



Karlsruhe Reports in Informatics 2010,21

Edited by Karlsruhe Institute of Technology,
Faculty of Informatics
ISSN 2190-4782

On d-regular Schematization of Embedded Paths

Andreas Gemsa, Martin Nöllenburg,
Thomas Pajor, Ignaz Rutter

2010

KIT – University of the State of Baden-Wuerttemberg and National
Research Center of the Helmholtz Association



Fakultät für Informatik

Please note:

This Report has been published on the Internet under the following
Creative Commons License:

<http://creativecommons.org/licenses/by-nc-nd/3.0/de>.

On d -regular Schematization of Embedded Paths

Andreas Gemsa¹, Martin Nöllenburg^{*1,2}, Thomas Pajor¹, and Ignaz Rutter¹

¹ Institute of Theoretical Informatics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

² Department of Computer Science, University of California, Irvine, USA

Abstract. In the d -regular path schematization problem we are given an embedded path P (e.g., a route in a road network) and an integer d . The goal is to find a d -schematized embedding of P in which the orthogonal order of all vertices in the input is preserved and in which every edge has a slope that is an integer multiple of $90^\circ/d$. We show that deciding whether a path can be d -schematized is NP-hard for any integer d . We further model the problem as a mixed-integer linear program. An experimental evaluation indicates that this approach generates reasonable route sketches for real-world data.

1 Introduction

Angular or \mathcal{C} -oriented schematizations of graphs refer to a class of graph drawings, in which the admissible edge directions are limited to a given set \mathcal{C} of (usually evenly spaced) slopes. This includes the well-known class of orthogonal drawings and extends more generally to k -linear drawings, e.g., octilinear metro maps. Applications of schematic drawings can be found in various domains such as cartography, VLSI layout, and information visualization.

In many schematization scenarios the input is not just a graph but a graph with an initial drawing that has to be schematized according to the given set of slopes. This is the case, e.g., in cartography, where the geographic positions of network vertices and edges are given [6], in sketch-based graph drawing, where a sketch of a drawing is given and the task is to improve or schematize that sketch [3], or in dynamic graph drawing, where each drawing in a sequence of drawings must be similar to its predecessor [5]. For such a redrawing task it is crucial that the mental map [13] of the user is preserved, i.e., the output drawing must be as similar as possible to the input. Misue et al. [13] suggested preserving the *orthogonal order* of the input drawing as a simple criterion for maintaining a set of basic spatial properties of the input, namely the relative above/below and left/right positions of all pairs of input nodes. The orthogonal order has been used successfully as a means for maintaining the mental map [4, 7, 9, 11].

The motivation behind the work presented here is the visualization of routes in road networks as sketches for driving directions. An important property of a route sketch is that it focuses on road changes and important landmarks rather than exact geography and distances. Typically the start and destination lie in populated areas that are locally reached via a sequence of relatively short road segments. On the other hand, the majority of the route typically consists of long highway segments with no or only few road changes. This property makes it difficult to display driving directions for the whole route in a single traditional map since some areas require much smaller scales than others. The strength of route sketches for this purpose is that they are not drawn to scale but rather use space proportionally to the route complexity.

* Supported by grant NO 899/1-1 of the German Research Foundation (DFG)

Related Work. Geometrically, we can consider a route to be an embedded path in the plane. The simplification of paths (or polylines) in cartography is well studied and the classic line simplification algorithm by Douglas and Peucker [8] is one of the most popular methods. Two more recent algorithms were proposed for \mathcal{C} -oriented line simplification [12, 14]. These line simplification algorithms, however, are not well suited for drawing route sketches since they keep the positions of the input points fixed or within small local regions around the input points and thus edge lengths are more or less fixed. On the other hand, Agrawala and Stolte [1] presented a system called *LineDrive* that uses heuristic methods based on simulated annealing to draw route sketches. Their method allows distortion of edge lengths and angles. It does not, however, restrict the set of edge directions and does not give hard quality guarantees for the mental map such as the preservation of the orthogonal order.

A graph drawing problem that has similar constraints as drawing route sketches is the metro-map layout problem, in which an embedded graph is to be redrawn octilinearly. The problem is known to be NP-hard [15] but it can be solved successfully in practice by mixed-integer linear programming [16]. The existing methods covered in a survey by Wolff [17] do aim to keep the mental map of the input but no strict criterion like the orthogonal order is applied. Brandes and Pampel [4] studied the path schematization problem in the presence of orthogonal order constraints in order to preserve the mental map. They showed that deciding whether a rectilinear schematization exists that preserves the orthogonal order of the input path is NP-hard. They also showed that schematizing a path using arbitrarily oriented unit-length edges is NP-hard. Delling et al. [7] gave an efficient algorithm to compute \mathcal{C} -oriented drawings of monotone paths that preserve the orthogonal order and have minimum schematization cost. The schematization cost counts the number of edges that are not drawn with their closest \mathcal{C} -oriented direction. The authors also sketch a heuristic approach for schematizing non-monotone paths.

Contributions. In this paper we close the complexity gap of the path schematization problem that remained open between the hardness result of Brandes and Pampel [4] for rectilinear paths and the efficient algorithm of Delling et al. [7] for monotone \mathcal{C} -oriented paths. We prove that deciding whether a \mathcal{C} -oriented orthogonal-order preserving drawing of an embedded input path exists is NP-hard, even if the path is simple. This is true for every d -regular set \mathcal{C} of slopes that have angles that are integer multiples of $90^\circ/d$ for any integer d . The case $d = 1$ is covered by Brandes and Pampel [4] but their proof relies on the absence of diagonal edges and hence does not extend to other values of d . We show the hardness in the octilinear case $d = 2$ in Section 3 and subsequently, in Section 4, how this result extends to the general d -regular case for $d > 2$. We finally design and evaluate a mixed integer-linear program (MIP) for solving the d -regular path schematization problem in Section 5. Our experimental results show that routes in practice usually consist of only a small number of relevant road segments and that our MIP is indeed able to quickly generate reasonable sketches for those routes.

2 Preliminaries

A plane embedding of a graph $G = (V, E)$ is a mapping $\pi : G \rightarrow \mathbb{R}^2$ that maps every vertex $v \in V$ to a distinct point $\pi(v) = (x_\pi(v), y_\pi(v))$ and every edge $e = uv \in E$ to the line segment $\pi(e) = \overline{\pi(u)\pi(v)}$ such that no two edges e_1, e_2 cross in π except at common endpoints. For simplicity we also use the terms vertex and edge to refer to their images under an embedding.

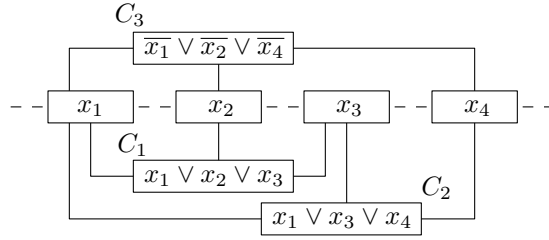


Fig. 1. A MONOTONE PLANAR 3-SAT instance with four variables and three clauses.

We measure the slope of an edge uv as the counterclockwise angle formed between the horizontal line through $\pi(u)$ and the line segment $\pi(uv)$. For a set of angles \mathcal{C} we say that a drawing is \mathcal{C} -oriented if the slope of every edge $e \in E$ is contained in \mathcal{C} . A set of slopes \mathcal{C} is called d -regular for an integer d if $\mathcal{C} = \mathcal{C}_d = \{i \cdot 90^\circ/d \mid i \in \mathbb{Z}\}$.

Let π and ρ be two embeddings of the same graph G . We say that ρ respects the *orthogonal order* [4] of π if for any two vertices u and $v \in V$ it holds that $x_\rho(u) \leq x_\rho(v)$ if $x_\pi(u) \leq x_\pi(v)$ and $y_\rho(u) \leq y_\rho(v)$ if $y_\pi(u) \leq y_\pi(v)$. In other words, the orthogonal order defines the relative above-below and left-right positions of any two vertices.

