Morphisms on Amiable Words

Adrian Atanasiu*

Faculty of Mathematics and Computer Science, Bucharest University, Str. Academiei 14, Bucharest 010014, Romania

Abstract. Using the fact that the Parikh matrix mapping is not an injective mapping, the paper investigates some properties of the set of words having the same Parikh matrix; these words are called "amiable" or "M - equivalent". The aim is to reduce the number of amiable words using a morphism which provides additional information about them.

1 Introduction

The idea of identifying binary sequences using as input data the number of a's and b's is quite old. Unfortunately this information (given by the Parikh mapping associated to the sequence) is insufficient: there are $\binom{x+y}{x}$ binary sequences α with $|\alpha|_a = x$ and $|\alpha|_b = y$ (we denote by $|\alpha|_w$ the number of appearances of the scattered sequence w in α). Once the Parikh matrix mapping ([6]) has been defined and especially when the Parikh matrix mapping associated with the binary sequences ([2]) has been studied (where $|\alpha|_{ab}$ is also taken into consideration), the number of sequences defined by the same characteristics has been drastically decreased. For instance, from the 184756 binary sequences α having $|\alpha|_a = 10$ and $|\alpha|_b = 10$, only 5448 do have $|\alpha|_{ab} = 50$; moreover, this is the most unfortunate choice (for other situations see some Examples from [2] and [3]). This represents almost 3% from all 184756 posible strings. Because this number still remains quite large, the possibility of identifying the sequences by using this procedure is reduced (especially for the balanced sequences, when $|\alpha|_a$ is almost equal to $|\alpha|_b$).
A remarkable improvement seems to be the use of some morphisms which distinguish the amiable binary words by their Parikh matrices. To this aim, the Section 3 of the paper analyses the morphisms $\phi : \Sigma_1^* \longrightarrow \Sigma_2^*$ where Σ_1 is a binary alphabet, and Σ_2 has at most 3 elements. Such a morphism (Istrail morphism) was studied in [3]. As a result, these morphisms distinguish the sequences in many classes of amiable words, but not totally. There are some words – like *abbabaab* and *baababba* – that remain amiable over any morphism with $|\Sigma_2| \leq 3$.

The last section proposes the construction of a morphism over a general alphabet, that is able to separate two arbitrary amiable words. The only problem which remains to be solved is the size of the alphabet Σ_2, which increases with the length of the analyzed words.

* e-mail *aadrian@gmail.com*