Available online at www.sciencedirect.com
ScienceDirect
Discrete Applied Mathematics 155 (2007) 2061–2068
DISCRETE APPLIED MATHEMATICS
www.elsevier.com/locate/dam

$\mathcal{NP}$-completeness of generalized multi-Skolem sequences

Gustav Nordh[1]

Department of Computer and Information Science, Linköpings Universitet, S-581 83 Linköping, Sweden

Received 30 April 2003; received in revised form 7 December 2005; accepted 14 May 2007
Available online 24 May 2007

Abstract

A Skolem sequence is a sequence $a_1, a_2, \ldots, a_{2n}$ (where $a_i \in A = \{1, \ldots, n\}$), each a_i occurs exactly twice in the sequence and the two occurrences are exactly a_i positions apart. A set A that can be used to construct Skolem sequences is called a Skolem set. The existence question of deciding which sets of the form $A = \{1, \ldots, n\}$ are Skolem sets was solved by Skolem [On certain distributions of integers in pairs with given differences, Math. Scand. 5 (1957) 57–68] in 1957. Many generalizations of Skolem sequences have been studied. In this paper we prove that the existence question for generalized multi-Skolem sequences is $\mathcal{NP}$-complete. This can be seen as an upper bound on how far the generalizations of Skolem sequences can be taken while still hoping to resolve the existence question.

Keywords: Skolem sequence; Design theory; $\mathcal{NP}$-completeness

1. Introduction

Skolem sequences were introduced by Skolem [6] in 1957, for the construction of Steiner triple systems. He considered sets of the form $A = \{1, 2, \ldots, n\}$ and asked whether one can always form a sequence with two copies of every element k in the set so that the two copies of k are placed k places apart in the sequence. Such sequences are called *Skolem sequences*. For example, the set $\{1, 2, 3, 4\}$ can be used to construct the sequence 42324311, but the set $\{1, 2, 3\}$ cannot be used to form such a sequence. A set that can be used to construct a Skolem sequence is called a *Skolem set*.

Many different aspects and generalizations of Skolem sequences have been studied. One reason for them being so well studied is that they have important applications in several branches of mathematics; Shalaby [5] describes applications in design theory and graph labelings.

Baker [1] introduced generalized Skolem sequences and used them to construct k-extended Skolem sequences. They have also been used in the construction of extended Langford sequences with small defects [3]. A *generalized Skolem sequence* is a sequence of positive integers and null symbols such that an integer appears exactly twice or not at all, and the two appearances of an integer j are j positions apart. If the integers in A can be used to construct a generalized Skolem sequence using only the positions in P, we say that (P, A) is a *generalized Skolem pair*. For example, $(\{1, 2, 4, 5, 7, 8\}, \{1, 5, 7\})$ is a generalized Skolem pair. The corresponding generalized Skolem sequence is 75011057 (0 occupies positions that are not in P). Note that a pair (P, A) is a generalized Skolem pair if and only

[1] Supported by the *National Graduate School in Computer Science* (CUGS), Sweden.
E-mail address: gusno@ida.liu.se.

0166-218X/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2007.05.007

if the positions in P can be partitioned into the differences in A, e.g., the example $(\{1, 2, 4, 5, 7, 8\}, \{1, 5, 7\})$ above is a generalized Skolem pair since $\{5-4, 7-2, 8-1\} = \{1, 5, 7\}$. Hence, we will refer to the elements in A as the differences in A.

We generalize the notion of generalized Skolem sequences slightly and allow the set of differences A to be a multiset. We call these sequences *generalized multi-Skolem sequences*, and the corresponding pair (P, A) *generalized multi-Skolem pair*. For example, $(\{1, 2, 4, 5, 7, 8\}, \{1, 6, 6\})$ is a generalized multi-Skolem pair. The corresponding generalized multi-Skolem sequence is 66011066 (0, occupies positions that are not in P). Linek and Shalaby [4] give some necessary conditions for the existence of generalized multi-Skolem pairs. Moreover, they state that a basic question is to decide which pairs (P, A) are generalized multi-Skolem pairs. We prove that the problem of deciding which pairs (P, A) are generalized multi-Skolem pairs is $\mathcal{NP}$-complete. The proof is a reduction from the $\mathcal{NP}$-complete problem Multiple Choice Matching [2]. We refer the reader to Garey and Johnson [2] for an in-depth treatment of the theory of $\mathcal{NP}$-completeness.

2. $\mathcal{NP}$-completeness of generalized multi-Skolem sequences

We prove that Generalized Multi-Skolem Sequences is $\mathcal{NP}$-complete by giving a reduction from the $\mathcal{NP}$-complete problem Multiple Choice Matching. Before presenting the reduction we define the two problems formally and prove some additional properties of Multiple Choice Matching that we will use in the reduction.

2.1. Generalized multi-Skolem sequences

Instance: A multiset A of positive integers, $|A| = m$, a set P of positive integers, $|P| = 2m$.

Question: Is (P, A) a generalized multi-Skolem pair? That is, can the positions in P be partitioned into the differences in A?

2.2. Multiple choice matching

Instance: A graph $G = (V, E)$, a partition of the edges E into disjoint sets $E_1, E_2, \ldots, E_m$, and a positive integer K.

Question: Is there a subset $M \subseteq E$ with $|M| \geqslant K$ such that no two edges in M share a common vertex and such that M contains at most one edge from each E_i, $1 \leqslant i \leqslant m$?

The definition of Multiple Choice Matching is taken from Garey and Johnson [2], where it is also stated that the problem remains $\mathcal{NP}$-complete even if each E_i contains at most two edges, and $K = |V|/2$. We will make use of these properties in the reduction. A set M, as defined above is called a multiple choice matching. We think of the edges in E_i as being labeled with the label i. Another property of Multiple Choice Matching that we need is $\mathcal{NP}$-completeness even in the restricted case where none of the edges with the same edge label share a common vertex. This is obviously true since two edges that share a common vertex can never be part of the same matching. Thus, we can simply assign one of these edges a new edge label not previously used in G, see Fig. 1. In the resulting graph none of the edges with the same edge label share a common vertex, and it is easy to see that the resulting graph has a Multiple Choice Matching if and only if the original graph has one.

To simplify the reduction we prove one final property of Multiple Choice Matching. Multiple Choice Matching is $\mathcal{NP}$-complete even if the number of different edge labels in the graph is greater than half the number of vertices in the graph. If m equals the number of different edge labels in the graph and the number of vertices is $2n$ (assuming that

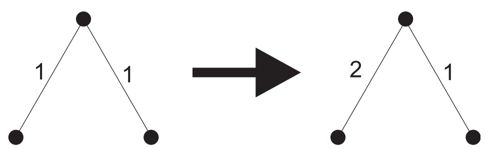

Fig. 1. The label 2 is a new one not previously used in the graph.