Let (G, π) be a graph G with a plane input embedding π . A d -regular *schematization* (or d -schematization) of (G, π) is a plane embedding ρ that is \mathcal{C}_d -oriented, that preserves the orthogonal order of π and where no two vertices are embedded at the same coordinates. We also call ρ *valid* if it is a d -schematization of (G, π) .

3 Hardness of 2-regular Path Schematization

In this section we show that the problem of deciding whether there is a 2-schematization for a given embedded graph (G, π) is NP-hard, even if G is a simple path. In the latter case we denote the problem as the (d -regular) PATH SCHEMATIZATION PROBLEM (PSP). We fix $d = 2$. To prove that 2-regular PSP is NP-hard we first show hardness of the closely related 2-regular UNION OF PATHS SCHEMATIZATION PROBLEM (UPSP), where (G, π) is a set \mathcal{P} of k embedded disjoint paths $\mathcal{P} = \{P_1, \dots, P_k\}$.

We show that 2-regular UPSP is NP-hard by a reduction from MONOTONE PLANAR 3-SAT, which is known to be NP-hard [2]. MONOTONE PLANAR 3-SAT is a special variant of PLANAR 3-SAT where each clause either contains exactly three positive literals or exactly three negative literals and additionally, the variable-clause graph admits a planar drawing such that all variables are on the x -axis, the positive clauses are embedded below the x -axis and the negative clauses above the x -axis. An example instance of MONOTONE PLANAR 3-SAT is depicted in Fig. 1. In a second step, we show how to augment the set of paths \mathcal{P} to form a single simple path P that has the same properties as \mathcal{P} and thus proves that PSP is NP-hard.

In the following, we assume that φ is a given MONOTONE PLANAR 3-SAT instance with variables $X = \{x_1, \dots, x_n\}$ and clauses $C = \{c_1, \dots, c_m\}$.

3.1 Hardness of the Union of Paths Schematization Problem

In the following we first introduce the different types of required gadgets and then show how to combine them in order to prove the hardness of UPSP.

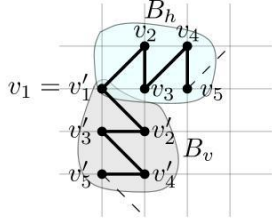


Fig. 2. Border gadget B that is rigid and regular. The horizontal component is denoted by B_h and the vertical one by B_v .

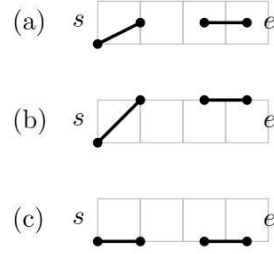


Fig. 3. The switch s is linked with edge e ; (a) shows the input embedding, in (b) e is pulled up and in (c) it is pulled down.

Border Gadget. For all our gadgets we need to control the placement of vertices on discrete positions. Thus our first gadget is a path whose embedding is unique up to scaling and translation. This path will induce a grid on which we subsequently place the remaining gadgets.

Let $\Delta_x(u, v)$ be the x -distance between the two vertices u and v . Likewise, let $\Delta_y(u, v)$ be the y -distance between u and v . We call a simple path P with embedding π *rigid* if a 2-schematization of (P, π) is unique up to scaling and translations. Further, we call an embedding π of P *regular* if there exists a length $\ell > 0$ such that for any two vertices u, v of P it holds that $\Delta_x(u, v) = z_x \cdot \ell$ and $\Delta_y(u, v) = z_y \cdot \ell$ for some $z_x, z_y \in \mathbb{Z}$. Thus in a regular embedding of P all vertices are embedded on a grid whose cells have side length ℓ . For our border gadget we construct a simple path B of appropriate length with embedding π that is both rigid and regular. Hence, π is essentially the unique 2-schematization of (B, π) , and after rescaling we can assume that all points of B lie on an integer grid. The border gadget consists of a horizontal component B_h and a vertical component B_v which share a common starting vertex $v_1 = v'_1$, see Fig. 2. The component B_h alternates between a 45° edge to the upper right and a vertical edge downwards. The vertices are placed such that their y -coordinates alternate and hence all odd, respectively all even, vertices have the same y -coordinate. The vertical component consists of a copy of the horizontal component rotated by 90° in clockwise direction around v_1 . To form B , we connect B_v and B_h by identifying their starting points v_1 and v'_1 .

Lemma 1. *The border gadget B with its given embedding π is rigid and regular.*

Proof. First note that in B_h due to the orthogonal order all vertices v_i with i odd and all vertices v_i with i even have the same y -coordinates, respectively, in any valid schematization. Hence $\Delta_y(v_i, v_{i+1})$ is the same for all i . We show that v_i and v_{i+1} for i odd, also have the same x -distance $\Delta_x(v_i, v_{i+1}) = \Delta_y(v_i, v_{i+1})$.

Since the edge v_1v_2 connects two vertices with different y -coordinates and since v_2v_3 must be embedded vertically, v_1v_2 must be embedded with an angle of 45° . And hence we have $\Delta_x(v_1, v_2) = \Delta_y(v_1, v_2)$. The same argument holds for all edges v_iv_{i+1} with i odd and hence B_h is both regular and rigid.

Since the vertical border gadget B_v is a copy of B_h it is also regular and rigid. Moreover, we have that $\Delta_x(v'_1, v'_2) = \Delta_x(v_1, v_2)$, i.e., the distances between vertices of B_v that lie on

the same horizontal or vertical lines are the same as for B_h . Hence the border gadget B is regular. \square

In the following we define the length of a grid cell induced by B as 1 so that we obtain an integer grid. We choose B long enough to guarantee that any vertex of our subsequent gadgets lies vertically and horizontally between two pairs of vertices in B .

The grid in combination with the orthogonal order gives us the following properties:

1. If we place a vertex v with two integer coordinates, i.e., on a grid point, its position in any valid embedding is unique;
2. if we place a vertex v with one integer and one non-integer coordinate, i.e., on a grid edge, its position in any valid embedding is on that grid edge;
3. if we place a vertex v with two non-integer coordinates, i.e., in the interior of a grid cell, then its position in any valid embedding is in that grid cell (including its boundary).

Basic Building Blocks. We will frequently make use of two basic building blocks that rely on the above grid properties.

The first one is a *switch*, i.e., an edge that has exactly two valid embeddings. Let $s = uv$ be an edge within a single grid cell where u is placed on a grid point and v on a non-incident grid edge. We call u the *fixed* and v the *free* vertex of s . Assume that u is in the lower left corner and v on the right edge of the grid cell. Then in any valid d -schematization s is either horizontal or diagonal, see Fig. 3 for an example.

The second basic concept is *linking* of vertices. We can synchronize two vertices in different and even non-adjacent cells of the grid by assigning them the same x - or y -coordinate. We call two vertices u and v *linked*, if in π either $x_\pi(u) = x_\pi(v)$ or $y_\pi(u) = y_\pi(v)$. Then the orthogonal order requires that u and v remain linked in any valid embedding. This concept allows us to transmit information on local embedding choices over distances. Two edges $e_i = uu'$ and $e_j = vv'$ are *linked* if there is a vertex of e_i that is linked to a vertex of e_j . We use linking of edges in combination with switches. Namely, we link a switch s via its free vertex with another edge e , as illustrated in Fig. 3. Then the choice of the embedding of s determines one of the two coordinates of the linked vertices in e . In the case depicted in Fig. 3 the switch s determines the y -coordinate of both vertices of e ; we say that s *pulls* e up (Fig. 3(b)) or down (Fig. 3(c)). Such a switch is called a *vertical switch*. Analogously, edges can be pulled to the left and to the right by a *horizontal switch*.

