Repetitions in Strings: A "Constant" Problem

Hideo Bannai
M&D Data Science Center, Tokyo Medical and Dental University, Japan

Abstract

Repeating structures in strings is one of the most fundamental characteristics of strings, and has been an important topic in the field of combinatorics on words and combinatorial pattern matching since their beginnings. In this talk, I will focus on squares and maximal repetitions and review the "runs" theorem [1] as well as related results (e.g. [5, 6, 7, 3, 2, 4]) which address the two main questions: how many of them can be contained in a string of given length, and algorithms for computing them.

2012 ACM Subject Classification Mathematics of computing → Combinatorics on words

Keywords and phrases Maximal repetitions, Squares, Lyndon words

Digital Object Identifier 10.4230/LIPIcs.CPM.2021.1

Category Invited Talk

Funding *Hideo Bannai*: Supported by JSPS KAKENHI Grant Number JP20H04141.

References

1 Hideo Bannai, Tomohiro I, Shunsuke Inenaga, Yuto Nakashima, Masayuki Takeda, and Kazuya Tsuruta. The "runs" theorem. *SIAM J. Comput.*, 46(5):1501–1514, 2017. doi:10.1137/15M1011032.

2 Hideo Bannai, Takuya Mieno, and Yuto Nakashima. Lyndon words, the three squares lemma, and primitive squares. In Christina Boucher and Sharma V. Thankachan, editors, *String Processing and Information Retrieval - 27th International Symposium, SPIRE 2020, Orlando, FL, USA, October 13-15, 2020, Proceedings*, volume 12303 of *Lecture Notes in Computer Science*, pages 265–273. Springer, 2020. doi:10.1007/978-3-030-59212-7_19.

3 Philip Bille, Jonas Ellert, Johannes Fischer, Inge Li Gørtz, Florian Kurpicz, J. Ian Munro, and Eva Rotenberg. Space efficient construction of Lyndon arrays in linear time. In Artur Czumaj, Anuj Dawar, and Emanuela Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, July 8-11, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 168 of *LIPIcs*, pages 14:1–14:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.ICALP.2020.14.

4 Jonas Ellert and Johannes Fischer. Linear time runs over general ordered alphabets. *CoRR*, abs/2102.08670, 2021. arXiv:2102.08670.

5 Aviezri S. Fraenkel and Jamie Simpson. How many squares can a string contain? *J. Comb. Theory, Ser. A*, 82(1):112–120, 1998. doi:10.1006/jcta.1997.2843.

6 Yuta Fujishige, Yuto Nakashima, Shunsuke Inenaga, Hideo Bannai, and Masayuki Takeda. Almost linear time computation of maximal repetitions in run length encoded strings. In Yoshio Okamoto and Takeshi Tokuyama, editors, *28th International Symposium on Algorithms and Computation, ISAAC 2017, December 9-12, 2017, Phuket, Thailand*, volume 92 of *LIPIcs*, pages 33:1–33:12. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.ISAAC.2017.33.

7 Ryo Sugahara, Yuto Nakashima, Shunsuke Inenaga, Hideo Bannai, and Masayuki Takeda. Computing runs on a trie. In Nadia Pisanti and Solon P. Pissis, editors, *30th Annual Symposium on Combinatorial Pattern Matching, CPM 2019, June 18-20, 2019, Pisa, Italy*, volume 128 of *LIPIcs*, pages 23:1–23:11. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.CPM.2019.23.