Diskrete Mathematik und Optimierung

Stephan Dominique Andres:

Greedy versus Recursive Greedy: Uncorrelated Heuristics for the Binary Paint Shop Problem

Technical Report feu-dmo064.19
Contact: dominique.andres@fernuni-hagen.de

FernUniversität in Hagen
Fakultät für Mathematik und Informatik
Lehrgebiet für Diskrete Mathematik und Optimierung
D – 58084 Hagen

2000 Mathematics Subject Classification: 68W40, 68R15
Keywords: binary paint shop problem, greedy heuristic, recursive greedy heuristic, red first heuristic, greedy-perfect, recursive-greedy-perfect, red-first-perfect, APX-hardness

Greedy versus Recursive Greedy: Uncorrelated Heuristics for the Binary Paint Shop Problem

Stephan Dominique Andres

Faculty of Mathematics and Computer Science,
FernUniversität in Hagen
Universitätsstr. 1, 58084 Hagen, Germany
dominique.andres@fernuni-hagen.de

27 November 2019

Abstract

It is well-known that there are instances of the binary paint shop problem for which the recursive greedy heuristic is better than the greedy heuristic. In this note, we give an example of a family of instances where the greedy heuristic is better than the recursive greedy heuristic, thus showing that these heuristics are uncorrelated.

Key words: binary paint shop problem, greedy heuristic, recursive greedy heuristic, red first heuristic, greedy-perfect, recursive-greedy-perfect, red-first-perfect, APX-hardness

MSC 2000: 68W40, 68R15

1 Introduction

In the *binary paint shop problem*, we are given two colours (red and blue) and a *double occurrence word*, i.e. a word $w \in \Sigma^{2n}$ of length $2n$ over an alphabet Σ of size $|\Sigma| = n$, in which every character of Σ occurs exactly twice as a letter of the word w. A *feasible colouring* of w is an assignment of colours to the letters of w such that every character of Σ is coloured once in red and once in blue. For such a feasible colouring, a *colour change* is an index $i \in \{1, \ldots, 2n-1\}$ such that the colours of w_i and w_{i+1} are different. We are looking for a feasible colouring that minimises the number of colour changes.

This problem is a special case of the so-called *paint shop problem* that was motivated by an application in car manufacturing. Both problems were introduced by Epping et al. [5].

Bonsma et al. [3] proved that the binary paint shop problem is APX-hard. Thus, it is NP-complete. By another reduction this APX-hardness result was reproved by Meunier and Sebő [9]. It is not known whether the problem is in

APX. Gupta et al. [6] proved that the problem is not in APX if the Unique Games Conjecture proposed by Subhash Khot [7] is true.

A straightforward heuristic, the *greedy heuristic*, which scans the word from left to right, was already introduced by Epping et al. [5] and considered meanwhile by several authors [1, 2, 3, 5, 8, 9, 10]. Amini et al. [1] proved that the expected value of colour changes for the greedy heuristic is at most $2n/3$ and conjectured that it is asymptotically $n/2$. Andres and Hochstättler [2] proved this conjecture. Moreover, they considered two other heuristics, the *red first* and the *recursive greedy heuristic* and determined the asymptotic behaviour of the expected number of colour changes for these heuristics as $2n/3$ and $2n/5$. Recently, Šámal et al. [11] announced a proof showing that the expected number of colour changes for optimal colouring is at least $107n/500 + o(n)$.

The better asymptotic behaviour of the recursive greedy heuristic raises the question whether, for any instance of the binary paint shop problem, the recursive greedy heuristic is better than the greedy heuristic. In this paper, we answer this question to the negative by giving a counterexample of a family of instances with a linear gap in the number of colour changes.

The rest of the paper is structured as follows. In Section 2, we define the mentioned heuristics and give an example of non-optimality for all of these heuristics. Section 3 is devoted to the counterexample. In the final Section 4 we make some observations and conjectures concerning the relation between the heuristics and optimal colouring.

2 Heuristics for the Binary Paint Shop Problem

We consider the following three heuristics for an instance $s = (s_1, \dots, s_{2n})$ of the binary paint shop problem. Denote the available colours as *red* and *blue*.

Greedy heuristic Start with colour red. Scan the letters s_i of the word from left ($i = 1$) to right ($i = 2n$) and change the colour only when it is necessary.

Red First heuristic The first occurrence of each character is coloured red, the second blue.