Variable Gadgets. The variable gadget for a variable x is a simple structure consisting of a horizontal switch and a number of linked *connector vertices* on consecutive grid lines below the switch, one for each appearance of x or $\neg x$ in a clause of φ . We denote the number of appearances as $t(x)$. All connector vertices share the same x -coordinate in π . Each one will be connected to a clause with a diagonal *connector edge*. A connector edge spans the same number of grid cells horizontally and vertically and hence can only be embedded at a slope of 45° . In order to have the correct slope, the positions of both endpoints within their grid cell must be the same. The upper vertices will connect to the gadgets of the negative clauses and the lower vertices to the gadgets of the positive clauses. The variable gadget has two states, one in which the switch pulls all vertices to the left (defined as **true**), and one in which it pulls them to the right (defined as **false**). Figure 4 shows an example. A variable gadget g_x for a variable x takes up one grid cell in width and $t(x) + 1$ grid cells in height.

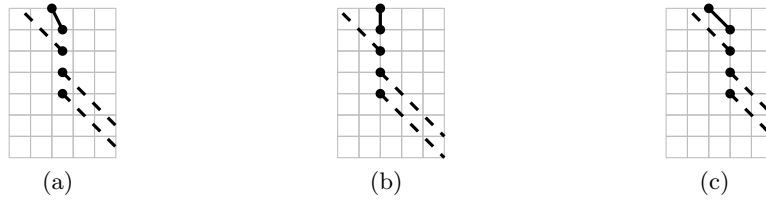


Fig. 4. (a) The input embedding of a variable gadget with three connector vertices, (b) the state `true`, and (c) the state `false`.

Clause Gadgets. The general idea behind the clause gadget is to create a set of paths whose embedding is influenced by three connector edges. Each of those edges carries the truth value of the connected variable gadget. The gadget consists of three diagonal edges e_2, e_3, e_4 , two vertical edges e_1 and e_5 and four switches s_1, s_2, s_3, s_4 , see Fig. 5. For positive clauses, we want that if all its connector edges are pulled to the right, i.e., all literals are `false`, there is no valid embedding of the clause gadget. This is achieved by placing the edges of the gadget in such a way that there are certain points, called *critical points*, where two different non-adjacent edges can possibly place a vertex. Further, we ensure with the help of switches that each of the three diagonal edges of the clause gadget has exactly one vertex embedded on a critical point in a valid embedding.

Key to the gadget is the *critical edge* e_3 that is embedded with one fixed vertex on a grid point and one loose vertex in the grid cell A to the top left of the fixed vertex. Edge e_3 is linked with switches s_2 and s_4 . These switches ensure that the loose vertex must be placed on a free corner of the grid cell. The three free corners are all critical points. It is clear that there is a valid embedding of e_3 if and only if one of the three critical points is available.

We use the remaining edges of the gadget to block one of the critical points of A for each literal that is `false`. The lower vertex of the middle connector edge is placed just to the left of the upper left corner of A such that it occupies a critical point if it is pulled to the right. Both of the left and right connector edges have another linked vertical edge e_1 and e_5 appended each. Now if e_1 is pushed to the right by its connector edge, then edge e_2 is pushed upward since e_1 and e_2 share a critical point. Due to switch s_1 edge e_2 blocks the lower left critical point of A in that case. Similarly, edge e_4 blocks the upper right critical point of A if the third literal is `false`.

The clause gadget for a negative clause works analogously such that a critical point is blocked for each connector edge that is pulled to the left instead of to the right. It corresponds basically to the positive clause gadget rotated by 180° .

Lemma 2. *A clause gadget has a valid embedding if and only if at least one of its literals is `true`.*

Proof. To see this we note that there are five critical points in the gadget and the connector edge of each literal that is `false` blocks one of the critical points. The three edges e_2, e_3, e_4 must also occupy one critical point each. So if all literals are `false`, there are only two critical points remaining for three edges and hence there is no valid embedding of the gadget. Conversely, if there is a valid embedding, then one of the connector edges does not block a critical point and hence the literal that it represents is `true`. \square

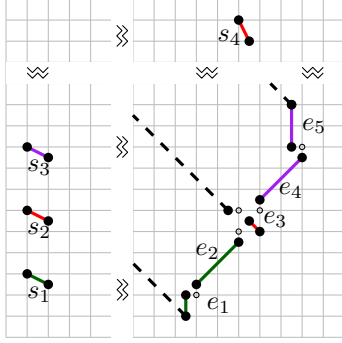


Fig. 5. Sketch of a clause gadget. Critical points are marked as white disks.

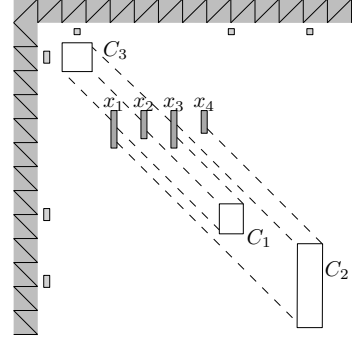


Fig. 6. Sketch of the gadgets for the instance φ from Fig.1.

The size of a clause gadget (not considering the switches) depends on the horizontal distances Δ_1, Δ_2 between the left and middle connector edge and between the middle and right connector edge, respectively. The width of a gadget is 6 and its height is $\Delta_1 + \Delta_2 - 5$.

Gadget Placement. The top and left part of our construction is occupied by the border gadget that defines the grid. The overall shape of the remaining construction is similar to a slanted version of the standard embedding of an instance of MONOTONE PLANAR 3-SAT as in Fig. 1.

We place the individual variable gadgets horizontally aligned in the center of the drawing. Since variable gadgets do not move vertically there is no danger of accidentally linking vertices of different variable gadgets by placing them at the same y -coordinates. For the clause gadgets to work as desired, we must make sure that any two connector edges to the same clause gadget have a minimum distance of seven grid cells. This can easily be achieved by spacing the variable gadgets horizontally so that connector edges of adjacent variables cannot come too close to each other.

To avoid linking between different clause gadgets or clause gadgets and variable gadgets, we place each of them in its own x - and y -interval of the grid with the negative clauses to the top left of the variables and the positive clauses to the bottom right. The switches of each clause gadget are placed at the correct positions just next to the border gadget. Figure 6 shows a sketch of the full placement of the gadgets for a MONOTONE PLANAR 3-SAT instance φ . Since the size of each gadget is polynomial in the size of the formula φ , so is the whole construction. The above construction finally yields the following theorem.

Theorem 1. *The 2-regular UNION OF PATHS SCHEMATIZATION PROBLEM is NP-hard.*

Proof. For any given MONOTONE PLANAR 3-SAT instance φ with variables $X = \{x_1, \dots, x_n\}$ and clauses $C = \{c_1, \dots, c_m\}$ we construct and place the paths for the border, variable and clause gadgets as in Section 3.1. From Lemma 2 we know that a clause gadget has a valid embedding if and only if it contains a true literal. So the whole construction has a valid embedding if and only if every clause gadget has a valid embedding. This is the case if and only if φ is satisfiable.

Both the size of the construction and the time to create it are polynomial in n and m , the number of variables and clauses in φ . \square

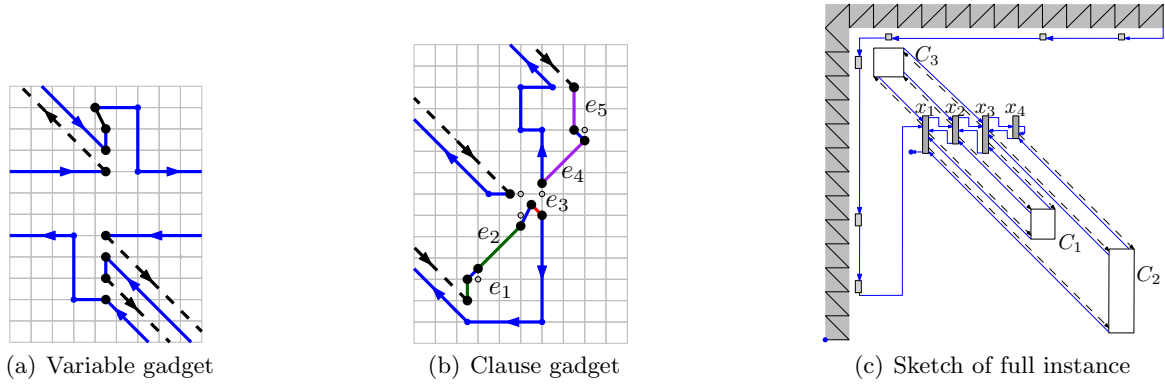


Fig. 7. Augmented gadget versions. Additional path edges are highlighted in blue.

