The binary paint shop problem

Robert Šámal (Charles University)

Joint work with J. Hančl, A. Kabela, M. Opler, J. Sosnovec, P. Valtr

MCW, Prague Jul 30, 2019

CSI

IÚUK

Outline

Introduction

Our results

The problem

- **double occurrence word** – every letter occurs twice

$$w = ADEBAFCBCDEF$$

- **want:** color all letters red&blue, every letter once red and once blue

$ADEBAFCBCDEF$ 4 changes

- **goal:** minimize the number of **color changes**

$ADEBAFCBCDEF$ 4 changes

$ADEBAFCBCDEF$ 2 changes

$$\gamma(w) = 2$$

Trivial observations

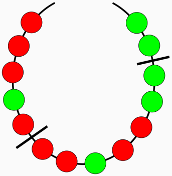
- $w_1 = A_1 A_1 A_2 A_2 \dots A_n A_n \qquad \gamma(w_1) = n$
- $w_2 = A_1 A_2 \dots A_n A_1 A_2 \dots A_n \qquad \gamma(w_2) = 1$
- W_n – set of words with letters $A_1, \dots, A_n$, each of them twice.

Natural questions

- value for nontrivial cases?
- algorithms?
- random $w \in W_n$?
- connection to some other parameters?
- motivation?

Motivation and previous results

- **paint shop**: a factory where a sequence of cars needs to be painted, for each sub-type we want one of each color, it is practical not to change the color too often.
- **necklace splitting**: [Image by Wikipedia user Kilom691, CC BY-SA 4.0]



Two (possibly more) thieves want to split a necklace with various types of gem-stones, using minimum number of cuts. N.Alon's theorem is more general, here it gives just $\gamma(w) \leq n$ for $w \in W_n$.

Hard problem

- APX-hard [Bonsma, Epping, Hochstättler (06); Meunier, Sebő (09)]
- Thus, the decision problem is NP-complete.
- some polynomial instances identified by Meunier and Sebő (09)

Heuristics

Results by Andres&Hochstättler, 2010.

- **greedy** – $g(w)$ – going from left to right, change color only if you must.

$$\mathbb{E}_{w \in W_n} g(w) = \mathbb{E}_n g(w) = 0.5n + o(n)$$

- **recursive greedy** – $rg(w)$ – remove the last letter, color recursively, choose the better way for the extra letter

$$\mathbb{E}_n rg(w) = 0.4n + o(n)$$

Outline

Introduction

Our results

Lower bounds

Observation

$$\gamma(w) \geq \alpha(G(w))$$

where $G(w)$ is the interval graph corresponding to the word w.

Scheinerman (1988) proved that for a random interval graph on n vertices, $\alpha \geq C\sqrt{n}$. Thus:

Corollary

$$\mathbb{E}_n\gamma \geq C\sqrt{n}$$

Linear lower bound

Theorem

$$\mathbb{E}_n\gamma \geq 0.214n - o(n)$$

This disproves a conjecture by Meunier, Neveu (2012). The conjecture was also mentioned at MCW 2012 (Andres) and MCW 2017 (Hochstättler).

Lower bound proof

- $w \in W_n$ – a random element
- will show $\Pr[\gamma(w) \le k] \le p$.
- This will prove that $\mathbb{E}_n \gamma \ge (1-p)k$.
- $C_n^{\le k}$ – colorings of $1, \ldots, 2n$ using n red and n blue, with at most k color changes.

$$\begin{aligned}
\Pr[\gamma(w) \le k] &= \Pr[w \text{ has a legal coloring in } C_n^{\le k}] \\
&\le \sum_{C \in C_n^{\le k}} \Pr[C \text{ is legal for } w] \\
&= \sum_{C \in C_n^{\le k}} \frac{n!^2}{(2n)!/2^n} \\
&= \cdots \le \frac{\sqrt{4n}}{2^n} \left(\frac{e \cdot 2n}{k}\right)^k
\end{aligned}$$

$p :=$ the latter, $k := 0.214n$ ... done.

Concentration

Theorem

Let w be a random element of W_n. Let $\gamma_n = \mathbb{E}_n \gamma$.

$$\Pr\left[|\gamma(w) - \gamma_n| \geq \sqrt{n \log n}\right] \leq 2n^{-1/8}$$

Concentration

Theorem

Let w be a random element of W_n. Let $\gamma_n = \mathbb{E}_n \gamma$.

$$\Pr\left[|\gamma(w) - \gamma_n| \geq \sqrt{n \log n}\right] \leq 2n^{-1/8}$$

Proof.

- Standard application of Azuma inequality.
- We let X_k be the expectation of $\gamma(w)$ after the positions of the letters $A_1, \ldots, A_k$ have been fixed.
- $X_0, X_1, \ldots, X_n$ is a martingale.
- $|X_k - X_{k+1}| \leq 2$.
- Azuma inequality gives the rest.

□

Improved upper bounds – theorem

Theorem

$$\gamma_n \leq (\frac{2}{5} - \varepsilon)n$$

for $\varepsilon \approx 1.64 \times 10^{-6}$.

Proof.

We run the recursive greedy algorithm, then observe that there is a linear number of local changes. □

Improved upper bounds – star heuristic

We propose a new heuristics – star heuristics. According to numerical evidence and rather convincing arguments, we believe that

$$\mathbb{E}_n s \leq 0.361n$$

Improved upper bounds – star heuristic

We propose a new heuristics – star heuristics. According to numerical evidence and rather convincing arguments, we believe that

$$\mathbb{E}_n s \leq 0.361n$$

1. Similarly as in the recursive greedy, we take away the last letter and its second copy, we repeat.
2. We let the resulting words be $w_n = w, w_{n-1}, \ldots, w_1 = AA$.
3. Then we go forward, producing the coloring using red, blue, and * with the following condition:
4. The two copies of a letter must either be red/blue, blue/red or */*. We use the latter, if both red/blue and blue/red yield the same number of color changes.

Improved upper bounds – star heuristic

1. Similarly as in the recursive greedy, we take away the last letter and its second copy, we repeat.
2. We let the resulting words be $w_n = w, w_{n-1}, \ldots, w_1 = AA$.
3. Then we go forward, producing the coloring using red, blue, and * with the following condition:
4. The two copies of a letter must either be red/blue, blue/red or */*. We use the latter, if both red/blue and blue/red yield the same number of color changes.
5. To get the coloring of w_{k+1} from that of w_k
   - do the greedy consideration of the new letter (possibly deciding about some *-colored letters).
   - possibly recolor the penultimate letter (and its copy) by a *.

Better bounds – open problem

Based on experiments (using a heuristics impossible to analyze), we believe the true value of γ_n is around $0.3n$. However, we have only the following bounds proved rigorously

$$0.214 \leq \lim \frac{\gamma_n}{n} \leq 0.4 - \varepsilon$$

We can imagine the upper bound can be decreased to around 0.361 with more work.

Question

What is $\lim \frac{\gamma_n}{n}$*? Does the limit even exist?*