofs from the book* (4th ed.). Berlin: Springer.

[CJN] Carlsson, S., Jonsson, H., Nilsson, B. J. (1993). Finding the shortest watchman route in a simple polygon. *Algorithms and Computation Lecture Notes in Computer Science*, 22(3), 58-67. doi:10.1007/3-540-57568-5_235

[EL] Erickson, L.H., LaValle, S.M. (2010). *A chromatic art gallery problem.*

[S] Safak, G. (2009). *The Art-Gallery Problem: A Survey and an Extension* (Master's thesis, Royal Institute of Technology, 2009). Sweden: KTH Computer Science and Communication.