3.2 Hardness of the Path Schematization Problem

Finally, to prove that 2-regular PSP is also NP-hard we have to show that we can augment the union of paths constructed above to form a single simple path that still has the property that it has a valid embedding if and only if the corresponding MONOTONE PLANAR 3-SAT formula φ is satisfiable.

The general idea is to start the path at the lower end of the border gadget and then collect all the switches next to the border gadget. From there we enter the upper parts of the variable gadgets and walk consecutively along all the connector edges into the negative clause gadgets. Once all negative connectors have been traversed, the path continues along the positive connectors and into the positive clauses. The additional edges and vertices must be placed such that they do not interfere with any functional part of the construction.

The only major change is that we double the number of connector vertices in each variable gadget and add a parallel dummy edge for each connector edge. That way we can walk into a clause gadget along one edge and back into the variable gadget along the other edge. We also need a clear separation between the negative and positive connector vertices, i.e., the topmost positive connector vertices of all variable gadgets are assigned the same y -coordinate and we adjust the required spacing of the variable gadgets accordingly. Figure 7(a) shows an example of an augmented variable gadget with one positive and two negative connector edges.

The edges e_1 – e_5 of each clause gadget are inserted into the path in between the leftmost connector edge and its dummy edge and in between the rightmost connector edge and its dummy edge, respectively. The details are illustrated in Fig. 7(b). It is clear that by this construction we obtain a single simple path P that contains all the gadgets of our previous reduction. Moreover, the additional edges are embedded such that they do not interfere with the gadgets themselves. Rather they can move along with the flexible parts without occupying grid points that are otherwise used by the gadgets. Hence we can summarize:

Theorem 2. *The 2-regular PATH SCHEMATIZATION PROBLEM is NP-hard.*



Fig. 8. (a) A meta cell with its three minor cells for $d = 4$; (b) input embedding of a vertical switch; (c)+(d) the two possible output embeddings of the vertical switch.

4 Hardness of d -regular PSP for $d > 2$

In Section 3 we have established that the 2-regular PATH SCHEMATIZATION PROBLEM is NP-hard. Here we show that d -regular PSP is NP-hard for all $d > 2$ by modifying the gadgets of our reduction for $d = 2$.

Theorem 3. *The d -regular PATH SCHEMATIZATION PROBLEM is NP-hard for any $d > 2$.*

Proof. An obvious problem for adapting the gadgets is that due to the presence of more than one diagonal slope the switches do not work any more for uniform grid cells. In a symmetric grid, we cannot ensure that the free vertex of a switch is always on a grid point in any valid d -schematization. This means that we need to devise a different grid in order to make the switches work properly. We construct a new border gadget that induces a grid with cells of uniform width but non-uniform heights. We call the cells of this grid *minor cells* and form groups of $d - 1$ vertically consecutive cells to form *meta cells*. Then all meta cells again have uniform widths and heights. For an example of a meta cell and its minor cells see Figure 8(a).

The $d - 1$ different heights of the minor cells are chosen such that the segments connecting the upper left corner of a meta cell to the lower right corners of its minor cells have exactly the slopes that are multiplies of $(90/d)^\circ$ and lie strictly between 0° and 90° . This restores the functionality of switches whose fixed vertex is placed on the upper left corner of a meta cell and whose free vertex is on a non-adjacent grid line of the same meta cell. See Figure 8(b) for an example of a vertical switch.

Border Gadget. As before the border gadget consists of a horizontal part B_h and a vertical part B_v . The horizontal part B_h of the border gadget consists of two smaller gadgets that we will describe next.

The *spiral gadget* is a single path consisting of $2d - 2$ edges. Its main part is the subpath that is formed by the first d edges $E_S = (e_1, \dots, e_d)$ with vertices v_1, \dots, v_{d+1} . The purpose of the remaining edges only is to ensure that several gadgets can be joined to a single path. Although the spiral does not have a unique embedding (up to scaling and translation) we do at least know that the edge e_1 uses the largest non-vertical slope. We construct it as follows.

First, we embed the edge e_1 such that v_2 is above and to the left of v_1 . The whole gadget will fit into the bounding box spanned by these two points. We embed v_{i+1} with $2 \leq i < d$ horizontally between v_{i-1} and v_i . If $i + 1$ is odd (even) we put v_{i+1} above (below) the edge e_{i-1} and below (above) the vertex v_i . Vertex v_{d+1} is embedded on the same vertical line as v_d and between v_{d-1} and v_{d-2} . We place the remaining vertices such that the whole path is crossing-free and the last edge runs vertically upward with the last vertex of the spiral in the upper right corner of the bounding box spanned by e_1 .

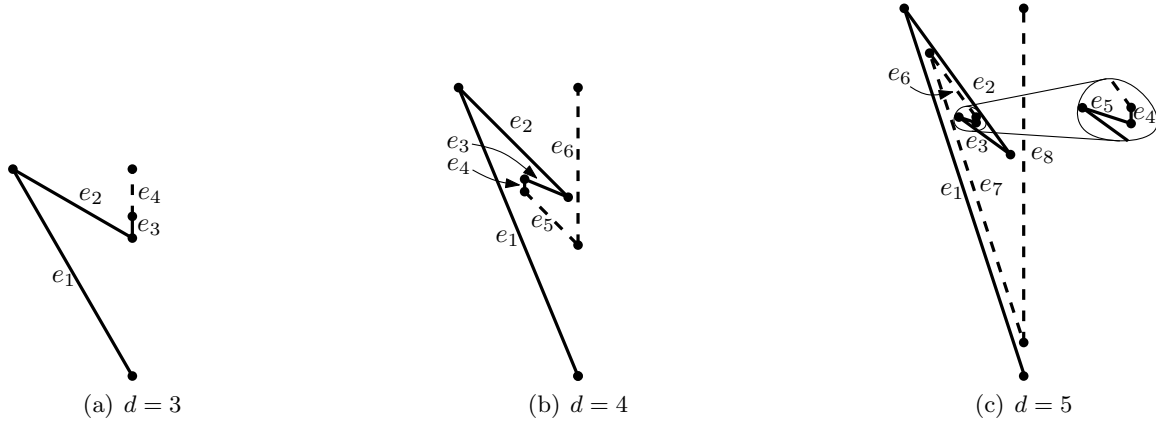


Fig. 9. Spiral gadget for different values of d .

In each valid schematization of the spiral for a fixed d we have that the edges $e_i \in E_S$ with $1 \leq i \leq d-1$ are embedded diagonally and with strictly decreasing slopes. This can be seen as follows. Each edge e_i with $i > 1$ is contained in the bounding box of e_{i-1} . Moreover, if v_{i-1} is the upper left corner of this bounding box then v_{i+1} is embedded below e_{i-1} and if v_{i-1} is the lower right corner then v_{i+1} is embedded above e_{i-1} . Hence, the edge e_i must be embedded with smaller slope than the diagonal of this bounding box, which is e_{i-1} . Moreover, edge e_1 is embedded diagonally as v_3 prevents a vertical embedding. Similarly, also edge e_{d-1} is embedded diagonally as vertex v_{d+1} prevents a horizontal embedding. Hence all edges e_i with $1 \leq i \leq d-1$ are embedded diagonally and with strictly decreasing slopes. As there are only $d-1$ different diagonal slopes in a d -schematization the slopes of these edges are unique.

