Doctopic: Algorithms, automata, complexity and games ARTICLE IN PRESS TCS:11406

Theoretical Computer Science ••• (••••) •••–•••

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Trimming and gluing Gray codes ☆

Petr Gregor [a,*], Torsten Mütze [b]

[a] *Department of Theoretical Computer Science and Mathematical Logic, Charles University, 11800 Prague 1, Czech Republic*
[b] *Institut für Mathematik, TU Berlin, 10623 Berlin, Germany*

ARTICLE INFO

Article history:
Received 6 April 2017
Received in revised form 16 October 2017
Accepted 5 December 2017
Available online xxxx
Communicated by M.J. Golin

Keywords:
Gray code
Subset
Combination
Loopless algorithm
Hypercube

ABSTRACT

We consider the algorithmic problem of generating each subset of $[n] := \{1, 2, \ldots, n\}$ whose size is in some interval $[k, l]$, $0 \le k \le l \le n$, exactly once (cyclically) by repeatedly adding or removing a single element, or by exchanging a single element. For $k = 0$ and $l = n$ this is the classical problem of generating all 2^n subsets of $[n]$ by element additions/removals, and for $k = l$ this is the classical problem of generating all $\binom{n}{k}$ subsets of $[n]$ by element exchanges. We prove the existence of such cyclic minimum-change enumerations for a large range of values n, k, and l, improving upon and generalizing several previous results. For all these existential results we provide optimal algorithms to compute the corresponding Gray codes in constant $\mathcal{O}(1)$ time per generated set and $\mathcal{O}(n)$ space. Rephrased in terms of graph theory, our results establish the existence of (almost) Hamilton cycles in the subgraph of the n-dimensional cube Q_n induced by all levels $[k, l]$. We reduce all remaining open cases to a generalized version of the middle levels conjecture, which asserts that the subgraph of Q_{2k+1} induced by all levels $[k-c, k+1+c]$, $c \in \{0, 1, \ldots, k\}$, has a Hamilton cycle. We also prove an approximate version of this generalized conjecture, showing that this graph has a cycle that visits a $(1-o(1))$-fraction of all vertices.

1. Introduction

Generating all objects in a combinatorial class such as permutations, subsets, combinations, partitions, trees, strings etc. is one of the oldest and most fundamental algorithmic problems, and such generation algorithms appear as core building blocks in a wide range of practical applications, see the survey [30]. In fact, half of the most recent volume [19] of Donald Knuth's seminal series *The Art of Computer Programming* is devoted entirely to this fundamental subject. The ultimate goal for algorithms that efficiently generate each object of a particular combinatorial class exactly once is to generate each new object in constant time. Such optimal algorithms are sometimes called *loopless algorithms*, a term coined by Ehrlich in his influential paper [8]. Note that a constant-time algorithm requires in particular that consecutively generated objects differ only in a constant amount, e.g., in a single transposition of a permutation, in adding or removing a single element from a set, or in a single tree rotation operation. These types of orderings have become known as *combinatorial Gray codes*. Here are two fundamental examples for this kind of generation problems: (1) The so-called *reflected Gray code* is a method to generate all 2^n many subsets of $[n] := \{1, 2, \ldots, n\}$ by repeatedly adding or removing a single element. It is named after

☆ An extended abstract of this paper has appeared in the Proceedings of the 34th International Symposium on Theoretical Aspects of Computer Science (STACS 2017) [12].

* Corresponding author.
E-mail addresses: gregor@ktiml.mff.cuni.cz (P. Gregor), muetze@math.tu-berlin.de (T. Mütze).

https://doi.org/10.1016/j.tcs.2017.12.003
0304-3975/© 2017 Published by Elsevier B.V.

Please cite this article in press as: P. Gregor, T. Mütze, Trimming and gluing Gray codes, Theoret. Comput. Sci. (2018), https://doi.org/10.1016/j.tcs.2017.12.003

Frank Gray, a physicist and researcher at Bell Labs, and appears in his patent [13]. The reflected Gray code has many interesting properties, see [19, Section 7.2.1.1], and there is a simple loopless algorithm to compute it [8,1]. (2) Of similar importance in practice is the problem of generating all $\binom{n}{k}$ many k-element subsets of $[n]$ by repeatedly exchanging a single element. Also for this problem, loopless algorithms are well-known [34,8,1,10,9,27,3,17] (see also [19, Section 7.2.1.3]).

In this work we consider far-ranging generalizations of the classical problems (1) and (2). Specifically, we consider the algorithmic problem of generating all, or almost all, subsets of $[n]$ whose size is in some interval $[k,l]$, where $0 \le k \le l \le n$, by repeatedly adding or removing a single element, or by exchanging a single element, as further detailed later. The classical problems (1) and (2) can be seen as the special cases where $k=0$ and $l=n$, or where $k=l$, respectively. The entire parameter range in between those special cases offers plenty of room for surprising discoveries and hard research problems, as Fig. 2 illustrates.

