On the Complexity of Point Recolouring in Geometric Graphs

Henk Meijer
Roosevelt Academy
Middelburg, The Netherlands
h.meijer@roac.nl

Yurai Núñez-Rodríguez
Queen's University
Kingston, ON Canada
yurai@cs.queensu.ca

David Rappaport *
Queen's University
Kingston, ON Canada
daver@cs.queensu.ca

Abstract

Given a collection of points representing geographic data we consider the task of delineating boundaries based on the features of the points. Assuming that the features are binary, for example, red or blue, this can be viewed as determining red and blue regions, or states. Due to regional anomalies or sampling error, we may find that reclassifying, or recolouring, some points may lead to a more rational delineation of boundaries. In this note we study the maximal length of recolouring sequences where recolouring rules are based on neighbour relations and neighbours are defined by a geometric graph. We show that the difference in the maximal length of recolouring sequences is striking, as it can range from a linear bound for all trees, to an infinite sequence for some planar graphs.

1 Introduction

Given a set of planar points partitioned into red and blue subsets, a red-blue separator is a boundary that separates the red points from the blue ones. There has been considerable investigation of methods for obtaining such red-blue separating boundaries. In his PhD thesis, Seara [8], examines various means for red-blue separation. For the case of red-blue separation with the minimum perimeter polygon the problem is known to be NP-hard [3, 1]. A somewhat related topic is to obtain a balanced subdivision of red and blue points, that is, faces of the subdivision contain a prescribed ratio of red and blue points. Kaneko and Kano [4] give a comprehensive survey of results pertaining to red and blue points in the plane, including results on balanced subdivisions.

For some applications one is willing to reclassify points by recolouring them so as to obtain a more reasonable boundary. For example Chan [2] shows that finding a red-blue separating line with the minimum number of reclassified points takes $O((n+k^2)\log k)$ expected time, where k is the number of recoloured points.

In Reinbacher et al. [7] a heuristic algorithm is presented for obtaining a better delineating boundary that recolours points. The input is a triangulated set of n planar red-blue points. For a point p to be recoloured it needs to be "surrounded" by points of the opposite colour. Reinbacher et al. show experimental results on delineating boundaries after recolouring.

A *surrounded point* is realized when there is a contiguous set of oppositely coloured neighbours of p, in the triangulation, that span a radial angle greater than $180°$. As the recolouring occurs in an iterative sequence it is not clear that the process will ever come to an end. However, Reinbacher et al. show that no sequence that iteratively recolours surrounded points will ever visit the exact same colouring of the points more than once. Thus the maximum number of recolourings is bounded by the total number of possible colourings which is $2^n - 1$. This bound was improved by Núñez-Rodríguez and Rappaport [5] by proving that any recolouring sequence has $O(n^2)$ length. This bound is, in fact, tight.

In this note, we present bounds on recolouring for other types of geometric graphs. Our main results include bounds on the number of recolourings for graphs of maximum degree 3, bounds on the number of recolourings of trees, and examples of infinite recolouring sequences on planar and non-planar graphs.

In the next section we precisely describe the recolouring problem. We follow, in the subsequent section, with our new results on the length of recolouring sequences of geometric graphs, such as planar graphs, non-planar graphs, and trees. The last section discusses some extensions of our results. Most of our proofs have been included in the Appendix.

2 Preliminaries

We are given as input a drawing of a graph, $D = (S, E)$, where S is a set of points in the plane partitioned into blue points and red points and $E \subseteq S \times S$ is the set of edges of the graph. The edges are represented by straight line segments. An edge incident to points p and q is denoted as $\overline{pq}$. We assume throughout, for simplicity of exposition, that the points are in general position. We colour the edges of D red if its two incident points are red, and blue if its two incident points are

*Supported by an NSERC of Canada Discovery Grant

blue. If one of the incident points is red and the other is blue we mix the colours to obtain a magenta edge.

Definition 1 *Let the edges of D be coloured as above. Then the* ***magenta angle*** *of a point $p \in S$ is:*

- $0°$, *if p has at most one radially consecutive incident magenta edge,*
- $360°$, *if p has degree greater than one and is only incident to magenta edges,*
- *the maximum angle between two or more radially consecutive incident magenta edges, otherwise.*

Notice that, according to the previous definition, a point with only one neighbour in D has magenta angle $0°$ regardless of the colour of its neighbour (See Figure 1). A *surrounded* point is one with magenta angle larger than $180°$. Therefore, a point of degree zero or one is never surrounded nor recoloured. We say an edge *is in* the magenta angle of a point if it is incident to the point and falls within the span of the magenta angle.