We can place two copies of a spiral vertically above each other such that corresponding vertices have the same x -coordinates and join them by identifying the first vertex of the upper copy with the last vertex of the lower copy. The fact that the slope of the first edge e_1 is fixed shows that this structure is in a sense regular; both spirals have the same width and height.

The spiral provides us with a simple way of creating a construction that contains evenly spaced points. Although, at a first glance, it seems that we might be able to use the spiral gadgets directly to form a grid this is not the case since the vertices in the interior of the gadget would influence the embedding of the remaining gadgets. Hence, we need another gadget, the *stairs gadget*, which overcomes this drawback, but is not rigid by itself.

A stairs gadget consists of $d-1$ diagonal edges placed in a zig-zag pattern from bottom to top in such a way that per gadget only two x -coordinates are used (see Figure 10(a) and 10(b)). We wish that in any valid schematization the edges $e_i = v_i v_{i+1}$, $i = 1, \dots, d-1$ are all embedded diagonally and their slopes are increasing.

To achieve that the first edge is not embedded horizontally we add an additional vertex u_1 that shares its y -coordinate with v_1 and its x -coordinate with v_2 . For each i with $2 \leq i \leq d-1$ we construct two spirals that are joined as described above and such that the starting point of the first spiral has the same y -coordinate as v_i and its endpoint has the same y -coordinate as v_{i+1} . We then place an additional point u_i vertically between v_{i-1} and v_{i+1} that has the same y -coordinate as the endpoint of the upper spiral.

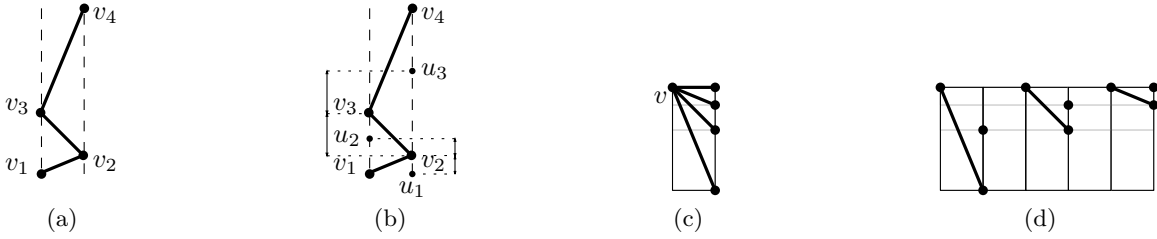


Fig. 10. Substructures of the border gadget for $d = 4$. (a) The stairs gadget; (b) stairs gadget with additional vertices; (c) non-path vertical border gadget; (d) transformation of (c) into a union of paths.

The properties of the spirals imply that the vertical distance between u_i and v_i is the same as the vertical distance between v_{i-1} and v_i in any valid d -schematization. Since v_{i+1} must be embedded strictly above u_i the edge e_{i+1} must be embedded steeper than e_i . As we have only $d - 1$ diagonal slopes the fact that the stairs gadget has $d - 1$ edges shows that each of them is used exactly once and in increasing order along the path.

Placing k identical stairs directly next to each other yields k unit-width grid rows. If we link the corresponding vertices of the stairs gadgets with each other we need to place the vertices u_i only in one of the stairs gadgets. Since we want to combine all gadgets to a path in the end, we use this observation and create a single stairs gadget with the additional vertices to the left of all other stairs gadgets. We can now create a path starting in this leftmost stairs gadget and then connecting the remaining $k - 1$ stairs gadgets.

Now that we have established the horizontal part B_h of the border gadget we can assume that all grid columns are of equal width 1. Hence the meta grid cells have height $\tan((d - 1)/d \cdot 90^\circ)$. We want to choose the heights of the minor cells such that the line segments from the upper left corner of the meta cell to all grid points on the right side of the meta cell have a slope in \mathcal{C}_d , see Figure 10(c). This property is necessary for constructing switches as explained later.

Unfortunately, the gadget in Figure 10(c) is not a path and hence we cannot use it to induce the required grid structure directly. Instead, we “stretch” the structure and use the concept of linking in order to yield the same result. We use $d - 1$ grid columns—one for each diagonal slope—with one “empty” grid column between any two non-empty grid columns (see Figure 10(d)). To induce ℓ grid rows we repeat this structure ℓ times.

The horizontal and vertical parts of the border gadget are combined by placing the vertical parts such that all vertices are placed on grid lines. A sketch of the complete border gadget for $d = 4$ is shown in Figure 11.

Switches. Since the border gadget induces a non-regular grid the switches must be adjusted. Each switch starts at the upper left corner of a meta cell. For a horizontal switch the free vertex is placed on the lower edge of the meta cell. For a vertical switch the free vertex is placed on the right edge of the meta cell in between two horizontal grid lines.

Variable and Clause Gadgets. The variable gadget is unchanged given that all its vertices are placed on meta grid lines. The overall structure and functionality of the clause gadgets remains unchanged as well. However, we need to do some slight modifications, see Figure 12. The critical edge e_3 is placed in the topmost minor cell of a meta cell. Due to the nature

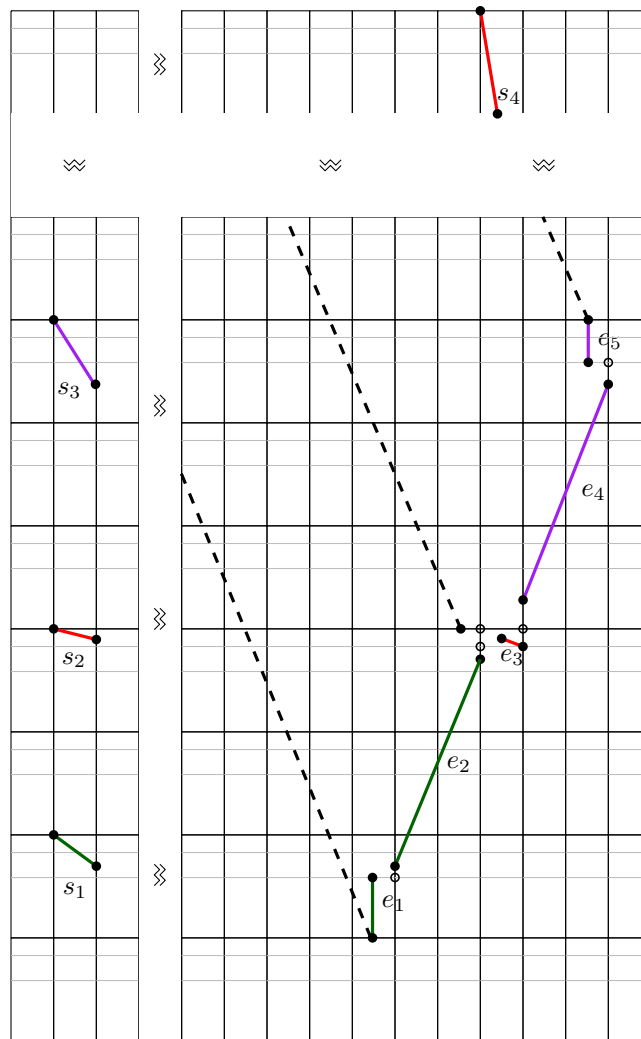


Fig. 12. Clause gadget for $d = 4$.

formulate the problem as a mixed integer linear program (MIP) and then justify our approach through an experimental study by showing that for the application of drawing sketches for real-world driving directions our MIP yields good results in reasonable time frames.

5.1 Modeling

Let $\mathcal{I} = (P, \pi, d)$ be an input instance of PSP consisting of a path P , an input embedding π and an integer d . For an embedding π of P and an edge e of P we denote by $\alpha_\pi(e)$ the slope of e in π . Besides deciding whether \mathcal{I} admits a valid d -schematization, we transform PSP into the following optimization problem in order to find a valid d -schematization that is as similar to the input path as possible.