Recursive greedy heuristic Let $w = w_1 w_2 w_3 \dots w_{2n-1} w_{2n}$ be a double occurrence word over an alphabet with $n \geq 1$ characters. If $n = 1$, then the word w consists of two equal letters, and we colour the first one red and the second one blue. If $n \geq 2$, we delete both occurrences of the last letter w_{2n}, which forms a subword w' over an alphabet with $n - 1 \geq 1$ characters. Then we colour w' recursively using the recursive greedy heuristic. The corresponding letters of w receive the same colours as their image in w'. So we are left with colouring the two occurrences of w_{2n}. If the first occurrence of w_{2n} is between two letters of the same colour (1.1) or has only one coloured neighbour (1.2), then we colour it with this colour and the second occurrence with the other. Otherwise we have $w_{2n-1} \neq w_{2n}$ and we colour (the second occurrence of) w_{2n}

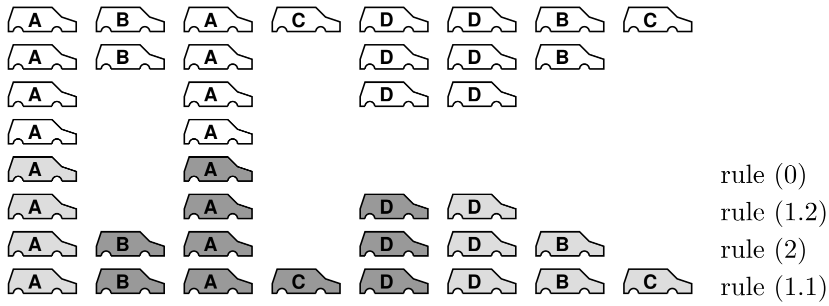

Figure 1: The dynamics of the recursive greedy algorithm.

with the colour of w_{2n-1} (2). Note that then by colouring the first occurrence of w_{2n} with the other colour we do not produce a new colour change, since this letter is between both colours. In Figure 1 we provide a simple example that uses all four rules (0), (1.1), (1.2) and (2). Instead of the colours red and blue we use the colours light gray and dark gray.

Heuristics versus optimality The number of colour changes produced by the three heuristics and the optimal colouring may differ significantly. Moreover, there are examples, where all three heuristics are arbitrarily bad. Example 1 combines examples that are already given in [2, 3, 5].

Example 1. *For every $k \in \mathbb{N}$ with $k \geq 2$, the instance*

$$A_1 \dots A_k C C B_1 \dots B_k (A_1 B_1) \dots (A_k B_k) D_1 \dots D_k (D_1 E_1) \dots (D_k E_k) E_1 \dots E_k$$

of length $2n = 8k + 2$ (with $n = 4k + 1$ different characters) of the binary paint shop problem is coloured with

- *5 colour changes optimally;*
- *$2k + 5$ colour changes by the red first heuristic;*
- *$2k + 2$ colour changes by the greedy or recursive greedy heuristic.*

In the following, underlined letters are red, overlined letters blue.

Proof of Example 1. In any feasible colouring, there must be a colour change between the two Cs, a colour change between the two B_1s and an odd number of colour changes between the two A_2s. All in all, there must be at least three colour changes between the two A_2s. Furthermore, there must be one colour change between the two D_1s and another one between the two E_1s. Thus, in any feasible colouring, we need at least five colour changes. An optimal colouring with five colour changes is the following.

$$\underline{A_1 \dots A_k C}|\overline{C}|\underline{B_1 \dots B_k}|\overline{A_1 B_1 \dots A_k B_k D_1 \dots D_k}|\underline{D_1 E_1 \dots D_k E_k}|\overline{E_1 \dots E_k}$$

The red first heuristic produces the following coloring.

$$\underline{A_1 \dots A_k C}|\overline{C}|\underline{B_1 \dots B_k}|\overline{A_1 B_1 \dots A_k B_k}|\underline{D_1 \dots D_k}|\overline{D_1}|\underline{E_1}| \dots \overline{D_k}|\underline{E_k}|\overline{E_1 \dots E_k}$$

The greedy and the recursive greedy heuristic colour the word as follows.

$$\underline{A_1 \dots A_k C}|\overline{C B_1 \dots B_k A_1}|\underline{B_1}| \dots \overline{A_k}|\underline{B_k D_1 \dots D_k}|\overline{D_1 E_1 \dots D_k E_k}|\underline{E_1 \dots E_k}$$

$\square$

3 The counterexample

There are simple instances where the recursive greedy heuristic is better than the greedy heuristic, the smallest such example being $ACABBC$. One might suppose that the recursive greedy heuristic is always better than or as good as the greedy heuristic. However, this is not true, as the following proposition shows.