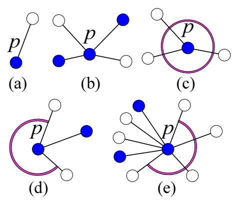

Figure 1: Examples of magenta angles α of point p. Magenta angles larger than $0°$ are represented by arcs: (a), (b) $\alpha = 0°$, (c) $\alpha = 360°$, (d), (e) $180° < \alpha < 360°$.

The strategy of reclassification by recolouring, recolours a surrounded point p at a time. The sequence in which surrounded points are recoloured can be driven by a mixed criterion, such as recolouring the surrounded point with the largest magenta angle, or with the largest number of edges in the magenta angle. According to Reinbacher et al. [7], there exist mixed criteria that always produce recolouring sequences of linear size. In the sequel, we assume surrounded points are recoloured in an arbitrary manner, in order to find bounds for any, and all, possible recolouring strategies. The recolouring process stops when there are no more surrounded points.

3 Bounded and Unbounded Recolouring Sequences

At all times the graphs are assumed to be connected because, in general, each connected component can be considered independently. We use the term *drawing*, to refer to the drawing of a graph as defined in the previous section. There are families of drawings for which every recolouring sequence is finite and others that allow infinite recolouring sequences. We characterize some of these families. A family of drawings with finite recolouring sequence comes from graphs with maximum degree 3.

Theorem 1 *Let D be a drawing with n bi-chromatic points. If D has maximum point degree 3, then the length of any recolouring sequence of D is $O(n)$.*

The proof of Theorem 1 (See Appendix) only relies on the number of point neighbours. Thus, the result also holds for non-planar and more general drawings with maximum degree 3.

3.1 Planar Drawings

As opposed to triangulations, planar drawings may have non-convex faces and points of degree one. One may think of obtaining bounds for planar graphs based on the fact that a planar graph is a subgraph of a triangulation. However, this does not seem to help since there are simple examples where a subgraph can have either a larger or a smaller number of recolourings in the worst case over all initial colourings.

In fact, a recolouring sequence of a planar graph can be infinite. Figure 2 shows an example of a graph and a colour configuration that lead to an infinite recolouring sequence. Observe from Figure 2 that the initial colouring repeats after a number of steps (recolourings). Notice that only certain recolouring sequences are infinite in this example. The drawing in Figure 2 can be made 2-connected and the minimum point degree can be increased by carefully adding more edges incidents to the points of degree one, without affecting the recolouring sequence.

3.2 Non-Planar Drawings

At this point, it is obvious that one can also construct non-planar drawings with infinite recolouring sequences since planar drawings allow so. Nevertheless, the examples shown for infinite recolouring sequences on planar drawings include points that never change colour. We show an example of a non-planar drawing with infinite recolouring sequence where every point changes colour infinitely many times (see Figure 3). If similar examples can be built for planar drawings, these have not yet been found.

3.3 Trees

In this subsection we use the term *tree drawing* to refer to a straight line drawing of a tree (not necessarily planar). A trivial example of a tree drawing that has $O(n)$ recolouring sequence is a “jigsaw” path with points alternately coloured. In such example, all blue points

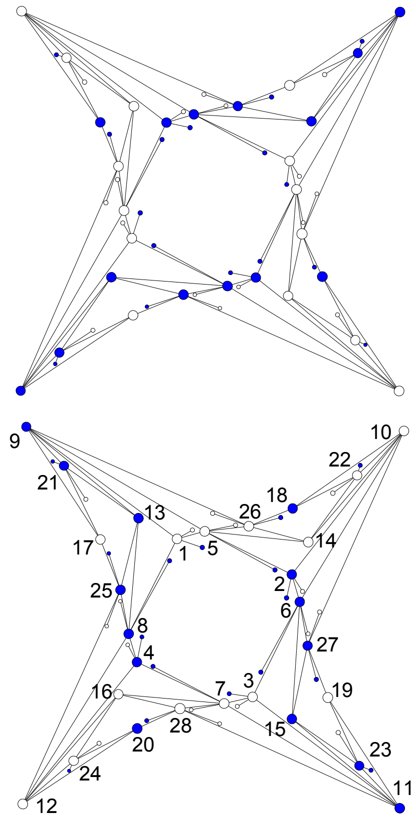

Figure 2: Planar drawing D with infinite recolouring sequence. Points represented by smaller circles never change colour. Top: initial colouring of D. Bottom: D after 28 recolourings. The labels indicate the order in which points are recoloured. Notice that the bottom drawing is a rotation of the top drawing. This indicates that the recolourings can repeat infinitely many times.