Problem 1. Given a simple embedded path $(P = (V, E), \pi)$ and an integer $d \geq 1$, find a valid d -schematization ρ such that

- (i) for every edge $e \in E$ the deviation of $\alpha_\rho(e)$ from $\alpha_\pi(e)$ is minimal,
- (ii) for every edge $e \in E$ the length $\ell(e)$ is at least a minimum length $\ell_{\min}(e) > 0$, and
- (iii) the total length of the path $\sum_{e \in E} \ell(e)$ is minimized.

In the following, we introduce our MIP that solves Problem 1 optimally.

Coordinate System. In order to handle the valid slopes of each edge with respect to \mathcal{C}_d , we extend the standard Cartesian coordinate system to include one axis for every slope contained in \mathcal{C}_d similarly to [16]. Therefore, we introduce continuous variables $\text{pos} : V \times \mathcal{C}_d \rightarrow \mathbb{R}$ for every vertex and every possible slope. The value for each coordinate axis $\gamma \in \mathcal{C}_d$ is then defined by $\text{pos}(v, \gamma) = \cos \gamma \cdot x(v) + \sin \gamma \cdot y(v)$. Note, that these are linear constraints, as the values of $\sin \gamma$ and $\cos \gamma$ are constants for fixed d . Moreover, our coordinate system allows us to introduce binary variables $\text{dir} : E \times \mathcal{C}_d \rightarrow \{0, 1\}$ for deciding whether in the output embedding ρ an edge $e \in E$ is schematized with slope $\gamma \in \mathcal{C}_d$, namely, if and only if $\text{dir}(e, \gamma) = 1$. To enforce that every edge is schematized according to exactly one slope $\gamma \in \mathcal{C}_d$, we add $\sum_{\gamma \in \mathcal{C}_d} \text{dir}(e, \gamma) = 1$ as a constraint for every edge $e \in E$.

Orthogonal Order. To maintain the orthogonal order between all pairs of vertices $v_i, v_j \in V$, we can make use of the following observation. Let $v_1 \prec \dots \prec v_n$ be the input x -order of the vertices (with ties broken arbitrarily). Then we introduce a linear number of orthogonal order constraints as follows. We set $\text{pos}(v_i, 0^\circ) \leq \text{pos}(v_{i+1}, 0^\circ)$ if $x_\pi(v_i) < x_\pi(v_{i+1})$ and $\text{pos}(v_i, 0^\circ) = \text{pos}(v_{i+1}, 0^\circ)$ if $x_\pi(v_i) = x_\pi(v_{i+1})$. Analogously, the same constraints are independently introduced for the y -order of the vertices and the values $\text{pos}(v, 90^\circ)$ for all $v \in V$.

Edge Slopes and Lengths. To ensure that every edge $e \in E$ is embedded according to the slope γ for which $\text{dir}(e, \gamma) = 1$ holds, we make use of our extended coordinate system in the following way. The edge $e = uv$ is embedded according to $\gamma \in \mathcal{C}_d$ if and only if u and v have the same coordinate on both orthogonal axes $\gamma + 90^\circ$ and $\gamma + 270^\circ$. Thus, for every $e = uv \in E$ and every $\gamma \in \mathcal{C}_d$ we add the constraints $\text{pos}(v, \gamma + 90^\circ) - \text{pos}(u, \gamma + 90^\circ) \leq M(1 - \text{dir}(e, \gamma))$ and $\text{pos}(v, \gamma + 270^\circ) - \text{pos}(u, \gamma + 270^\circ) \leq M(1 - \text{dir}(e, \gamma))$. Here, M is a sufficiently large constant, which we choose as the maximum possible distance in our coordinate space. This has the effect that the constraints are trivially satisfied for $\text{dir}(e, \gamma) = 0$ and that they enforce

the correct slope for $\text{dir}(e, \gamma) = 1$. Additionally, we can set the minimum length for every edge by adding constraints $\text{pos}(v, \gamma) - \text{pos}(u, \gamma) \geq -M(1 - \text{dir}(e, \gamma)) + \ell_{\min}(e)$. Again, if M is sufficiently large, these constraints are relevant only if $\text{dir}(e, \gamma) = 1$. In order to minimize the total path length, we define variables $\text{len} : E \rightarrow \mathbb{R}_0^+$ as an upper bound on the length of each edge $e = uv \in E$ by introducing for every $\gamma \in \mathcal{C}_d$ a constraint $\text{pos}(v, \gamma) - \text{pos}(u, \gamma) \leq \text{len}(e)$. Note that minimizing $\text{len}(e)$ through the objective functions will result in tight upper bounds on the actual edge lengths, i. e., $\text{len}(e) = \ell(e)$.

To minimize the deviation of the schematized slope of every edge $e \in E$ to its input slope, we add continuous variables $\text{dev} : E \rightarrow \mathbb{R}_0^+$ that represent the deviation cost of e according to its embedding in ρ . We set $\text{dev}(e) = \sum_{\gamma \in \mathcal{C}_d} c(e, \gamma) \text{dir}(e, \gamma)$, where $c : E \times \mathcal{C}_d \rightarrow \mathbb{R}_0^+$ is some user-defined deviation cost function. In our implementation we set $c(e, \gamma)$ to the distance in terms of the number of slopes in \mathcal{C}_d between the selected slope and the slope in \mathcal{C}_d that is closest to the input slope $\alpha_\pi(e)$.

Planarity. In the output embedding ρ , edges may be stretched and embedded with slopes different from the input embedding. Thus, we have to ensure that the output embedding is still crossing-free. Hence, we introduce two types of planarity constraints. First, for two adjacent edges $e_i = uv$ and $e_j = vw$ we observe that e_i and e_j only intersect, if $\gamma(e_i) = \gamma(e_j) + 180^\circ$. Thus, for every pair of adjacent edges e_i, e_j and every $\gamma \in \mathcal{C}_d$ we add a constraint $\text{dir}(e_i, \gamma) + \text{dir}(e_j, \gamma + 180^\circ) \leq 1$.

For the remaining non-adjacent pairs of edges $e_i = u_i v_i, e_j = u_j v_j$ we can make use of the following observation. The edges e_i and e_j do not intersect if there is at least one coordinate axis $\gamma \in \mathcal{C}_d$ according to which both u_j and v_j lie beyond u_i and v_i . Thus, we introduce binary variables $\text{sep} : E \times E \times \mathcal{C}_d \rightarrow \{0, 1\}$ selecting for each pair of non-adjacent edges the axis γ for which the previous observation needs to be true. We can now enforce planarity by adding four constraints for every pair e_i, e_j and every $\gamma \in \mathcal{C}_d$. We set $\text{pos}(X, \gamma) - \text{pos}(Y, \gamma) \geq -M(1 - \text{sep}(e_i, e_j, \gamma)) + \delta$, where $X \in \{u_i, v_i\}$, $Y \in \{u_j, v_j\}$, and $\delta \in \mathbb{R}_0^+$ is a user-defined minimum distance to be kept between non-adjacent edge-pairs.

Note that this approach to enforce planarity requires $O(|E|^2)$ binary variables and constraints. However, in our experiments on real-world driving directions, we observe that most of the planarity constraints seem to be unnecessary in the sense that the optimal solution without planarity constraints already yields a crossing-free path. Thus, we may add planarity constraints in a lazy fashion by an iterative process: We start by computing a solution without planarity constraints. Then, we test whether the solution has any intersections and add planarity constraints for these specific conflicting edge pairs. We then recompute a solution with the new set of constraints until we finally obtain a crossing-free embedding ρ .

Objective Function. Subject to all the constraints described above we minimize the following objective function:

$$\text{Minimize} \quad \sum_{e \in E} \text{dev}(e) + \frac{1}{M} \sum_{e \in E} \text{len}(e).$$

Note that the coefficient $1/M$ ensures that our primary goal is to minimize the total angular deviation as modeled by $\sum_{e \in E} \text{dev}(e)$, while minimizing the total path length becomes the secondary goal.

