Available online at www.sciencedirect.com

ScienceDirect

Electronic Notes in DISCRETE MATHEMATICS

Electronic Notes in Discrete Mathematics 61 (2017) 231–237 www.elsevier.com/locate/endm

On universal partial words

Herman Z. Q. Chen

School of Science, Tianjin Chengjian University, P.R. China

Sergey Kitaev

Department of Computer and Information Sciences, University of Strathclyde, Glasgow, UK

Torsten Mütze [1]

Institut für Mathematik, TU Berlin, Germany

Brian Y. Sun

College of Mathematics and System Science, Xinjiang University, Urumqi, P.R. China

Abstract

A *universal word* for a finite alphabet A and some integer $n \geq 1$ is a word over A such that every word of length n appears exactly once as a (consecutive) subword. It is well-known and easy to prove that universal words exist for any A and n. In this work we initiate the systematic study of universal *partial* words. These are words that in addition to the letters from A may contain an arbitrary number of occurrences of a special 'joker' symbol $\diamond \notin A$, which can be substituted by any symbol from A. For example, $u = 0\diamond 011100$ is a universal partial word for the binary alphabet $A = \{0, 1\}$ and for $n = 3$ (e.g., the first three letters of u yield the subwords 000 and 010). We present results on the existence and non-existence of universal partial words in different situations (depending on the number of $\diamond$s and

http://dx.doi.org/10.1016/j.endm.2017.06.043

their positions), including various explicit constructions. We also provide numerous examples of universal partial words that we found with the help of a computer.

Keywords: universal word, partial word, De Bruijn graph, Eulerian cycle, Hamiltonian cycle

---

1 Introduction

De Bruijn sequences are a centuries-old and well-studied topic in combinatorics, and over the years they found widespread use in real-world applications, e.g., in the areas of molecular biology, computer security, computer vision, robotics and psychology experiments (detailed references are given in [7]). More recently, they have also been studied in a more general context by constructing so-called *universal cycles* for other fundamental combinatorial structures such as permutations or subsets of a fixed ground set (see e.g. [8,14,16]).

In the context of words over a finite alphabet A, we say that a word u is *universal for* A^n if u contains every word of length $n \geq 1$ over A exactly once as a (consecutive) subword. For example, for the binary alphabet $A = \{0, 1\}$ and for $n = 3$, $u = 0001011100$ is a universal word for A^3. Note that reversing a universal word, or permuting the letters of the alphabet yields a universal word again. The following classical result is the starting point for our work (see [10,15,17]).

Theorem 1.1 *For any finite alphabet A and any $n \geq 1$, there exists a universal word for A^n.*

The standard proof of Theorem 1.1 is really beautiful and concise, using the De Bruijn graph, its line graph and Eulerian cycles (see [8]).

1.1 Universal partial words

In this paper we consider universality of so-called *partial words*, words that in addition to letters from A may contain any number of occurrences of an additional special symbol $\diamond \notin A$. The idea is that every occurrence of $\diamond$ can be substituted by any symbol from A, so we can think of $\diamond$ as a 'joker' or 'wildcard' symbol. Formally, we define $A_\diamond := A \cup \{\diamond\}$ and we say that a word $v = v_1 v_2 \cdots v_n \in A^n$ appears as a *factor* in a word $u = u_1 u_2 \cdots u_m \in A_\diamond^m$ if

---

[1] Email: muetze@math.tu-berlin.de