Proposition 2. *For every $k \in \mathbb{N}$, there is an instance $s \in \Sigma^{2\cdot(4k+12)}$ of the binary paint shop problem, for which the number of colour changes produced is*

- $2k + 8$ *for the greedy heuristic and*
- $4k + 6$ *for the recursive greedy heuristic.*

Proof. We consider the alphabet

$$\Sigma = \{A, B, C, D, E, F, A', B', C', D', E', F'\} \cup \{X_i, Y_i \mid 1 \leq i \leq 2k\}$$

and construct the instance $s \in \Sigma^{8k+24}$ as the concatenation $s = ww'w''$ of the three words

$$\begin{aligned} w &= ACADBBCEFDEY_1Y_3Y_5 \cdots Y_{2k-1}F, \\ w' &= A'C'A'D'B'B'C'E'Y_2Y_4Y_6 \cdots Y_{2k}F'D'E'F', \\ w'' &= X_1X_1Y_1\ X_2X_2Y_2\ X_3X_3Y_3 \cdots X_{2k}X_{2k}Y_{2k}. \end{aligned}$$

The greedy heuristic produces 4 colour changes for each of the words w and w', so that the first and the last letter are coloured red:

$$\begin{array}{llllllllllll} w = & \underline{A}\ \underline{C} & | & \overline{A}\ \overline{D}\ \overline{B} & | & \underline{B} & | & \overline{C}\ \overline{E} & & \overline{F} & | & \underline{D}\ \underline{E}\ \underline{Y_1Y_3Y_5 \cdots Y_{2k-1}}\ \underline{F} \\ w' = & \underline{A'}\ \underline{C'} & | & \overline{A'}\ \overline{D'}\ \overline{B'} & | & \underline{B'} & | & \overline{C'}\ \overline{E'} & \overline{Y_2Y_4Y_6 \cdots Y_{2k}} & \overline{F'} & | & \underline{D'}\ \underline{E'}\ \underline{F'} \end{array}$$

Then, in each subword of w'' of the form

$$\underline{X_{2i-1}}|\overline{X_{2i-1}Y_{2i-1}X_{2i}}|\underline{X_{2i}Y_{2i}} \quad (i \in \{1, \dots, k\})$$

the greedy heuristic only needs two colour changes (between the two equal X-letters) and the first and the last letter of each such subword are coloured red.

In particular, the Y_i with odd i are coloured red and the Y_i with even i are coloured blue.

The recursive greedy heuristic produces 3 colour changes for each of the words w and w', so that the corresponding letters in w and w' (neglecting the letters Y_i) have different colours:

$$\begin{aligned} w \; &= \underline{A} \,|\, \overline{C}\; \overline{A}\; \overline{D}\; \overline{B} \,|\, \underline{B}\; \underline{C}\; \underline{E} \qquad\qquad\qquad\quad \underline{F}\; \underline{D} \,|\, \overline{E}\; Y_1 Y_3 Y_5 \cdots Y_{2k-1}\; \overline{F} \\ w' &= \overline{A'} \,|\, \underline{C'}\, \underline{A'}\, \underline{D'}\, \underline{B'} \,|\, \overline{B'}\, \overline{C'}\, \overline{E'}\; Y_2 Y_4 Y_6 \cdots Y_{2k}\; \overline{F'}\, \overline{D'} \,|\, \underline{E'} \qquad\qquad\qquad \underline{F'} \end{aligned}$$

In particular, all Y_i must be coloured blue by the recursive greedy heuristic. Thus, every subword of w'' of the form $\underline{X_i}|\overline{X_i}|\underline{Y_i}$ must be coloured red-blue-red, which gives us two colour changes for each such subword, or four colour changes for each pair of such words. $\square$

4 Final remarks and open questions

The following questions are still open to our knowledge.

Problem 3. *Are there series of instances for which one of the greedy or recursive greedy heuristic performs well (i.e. the number of colour changes is bounded by a polynomial in the number of optimal colour changes) but the other heuristic is arbitrarily bad?*

Problem 4. *Is it true that, for instances for which the greedy heuristic is optimal, the recursive greedy heuristic is optimal, too?*

In Table 1 we see examples for six cases of optimality/non-optimality for the three heuristics different from the cases that were excluded in case of an affirmative answer to Problem 4.

	red first: not optimal	red first: optimal
recursive greedy and greedy: not optimal	$BDAACEBCDEFGFHGIHI$ (=Example 1, $k = 2$)	$BDAACEBCDE$
recursive greedy: optimal; greedy: not	$ACABBC$	$CEAABDBCDE$
recursive greedy and greedy: optimal	$AABB$	AA

Table 1: Small non-trivial optimal and non-optimal instances of the three paint shop heuristics.