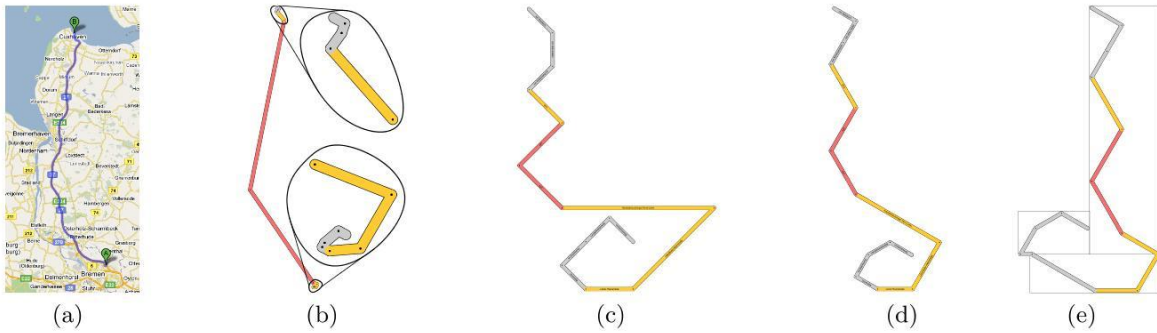


Fig. 13. A sample route from Bremen to Cuxhaven in Germany: (a) output of Google Maps, (b) simplified route after preprocessing, (c) output of our MIP for $d = 2$, (d) output of our MIP for $d = 3$, and (e) output of the algorithm by Delling et al. [7].

Experiments. We have shown that PSP can be formulated as a MIP that is similar to a MIP model for drawing octilinear metro maps [16]. This model was implemented for simple input paths and we evaluate its performance by schematizing 1 000 quickest routes in the German road network, where we select source and target nodes uniformly at random. The average path length is 1020.7 nodes, hence, a schematization would yield a route sketch with far too much detail. Therefore, in a preprocessing step, we simplify the path using the Douglas-Peucker algorithm [8] but keep all important nodes such as points of road and road category changes. Moreover, we shortcut self-intersecting subpaths whose lengths are below a threshold. This is to remove over- or underpasses near slip roads of highways, which are considered irrelevant for route sketches or would be depicted as an icon rather than a loop. Figure 13 illustrates an example of a route sketch obtained by our MIP. It clearly illustrates how a schematic route sketch, unlike the geographic map, is able to depict both high-detail and low-detail parts of the route in a single picture. Comparing Figures 13(c) and 13(d) indicates that using the parameter $d = 2$ for the schematization process yields route sketches with a very high level of abstraction while using $d = 3$ results in route sketches which resemble the overall shape of the route much better. For example, the route sketch in Figure 13(c) seems to suggest that there is a 90° turn while driving on the highway but this is not the case. The route sketch depicted in Figure 13(c) resembles the original route much better. Generally, we found that using the parameter $d = 3$ yields route sketches with a higher degree of readability.

For comparison, we show the output computed by the method of Delling et al. [7] in Figure 13(e). Recall that their original algorithm takes as input only monotone paths. They describe, however, a heuristic approach to schematize non-monotone paths by subdividing them into maximal monotone subpaths that are subsequently merged into a single route sketch. In order to avoid intersections of the bounding boxes of the monotone subpaths, additional edges of appropriate length must be inserted into the path; the orthogonal order is preserved only within the monotone subpaths. This may lead to undesired effects in sketches of non-monotone routes, see Figure 13(e).

5.2 Experimental Evaluation

We performed two experiments that were executed on a single core of an AMD Opteron 2218 processor running Linux 2.6.27.23. The machine is clocked at 2.6 GHz, has 16 GiB of RAM

Table 1. Results obtained by running 1 000 random queries in the German road network and schematizing the simplified paths with our MIP. The tables reports average path lengths, percentage of infeasible instances, average number of iterations, and average running times.

(a) Varying ε with $d = 3$.					(b) Varying d with $\varepsilon = 2^{-3}$.			
ε	\sim length	% inf	\sim iter	\sim time [ms]	d	% inf	\sim iter	\sim time [ms]
2^{-1}	20.04	0.1	1.28	801.61	1	51.3	1.03	107.36
2^{-2}	20.50	0.1	1.29	621.66	2	0.7	1.19	363.49
2^{-3}	21.81	0.1	1.29	649.49	3	0.1	1.29	649.49
2^{-4}	25.16	0.2	1.26	772.57	4	0.1	1.25	1075.82
2^{-5}	32.81	0.3	1.27	1102.20	5	0.1	1.21	1347.86

and 2×1 MiB of L2 cache. Our implementation is written in C++ and was compiled with GCC 4.3.2, using optimization level 3. As MIP solver we use Gurobi 3.0.1.

In the first experiment we examine the performance of our MIP for $d = 3$ subject to the amount of detail of the route, as controlled by the distance threshold ε between input and output path in the path simplification step. We also set the threshold for removing self-intersecting subpaths to ε . Table 1(a) reports our experimental results for values of ε between 2^{-1} and 2^{-5} .

We observe that with decreasing ε , the lengths of the paths increase from 20.04 nodes to 32.81 nodes on average. This correlates with the running time of our MIP which is between 801.61 ms and 1 102.20 ms. Note that in practice a value of $\varepsilon = 2^{-3}$ is a good compromise between computation time and amount of detail. Further, we observe that adding planarity constraints in a lazy fashion pays off since we need less than 1.3 iterations on average. The number of infeasible instances, i. e., paths without a valid 3-schematization, is 0.1%.

The second experiment evaluates the performance of our MIP when using different values of d , see Table 1(b). We fix $\varepsilon = 2^{-3}$. While for rectilinear drawings with $d = 1$ we need 107.36 ms to compute a solution, the running time increases to 1 347.86 ms when using $d = 5$. Interestingly, more than half of the paths do not have a valid rectilinear schematization. By allowing one additional diagonal slope ($d = 2$), the number of infeasible instances significantly decreases to 0.7% and for $d = 3$ only 0.1%, i. e., a single instance, is infeasible.

Further Examples. In Figures 14 and 15 are two example routes as displayed in Google Maps (left), simplified route determined by the Douglas-Peucker algorithm (middle), and schematized by our MIP (right).

6 Conclusion

Motivated by drawing route sketches in road networks, we studied the problem of d -regular schematization of embedded paths (PSP). In our problem definition we aimed for two main goals: To preserve the user’s mental map we used the concept of orthogonal order, and to reduce the visual complexity we restricted the valid edge slopes to integer multiples of $(90/d)^\circ$. We analyzed the complexity of the problem and showed that PSP is NP-hard for $d \geq 2$, thus, closing the complexity gap between the hardness result of Brandes and Pampel [4] for the special case of $d = 1$, and the polynomial-time algorithm of Delling et al. [7] for monotone paths. In the second part of this work, we modeled the PSP as a mixed integer linear program (MIP). To generate easily readable drawings, in our MIP we minimized the total path length

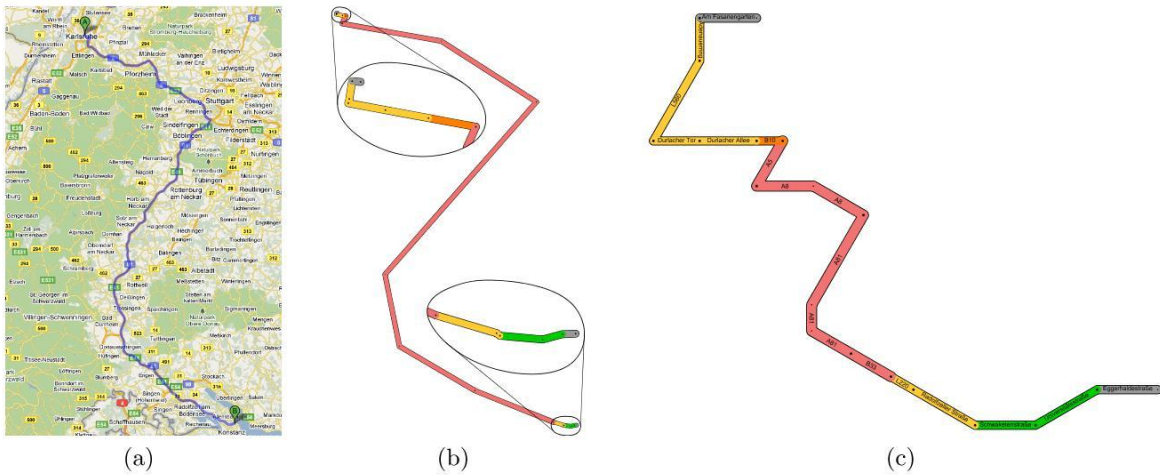


Fig. 14. Route from Karlsruhe to Konstanz. Parameters used: $\varepsilon = 2^{-3}$ and $d = 3$. Time to compute MIP solution: 262ms.

