

RECURRENT PARTIAL WORDS AND REPRESENTABLE SETS

FRANCINE BLANCHET-SADRI^(A) ALEKSANDAR CHAKAROV^(B) BOB CHEN^(C)
LUCAS MANUELLI^(D) SINZIANA MUNTEANU^(E) JARETT SCHWARTZ^(F)
SLATER STICH^(G)

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

^(B)*Department of Computer Science, University of Colorado at Boulder
430 UCB, Boulder, Colorado 80309-0430, USA*

^(C)*Department of Mathematics, University of California San Diego
9500 Gilman Drive #0112, La Jolla, CA, 92093-0112, USA*

^(D)*Department of Economics, Massachusetts Institute of Technology
50 Memorial Drive, E52-391, Cambridge, MA 02142, USA*

^(E)*School of Computer Science, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213-3891, USA*

^(F)*Electrical Engineering and Computer Sciences Department
Computer Science Division, University of California Berkeley
387 Soda Hall, Berkeley, CA 94720-1776, USA*

^(G)*Department of Mathematics, University of California Berkeley
970 Evans Hall, Berkeley, CA 94720, USA*

ABSTRACT

Partial words are sequences over a finite alphabet that may contain wildcard symbols, called holes, which match or are compatible with all letters; partial words without holes are said to be total words (or simply words). Given an infinite partial word w , the number of distinct total words over the alphabet that are compatible with factors of w of a given length, called subwords of w , refers to a measure of complexity of infinite partial words called subword complexity. This measure is of particular interest because partial words can be constructed with subword complexities not achievable by total words. In this paper, we consider the notion of recurrence over infinite partial words, i. e., we study whether all of the finite subwords of a given infinite partial word appear infinitely often, and we establish connections between subword complexity and recurrence in this more general framework. We consider the problem of representing languages by infinite partial words, i. e., we classify the languages which are the sets of subwords of infinite partial words. We show how to construct such representing partial words and give criteria for when they are unique. Non-recurrence turns out to be a necessary condition for a set to be uniquely representable, while non-ultimate recurrence is sufficient but not necessary. We also answer the question “Which representable sets are irreducible?”

Keywords: formal languages, combinatorics on words, partial words, subword complexity, recurrence, representable sets