

PATTERN AVOIDANCE IN PARTIAL WORDS DENSE WITH HOLES

KEVIN BLACK^(A) FRANCINE BLANCHET-SADRI^(B) IAN COLEY^(C)
BRENT WOODHOUSE^(C) ANDREW ZEMKE^(D)

^(A)*Department of Mathematics, Harvey Mudd College
301 Platt Boulevard, Claremont, CA 91711, USA*

^(B)*Department of Computer Science, University of North Carolina
P.O. Box 26170, Greensboro, NC 27402-6170, USA
blanchet@uncg.edu*

^(C)*Department of Mathematics, University of California at Los Angeles
Box 951555, Los Angeles, CA 90095-1555, USA*

^(D)*Department of Mathematics, Cornell University
310 Malott Hall, Ithaca, NY 14853-4201, USA*

ABSTRACT

A *partial word* is a sequence of symbols over a finite alphabet that may have some undefined positions, called *holes*, that match every letter of the alphabet. Previous work completed the classification of all binary patterns with respect to partial word avoidability. In this paper, we pose the problem of avoiding patterns in partial words very dense with holes. We define the concept of hole sparsity, a measure of the frequency of holes in a partial word, and determine the minimum hole sparsity for all unary patterns in the context of trivial and non-trivial avoidability. Results for more general patterns are also given. Furthermore, we discuss hole spacing and hole density for abelian powers.

Keywords: combinatorics on words, partial words, patterns, avoidable patterns, abelian powers, minimum hole sparsity, hole spacing, hole density

1. Introduction

Notions and techniques related to patterns such as repetitions in strings find applications in several areas of theoretical and applied computer science, notably in text processing, data compression, computational biology, string and pattern matching algorithms (see [12] for an overview on repetitions in strings). In pattern matching, several algorithms take advantage of the repetitions of the pattern to speed up the search of its occurrences in a text. On the other hand, non-repetitive sequences, those avoiding patterns such as squares, or square-free words, have been used to build several counterexamples in context-free languages, groups, lattice of varieties, partially