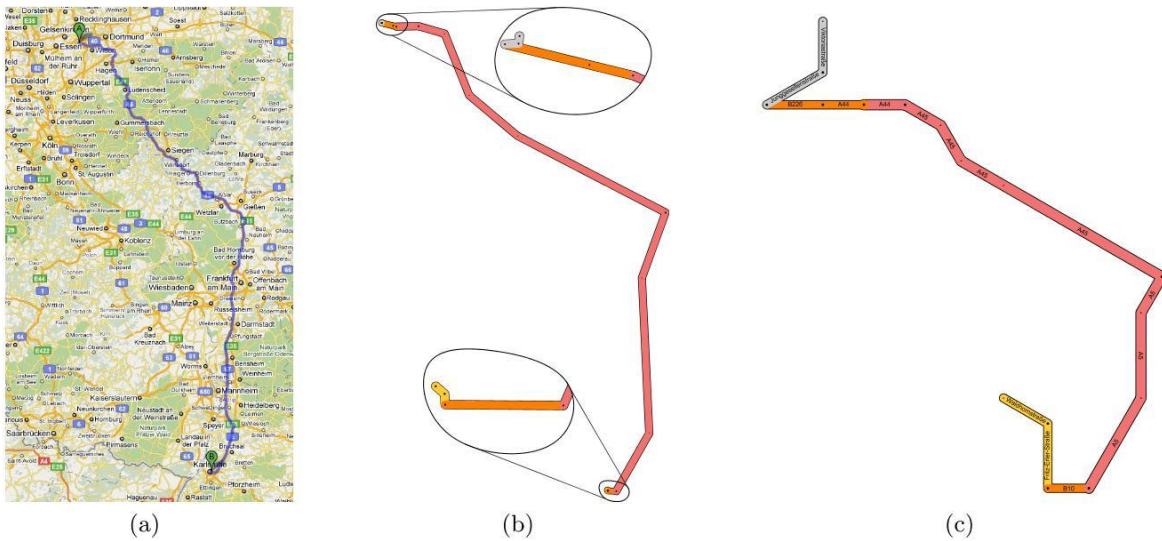


Fig. 15. Route from Bochum to Karlsruhe. Parameters used: $\varepsilon = 2^{-4}$, and $d = 3$. Time to compute MIP solution: 126 ms.

via the objective function while ensuring a certain minimum length for each individual edge. An experimental evaluation on real-world data in the German road network revealed that we are indeed able to compute solutions to the PSP within approximately 1 sec for reasonable values of $d \leq 5$, producing visually appealing drawings.

Using ideas of Nöllenburg and Wolff [16], our MIP can be further generalized to handle both non-simple paths and general graphs, e. g., a set of alternative routes.

Acknowledgements. We thank an anonymous referee of [10] for helpful comments.

References

1. M. Agrawala and C. Stolte. Rendering effective route maps: Improving usability through generalization. *Proc. 28th Ann. Conf. Computer Graphics and Interactive Techniques (SIGGRAPH'01)*, pp. 241–249. ACM, 2001.
2. M. de Berg and A. Khosravi. Finding perfect auto-partitions is NP-hard. *Proc. 25th European Workshop Comput. Geom. (EuroCG'08)*, pp. 255–258, 2008.
3. U. Brandes, M. Eiglsperger, M. Kaufmann, and D. Wagner. Sketch-driven orthogonal graph drawing. *Proc. 10th International Symposium on Graph Drawing (GD'02)*, pp. 1–11. Springer-Verlag, Lecture Notes Comput. Sci. 2528, 2002.
4. U. Brandes and B. Pampel. On the hardness of orthogonal-order preserving graph drawing. *Proc. 16th Internat. Symp. Graph Drawing (GD'08)*, pp. 266–277. Springer-Verlag, Lecture Notes Comput. Sci. 5417, 2009.
5. J. Branke. Dynamic graph drawing. *Drawing Graphs: Methods and Models*, chapter 9, pp. 228–246. Springer-Verlag, Lecture Notes Comput. Sci. 2025, 2001.
6. S. Cabello, M. d. Berg, S. v. Dijk, M. v. Kreveld, and T. Strijk. Schematization of road networks. *Proc. 17th Annual Symposium on Computational Geometry (SoCG'01)*, pp. 33–39. ACM Press, 3–5 June 2001.
7. D. Delling, A. Gemsa, M. Nöllenburg, and T. Pajor. Path schematization for route sketches. *Proc. 12th Scand. Symp. & Workshops on Algorithm Theory (SWAT'10)*, pp. 285–296. Springer-Verlag, Lecture Notes Comput. Sci. 6139, 2010.
8. D. H. Douglas and T. K. Peucker. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica* 10(2):112–122, 1973.
9. T. Dwyer, Y. Koren, and K. Marriott. Stress majorization with orthogonal ordering constraints. *Proc. 13th Internat. Symp. Graph Drawing (GD'05)*, pp. 141–152. Springer-Verlag, Lecture Notes Comput. Sci. 3843, 2006.
10. A. Gemsa, M. Nöllenburg, T. Pajor, and I. Rutter. On d -regular Schematization of Embedded Paths. To appear in *Proc. 37th International Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM'11)*. Springer-Verlag, Lecture Notes Comput. Sci., 2011.
11. K. Hayashi, M. Inoue, T. Masuzawa, and H. Fujiwara. A layout adjustment problem for disjoint rectangles preserving orthogonal order. *Proc. 6th Internat. Symp. Graph Drawing (GD'98)*, pp. 183–197. Springer-Verlag, Lecture Notes Comput. Sci. 1547, 1998.
12. D. Merrick and J. Gudmundsson. Path simplification for metro map layout. *Proc. 14th Internat. Symp. Graph Drawing (GD'06)*, pp. 258–269. Springer-Verlag, Lecture Notes Comput. Sci. 4372, 2007.
13. K. Misue, P. Eades, W. Lai, and K. Sugiyama. Layout adjustment and the mental map. *J. Visual Languages and Computing* 6(2):183–210, 1995.
14. G. Neyer. Line simplification with restricted orientations. *Proc. 6th Int. Workshop on Algorithms and Data Structures (WADS'99)*, pp. 13–24. Springer-Verlag, Lecture Notes Comput. Sci. 1663, 1999.
15. M. Nöllenburg. Automated drawing of metro maps. Tech. Rep. 2005-25, Fakultät für Informatik, Universität Karlsruhe, 2005, <http://digbib.ubka.uni-karlsruhe.de/volltexte/1000004123>.
16. M. Nöllenburg and A. Wolff. Drawing and Labeling High-Quality Metro Maps by Mixed-Integer Programming. *IEEE Transactions on Visualization and Computer Graphics*, 2010, <http://dx.doi.org/10.1109/TVCG.2010.81>. Preprint available online.
17. A. Wolff. Drawing subway maps: A survey. *Informatik - Forschung und Entwicklung* 22(1):23–44, 2007.