Palindromes in mathematics

Wojciech Plandowski
Department of Mathematics
Warsaw University, Poland

Keywords: Palindromes

1.Introduction

Everyone knows what a palindrome is. It is a text which is the same when reading forward and backward. One of the examples is the palindrome: "C is a basic". This palindrome can be extended to a palindrome of unbounded length: "C is a basic, is a basic, is a basic, ...".

Palindromes in mathematics have a little bit more precise definition. Namely, each letter including space is important. Lower-case and capital letters are considered as different letters. Using this definition "C is a basic" is not a palindrome because after the capital C is space while before the last c is not a space. Moreover, the first C is a capital letter while the last c is a lower-case letter. A palindrome in a mathematical sense is "cisabasic".

Palindromes in mathematical sense do not have to have a meaning, for instance the sequence of letters "abbabba" is a good palindrome. Palindromes are considered mainly in two branches of mathematics: combinatorics on words and text algorithms. We will study one example of occurrence of palindromes in each of these two branches.

2.Palindromes in Combinatorics on Words

Combinatorics on words studies properties of words including their structure and regularities which occur inside them. Palindrome is some kind of regularity. Palindromes in Combinatorics on Words occur in several theorems [2]. We will explain they relation to one of the most basic and important theorems in this branch of mathematics – Fine and Wilf's Theorem. To explain this theorem we have to introduce a few definitions.

A letter is either 'a' or 'b'. *A word* is any sequence of letters, for instance "abbabbabbabba". The number of letters in a word w is called *the length* of w and is denoted by |w|. The i-th letter of a word w is denoted by w[i]. We have w=w[1]w[2]...w[|w|]. *A period* of a word w is a number p such that w[i]=w[i+p], for 1≤i≤|w|-p. If p is a period of w, then w is of the form uuu..uu' where u=w[1]...w[p] and u'=w[1]...w[k], for some k≤p. It is not difficult to see that if p is a period of w, then 2p, 3p and all other multiples of p are periods of w.

Denote by gcd(p,q) the greatest number which divides both p and q. Then the Fine and Wilf Theorem can be formulated in the following way.

Fine and Wilf's Theorem. Let p and q are two periods of a word w. If p+q-gcd(p,q)≤|w|, then gcd(p,q) is also a period of w.

Observe here that if gcd(p,q) is period of w, then p and q, as multiples of gcd(p,q), are also periods of w. For gcd(p,q)=1, the Fine and Wilf Theorem can be formulated in the following way.

Fine and Wilf's Theorem (case gcd(p,q)=1). Let p and q are two periods of a word w. Let gcd(p,q)=1. If p+q-1≤|w|, then w consists only of letters a or only of letters b.

An interesting question is: what happens if p and q are periods of w, gcd(p,q)=1 and |w|=p+q-2? Does w has to consists only of letters a or only of letters b? The answer is: No. For each p, q such that gcd(p,q)=1 there is a word w of length p+q-2 which contains letters a and b and such that p and q are periods of w. This means that there is a sense of introducing the following definition. A word w is *central* if there are numbers p and q such that |w|=p+q-2 and p and q are periods of w. Such words are characterized by the following theorem.

Theorem A word w is central if and only if either consists only of letters a, or only of letters b, or w is a palindrome and w is of the form uabv where u and v are palindromes.

3.Palindromes in Text Algorithms

Text algorithms are a part of algorithmics which works on words called here texts. Palindromes in Text Algorithms has been studied in a few contexts of sequential algorithms [2] and parallel algorithms [1]. We will study one algorithmic problem which leads to a very useful data structure for sequential algorithms connected to palindromes. We start by introducing a useful notation. By w[i..j] we mean the word w[i]w[i+1]...w[j]. The problem we are interested in is a basic one.

Problem: Given a word w. Design a data structure which allows, for a given two positions i and j of the word w, check whether w[i..j] is a palindrome.

An obvious solution of our problem is to build a two-dimensional boolean array p[i,j], for $1 \le i \le j \le |w|$ such that p[i,j]=true if and only if w[i..j] is a palindrome. Using the array p we can answer the question whether w[i..j] is a palindrome in constant time. However, this solution has one drawback: the size of a data structure p is quadratic with respect to the length of w so any algorithm which computes this data structure has to work in at least quadratic time.

A solution proposed by Manacher is more tricky. From this moment we assume that we are interested only in even palindromes, that is j-i+1 in our problem is even. Extension to all palindromes is not difficult task and is left to the reader. We compute an array of integers R such that R[i] is the maximal number j such that w[i-j..i+j-1] is a palindrome. The size of the array is linear with respect to |w|.

R[i] is just a radius of a maximal even palindrome which is centered between positions i-1 and i in w. If we want to check whether w[i..j] is an even palindrome we just check whether $R[(j+i+1)/2] \ge (j-i+1)/2$, that is whether the radius of the maximal palindrome centered in the center of w[i..j] is greater than half of the length of w[i..j].

The computation of R in quadratic time is a very simple task. We leave it to the reader. Linear time algorithm is not simple and it requires some reasoning from combinatorics on words. The reader which is interested in it is encouraged to read [2].

Theorem The array R can be computed in linear time.

1. A. Apostolico, D. Breslauer, Z. Galil, "Parallel detection of all palindromes in a string".
2. L. Banachowski, A. Kreczmar, W. Rytter, Sprwdzanie własności syntaktycznych tekstów związanych z palindromami", Chapter 3.3 in "Analiza Algorytmów i Struktur Danych", WNT 1989 (in Polish), p. 176-188.
3. M. Lothaire, "Sturmian words", Chapter 2 in "Algebraic Combiatorics on Words", Cambridge University Press 2002. p.45-110.