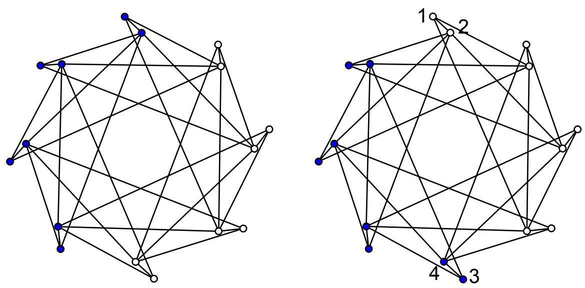

Figure 3: Non-planar drawing D with infinite recolouring sequence. Left: initial colouring of D. Right: D after 4 recolourings. The labels indicate the order in which points are recoloured. Notice that the right drawing is a rotation of the left drawing. This indicates that the recolourings can repeat infinitely many times.

can be coloured to red, leading to approximately $n/2$ recolourings.

It is not hard to prove that the number of recolourings of tree drawings is $O(n^2)$ by the same arguments used for triangulations (Theorem 9 [5]), with minor modifications. However, this bound is not tight. As a corollary of Theorem 1 we have that binary tree drawings have a linear number of recolourings. In the remainder of this section we prove that the number of recolourings of general tree drawings is also linear.

In order to prove a tight bound (Theorem 4) for the recolouring of tree drawings, we define a partial order on the recolourings involved in a recolouring sequence. Then we bound the total number of recolourings based on the number of minimal elements (sinks) of such partial order. This idea is explained and formalized in what follows.

We denote a recolouring event r, or simply a recolouring, as the event of a certain point p changing colour. Let $R = (r_1, \ldots, r_k)$ be a recolouring sequence where r_i denotes the recolouring at step i, $1 \leq i \leq k$, $k > 0$. We also denote $p(r)$ as the point that changes colour at recolouring r, and $N(r)$ the number of times that $p(r)$ has changed colour in R prior to event r.

Definition 2 *Let T be a tree drawing and let R be a recolouring sequence of T. The **history graph** of R is a directed graph $H = (R, I)$, $I \subseteq R \times R$ such that $(r_j, r_i) \in I$ if and only if $\overline{p(r_i)p(r_j)}$ is in the magenta angle associated to r_j.*

Observation 1 *By the definition of history graph all the edges are directed from later recolourings to earlier ones. Therefore, a history graph is a directed acyclic graph (DAG) and defines a partial order on the elements of the recolouring sequence.*

The following lemma formally states that for two consecutive recolourings of a point p to occur, it is required that at least two neighbours of p change colour in between.

Lemma 2 *Let T be a tree drawing with n bi-chromatic points, let R be a recolouring sequence of T, and let $H = (R, I)$ be the history graph of R. Consider a recolouring $r \in R$ with $N(r) > 0$. Then the outdegree of r is at least 2. Moreover, there exist two distinct neighbours of $p(r)$, $p_1, p_2 \in T$, with recolourings $s_1, s_2 \in R$, respectively, such that $(r, s_1) \in I$ and $(r, s_2) \in I$.*

Proof. Obviously, if a point is recoloured red (similarly blue) and was recoloured earlier in the sequence, the previous recolouring was to blue (red). The intersection between the magenta angles at the time it is surrounded by red (blue) and previously by blue (red) contains at least two edges since the corresponding magenta angles are greater than 180°. Therefore, there are at least two neighbours of p, $p_1, p_2 \in T$ that are recoloured at least once between two consecutive recolourings of p. $\square$

In the light of Lemma 2, we can state the following definition.

Definition 3 *Let R be a recolouring sequence of a tree drawing T and $H = (R, I)$ be the corresponding history graph. The* ***binary history graph*** *of R, $BH = (R, BI)$, $BI \subseteq I$, is a subgraph of the history graph where nodes have outdegrees 2 or 0: nodes with outdegree 0 correspond to first time recolourings; nodes with outdegree 2 correspond to subsequent recolourings. Consider a node r_k of degree 2 in the binary history tree. From Lemma 2 we know that there are two distinct neighbours of $p(r_k)$ that have been previously recoloured. Thus we choose the two outgoing edges of r_k (r_k, r_i), (r_k, r_j), such that i and j are the largest indices smaller than k for neighbours of r_k in the history graph where $p(r_i) \neq p(r_j)$.*

The motivation to define the binary history graph is to obtain a cycle-free subgraph of the history graph that involves all the recolourings. This is formalized in the next lemma.

Lemma 3 *Let T be a tree drawing, and R a recolouring sequence of T with binary history graph BH. BH has no directed or undirected cycles. Therefore, BH is a forest of trees.*

To obtain a bound on the size of binary history trees we show that the number of nodes is linear in the size of the corresponding tree drawing. This will lead us to conclude the results of the following theorem.

Theorem 4 *Let T be a tree drawing with n bi-chromatic points. The length of any recolouring sequence of T is $O(n)$.*

4 Extensions

4.1 Surrounded Threshold Greater Than 180°

Thus far we have assumed that a point is surrounded when its magenta angle is any value greater than 180°. Reinbacher et al. [7] show that a threshold value smaller than 180° allows for infinite recolouring sequences on very simple graphs –trees included. Some of our results hold for threshold values $\alpha > 180°$. Trees, for example, have linear recolouring sequences for any threshold $180° < \alpha < 360°$. Also Theorem 1 holds for any threshold $\alpha > 0°$. Other results do not seem to hold for any threshold value. For instance, it is not clear how large the value of α can be such that infinite recolouring sequences exist on planar graphs.

4.2 More than two colours

Suppose that the points come in more than two colours. We define the colour of an edge as the mixture of the colours of its endpoints. In a multi-coloured scenario we say that p is *surrounded* by a set of edges of a single mixed colour if the edges define a continuous angle greater than 180°. As we may intuitively observe, increasing the number of colours only lowers the chances of a point being surrounded without changing the fundamental nature of the problem. In fact, inspection shows that all of our previous results hold in a multi-coloured scenario. Thus, our recolouring bounds for a bi-chromatic set of points carry over to multi-coloured point sets.

5 Conclusions

We have re-examined a point recolouring method useful for reclassifying points to obtain reasonable subdividing boundaries. We show tight (linear) bounds on trees and graphs of maximum degree 3 for the longest possible sequence of recolourings. Planar and non-planar graphs have been shown to have infinitely many recolourings.

Some interesting questions remain open. First, can planar drawings have sequences of recolourings where all the points change colour infinitely many times? Also, what is the complexity of point recolouring in planar drawings and other geometric graphs when the threshold to consider a point as surrounded is greater than 180°?

References

[1] E. M. Arkin, S. Khuller, and J. S. B. Mitchell. Geometric knapsack problems. *Algorithmica*, 10:399–427, 1993.

[2] T. M. Chan. Low-dimensional linear programming with violations. in *Proceedings of the 43rd IEEE Symposium on Foundations of Computer Science*, 570–579, 2002.

[3] P. Eades and D. Rappaport. The complexity of computing minimum separating polygons. *Pattern Recogn. Lett.*, 14:715–718, 1993.

[4] A. Kaneko and M. Kano. Discrete geometry on red and blue points in the plane —a survey. in *Discrete and Computational Geometry, The Goodman-Pollack Festschrift*, Algorithms and Combinatorics, Springer, vol. 25, 551–570, 2003.

[5] Y. Núñez-Rodríguez, and D. Rappaport. Tight bounds for point recolouring. *Proc. 18th Canadian Conf. in Comp. Geometry (CCCG'06)*, Kingston, ON, Canada, 67–70, 2006.

[6] Joseph O'Rourke. Computational Geometry in C. *Cambridge University Press*, 1998.

[7] Iris Reinbacher, Marc Benkert, Marc van Kreveld, Joseph S.B. Mitchell, Jack Snoeyink, and Alexander Wolff. Delineating boundaries for imprecise regions. *Algorithmica*, vol. 50:3, 386–414, 2008.

[8] Carlos Seara. On Geometric Separability. *PhD. Thesis. Universitat Politècnica De Catalunya*, 2002.

A Appendix (Proofs)

Theorem 1 Let D be a drawing with n bi-chromatic points. If D has maximum point degree 3, then the length of any recolouring sequence of D is $O(n)$.

Proof. The proof follows from an observation on how the number of magenta edges decreases with each recolouring. Let p be a point that is being recoloured at a certain step along a recolouring sequence. By definition, at least two edges incident to p need to be magenta. These edges change to a solid colour after the recolouring of p. Before p is recoloured, at most one edge of solid colour can be incident to it given that $deg(p) \leq 3$. This edge, if it exists, will become magenta. Therefore, the number of magenta edges decreases by at least one with each recolouring. As the initial number of magenta edges is $O(n)$, the number of recolourings is also $O(n)$. $\square$

Lemma 3 Let T be a tree drawing, and R a recolouring sequence of T with binary history graph BH. BH has no directed or undirected cycles. Therefore, BH is a forest of trees.

Proof. Since BH is a subgraph of the history graph of R, BH is also a DAG, by Observation 1. Therefore, there are no directed cycles in BH. Any undirected cycle in BH would have at least one node r with two outgoing edges and one node s with two incoming edges. For the sake of contradiction, we assume that there exists such a cycle, C, in BH.

Consider the function $f : R^* \to V(T)^*$ such that $f(r_1, r_2, \ldots, r_k) = p(r_1), p(r_2), \ldots, p(r_k)$, $r_i \in R, 1 \leq i \leq k$. In particular, f maps a path in BH to a path in T. Let P_1 and P_2 be the two undirected paths that connect r and s in C. By the definition of binary history graph, the outgoing edges of r are incident to nodes t_1 and t_2 such that $p(t_1) \neq p(t_2)$. Thus, $|C| > 2$. Without loss of generality, let t_1 be in P_1 and t_2 be in P_2. Then paths $f(P_1)$ and $f(P_2)$ are different at points $p(t_1)$ and $p(t_2)$. This implies that there are two different paths in T connecting $p(r)$ and $p(s)$. Thus, we establish a contradiction. $\square$

Theorem 4 Let T be a tree drawing with n bi-chromatic points. The length of any recolouring sequence of T is $O(n)$.

Proof. Let R be any recolouring sequence of T, and let BH be the binary history graph of R. In order to prove this theorem we show that $|V(BH)| = |R|$ is $O(n)$. Let $V_k(BH)$ denote the set of nodes of degree k in BH, and $V_{k^+}(BH)$ be the set of nodes of degree at least k in BH. For accounting purposes, we split the nodes of BH into four classes: $V_0(BH)$, $V_1(BH)$, $V_2(BH)$, and $V_{3^+}(BH)$.

Nodes of degree 0 and 1 are all first-time recolourings (sinks) according to the definition of binary history graph, since these have 0 outgoing edges. Also, nodes of degree 2 are either sinks or sources because internal nodes have degree at least 3, that is, one or more incoming edges and two outgoing edges. The following transformation removes the sources of BH such that, in the resulting graph, all nodes of degree 2 are guaranteed to be sinks.

Let H' be a copy of BH, except that every source r and outgoing edges (r, t_1) and (r, t_2) in BH are replaced by the edge (t_1, t_2) in H'. We already know, from Lemma 3, that there are no undirected cycles in BH. Therefore, edges (t_1, t_2) or (t_2, t_1) could not have existed in BH. Notice that one edge is added in H' for each node removed. Thus,

$$|V(BH)| \leq |V(H')| + |E(H')| \leq 2|V(H')| - m, \tag{1}$$

where m is the number of connected components of H', given that H' is a forest of trees. The degrees of all the nodes remaining in H' is preserved. Therefore, $V_0(H') = V_0(BH)$, $V_1(H') = V_1(BH)$, and $V_2(H')$ only consist of sink nodes. At most n nodes can be sinks since at worst all nodes are recoloured for the first time. Consequently,

$$|V_0(H')| + |V_1(H')| + |V_2(H')| \leq n. \tag{2}$$

Thus, a linear bound on $|V_{3^+}(H')|$ entails a linear bound on $|V(H')|$. We derive such bound in what follows. From properties of graphs and, in particular, of forests of trees,

$$\sum_{r \in V(H')} deg(r) = 2|E(H')| = 2|V(H')| - 2m = 2(|V_0(H')| + |V_1(H')| + |V_2(H')| + |V_{3^+}(H')|) - 2m. \tag{3}$$

According to the definitions of V_k and V_{k^+},

$$\sum_{r \in V(H')} deg(r) \geq |V_1(H')| + 2|V_2(H')| + 3|V_{3^+}(H')|. \tag{4}$$

Equations (3) and (4) lead to

$$|V_{3^+}(H')| \leq 2|V_0(H')| + |V_1(H')| - 2m \leq 2(|V_0(H')| + |V_1(H')|) \tag{5}$$

Combining this with (2) we obtain

$$|V_{3^+}(H')| \leq 2n. \tag{6}$$

Finally, from (1), (2), and (6) we have

$$|V(BH)| \leq 2|V(H')| - m \leq 2(|V_0(H')| + |V_1(H')| + |V_2(H')| + |V_{3^+}(H')|) \leq 6n \tag{7}$$

□