We might define *h-perfect* instances of the binary paint shop problem concerning a heuristic h as those instances w where, for every subword of w, the heuristic produces an optimal solution. Rautenbach and Szigeti [10] characterised the perfect instances concerning the greedy heuristic.

Theorem 5 ([10])**.** *An instance of the the binary paint shop problem is greedy-perfect if and only if it does not contain a subword of the form*

$$ACABBC, \quad CAADBBCD, \quad \text{or} \quad CAABDBCD.$$

The following is obvious and was first stated by Brunner [4].

Observation 6 ([4])**.** *An instance of the binary paint shop problem is red-first-perfect if and only if it does not contain a subword of the form* $AABB$.

An interesting open question is the following.

Problem 7. *Characterise the recursive-greedy-perfect instances of the binary paint shop problem by forbidden subwords.*

Using computer search, Brunner [4] found all forbidden subwords of length at most 14 for recursive-greedy-perfectness. These are the following.

$$\mathbf{CAADBBCD} \qquad \mathbf{AD}B\mathbf{AC}B\mathbf{CD}$$
$$\mathbf{D}B\mathbf{AACE}B\mathbf{CDE} \qquad B\mathbf{DAACE}B\mathbf{CDE} \qquad ABA\mathbf{CE}B\mathbf{CDDE}$$
$$\mathbf{AD}B\mathbf{A}EB\mathbf{CCD}E \qquad D\mathbf{AC}E\mathbf{ABBC}DE \qquad \mathbf{AC}E\mathbf{ABB}D\mathbf{C}DE$$
$$\mathbf{CAABD}F\mathbf{BC}E\mathbf{D}EF \qquad \mathbf{CAABD}F\mathbf{BCD}EEF \qquad \mathbf{DAAC}BB\mathbf{FCD}EE\mathbf{F}$$
$$BAA\mathbf{C}FB\mathbf{E}G\mathbf{CDDE}FG \qquad \mathbf{FAA}EBB\mathbf{D}CC\mathbf{GD}E\mathbf{FG}$$

We observe that all these configurations contain a subword isomorphic to $ACABBC$, $CAADBBD$ or $CAABDBCD$ (indicated in bold). Thus, by Theorem 5, they are also forbidden for greedy-perfectness. This motivates us to formulate the following conjecture, which is weaker than an affirmative answer to Problem 4.

Conjecture 8. *If an instance of the binary paint shop problem is recursive-greedy-perfect, then it is greedy-perfect, too.*

We remark that the example from the proof of our main Proposition 2 is not a counterexample to Conjecture 8, since it contains the forbidden subword $ACABBC$.

References

[1] H. Amini, F. Meunier, H. Michel, A. Mohajeri: *Greedy colorings for the binary paintshop problem.* J. Discrete Algorithms (2010) 8–14.

[2] S. D. Andres, W. Hochstättler: *Some heuristics for the binary paint shop problem and their expected number of colour changes.* J. Discrete Algorithms 9 (2011) 203–211.

[3] P. Bonsma, T. Epping, W. Hochstättler: *Complexity results on restricted instances of a paint shop problem for words.* Discrete Appl. Math. 154 (2006) 1335–1343.

[4] C. Brunner: *Das Teilwortproblem für einige Heuristiken für das Binary-Paint-Shop-Problem* (in German). Bachelor's thesis, FernUniversität in Hagen (2016).

[5] T. Epping, W. Hochstättler, P. Oertel: *Complexity results on a paint shop problem.* Discrete Appl. Math. 136 (2004) 217–226.

[6] A. Gupta, S. Kale, V. Nagarajan, R. Saket, B. Schieber: *The approximability of the binary paintshop problem.* In P. Raghavendra et al. (Eds): APPROX/RANDOM 2013, LNCS 8096, pp 205–217, 2013.

[7] S. Khot: *On the power of unique 2-prover 1-round games.* Proc. STOC'02, May 19-21, 2002, Montreal Quebec, Canada (2002) 767–775.

[8] F. Meunier, B. Neveu: *Computing solutions of the paintshop-necklace problem.* Comp. Oper. Res. 39 (2012) 2666-2678.

[9] F. Meunier, A. Sebő: *Paintshop, odd cycles and necklace splitting.* Discrete Appl. Math. 157 (2009) 780–793.

[10] D. Rautenbach, Z. Szigeti: *Greedy colorings of words.* Discrete Math. 160 (2012) 1872–1874.

[11] R. Šámal, J. Hančl, A. Kabela, M. Opler, J. Sosnovec, P. Valtr: *The binary paint shop problem.* Slides of the talk of Robert Šámal at the Midsummer Combinatorial Workshop XXIV in Prague (July 30, 2019).