In a computer a subset of $[n]$ is conveniently represented by the corresponding characteristic bitstring x of length n, where all the 1s of x correspond to the elements contained in the set, and the 0s to the elements not contained in the set. E.g., for $n=5$ the subset $\{1,2,5\}$ corresponds to the bitstring 11001. The aforementioned subset generation problems can thus be rephrased as Hamilton cycle problems in subgraphs of the *cube* Q_n, the graph that has as vertices all bitstrings of length n, with an edge between any two vertices, i.e., bitstrings, that differ in exactly one bit. We refer to the number of 1s in a bitstring x as the *weight of* x, and we refer to the vertices of Q_n with weight k as the *k-th level* of Q_n. Note that there are $\binom{n}{k}$ vertices on level k. Moreover, we let $Q_{n,[k,l]}$, $0 \le k \le l \le n$, denote the subgraph of Q_n induced by all levels $[k,l]$. In terms of sets, the vertices of the cube Q_n correspond to subsets of $[n]$, and flipping a bit along an edge corresponds to adding or removing a single element. Continuing the previous example, moving from the vertex 11001 to 11101 corresponds to adding the element 3 to the set $\{1,2,5\}$, yielding the set $\{1,2,3,5\}$. The weight of a bitstring corresponds to the size of the set, and the vertices on level k correspond to all k-element subsets of $[n]$.

One of the hard instances of the aforementioned general enumeration problem in $Q_{n,[k,l]}$ is when $n=2k+1$ and $l=k+1$. The existence of a Hamilton cycle in the graph $Q_{2k+1,[k,k+1]}$ for any $k \ge 1$ is asserted by the well-known *middle levels conjecture*, raised independently in the 80's by Havel [15] and Buck and Wiedemann [2]. The conjecture has also been attributed to Dejter, Erdős, Trotter [20] and various others, and also appears in the popular books [35,19,4]. The middle levels conjecture has attracted considerable attention over the last 30 years [29,11,32,18,6,20,5,16,14,26,31,28], and a positive solution, i.e., an existence proof for a Hamilton cycle in $Q_{2k+1,[k,k+1]}$ for any $k \ge 1$, has been announced only recently.

Theorem 1 *([25]). For any $k \ge 1$, the graph $Q_{2k+1,[k,k+1]}$ has a Hamilton cycle.*

The following generalization of the middle levels conjecture was proposed in [14].

Conjecture 2 *([14]). For any $k \ge 1$ and $c \in \{0,1,\dots,k\}$, the graph $Q_{2k+1,[k-c,k+1+c]}$ has a Hamilton cycle.*

Conjecture 2 clearly holds for all $k \ge 1$ and $c=k$ as $Q_{2k+1,[0,2k+1]} = Q_{2k+1}$, so this is problem (1) from before. It is known that the conjecture also holds for all $k \ge 1$ and $c=k-1$ [7,21] and $c=k-2$ [14]. By Theorem 1 we know that it also holds for all $k \ge 1$ and $c=0$. As far as small cases are concerned, computer experiments show that $Q_{2k+1,[k-c,k+1+c]}$ indeed has a Hamilton cycle for all $k \le 6$ and all $c \in \{0,1,\dots,k\}$. The largest instance in this range not yet covered by the aforementioned general results is $Q_{13,[3,10]}$ with 8008 vertices.

Another generalization of Theorem 1 in a slightly different direction, which still remains a special case in our general framework, is the following result.

Theorem 3 *([24]). For any $n \ge 3$ and $k \in \{1,2,\dots,n-2\}$, the graph $Q_{n,[k,k+1]}$ has a cycle that visits all vertices in the smaller bipartite class.*

The idea for the proof of Theorem 3 based on induction over n was first presented in [15]. In that paper, the theorem was essentially proved conditional on the validity of the hardest case $n=2k+1$, the middle levels conjecture, which was established only much later, see Theorem 1. In [24], Theorem 3 was proved unconditionally, and the proof technique was refined further to also prove Hamiltonicity results for the so-called bipartite Kneser graphs, another generalization of the middle levels conjecture.

Conjecture 2 and Theorem 3 immediately suggest the following common generalization: For which intervals $[k,l]$ does the cube $Q_{n,[k,l]}$ have a Hamilton cycle? The graph $Q_{n,[k,l]}$ is bipartite with the two partition classes given by the parity of weight of the vertices, and it is clear that a Hamilton cycle can exist only if the two partition classes have the same size, which happens only for odd dimension n and between two symmetric levels k and $l=n-k$, the case covered by Conjecture 2, or for even dimension n and $[k,l]=[0,n]$. However, we may slightly relax this question, and ask for a long cycle. To this end, we denote for any bipartite graph G by $\nu(G)$ the number of vertices of G, and by $\delta(G)$ the difference between the larger and the smaller partition class. Note that in any bipartite graph G the length of any cycle is at most $\nu(G)-\delta(G)$, i.e., the length of a cycle that visits all vertices in the smaller partition class. We call such a cycle a *saturating cycle*, see Fig. 1 (b). Observe that if both partition classes have the same size, i.e., $\delta(G)=0$, then a saturating cycle is a Hamilton cycle. Hence saturating cycles naturally generalize Hamilton cycles for unbalanced bipartite graphs. The right common generalization of Conjecture 2 and Theorem 3 therefore is: