P-Complete Permutation Group Problems

Kenneth D. Blaha  Eugene M. Luks
Pacific Lutheran University  University of Oregon

Abstract

It was shown by Furst, Hopcroft, and Luks that a variant of Sims's elegant
algorithm for membership-testing in permutation groups could be implemented
in polynomial time. Because this well-known method employs a "sifting" pro-
cess which seems inherently sequential, McKenzie and Cook conjectured that
the membership problem was P-complete. Later, Babai, Luks, and Seress, re-
lying, in part, on the classification of finite simple groups, developed methods
that bypassed the sifting obstruction. However, the parallelizability of Sims's
method remained open. We now justify the earlier intuition by showing that
sifting is P-complete.

We also demonstrate the P-completeness of some other permutation group
problems. Among these is the problem of computing the proposed canonical
forms for the class of vertex-colored graphs with bounded color multiplicities.
This opens a gap, in parallel computation, between isomorphism-testing and
finding canonical forms, for the former problem is in NC for this graph class.

1 Introduction

In the late 60's Sim's introduced an efficient algorithm for membership-testing in
permutation groups [4]. It was later shown by Furst, Hopcroft, and Luks that Sims's
algorithm could be implemented in polynomial time [2]. This well-known method
employed a sifting procedure that appeared inherently sequential. Babai, Luks, and
Seress, relying, in part, on the classification of finite simple groups, developed meth-
ods that bypassed the sifting obstruction. However, the parallelizability of Sims's
algorithm remained open [1]. We prove that sifting is P-complete.

We also prove that the problem of computing the proposed canonical forms for
vertex-colored graphs with bounded color classes is P-complete. The interest in this
problem is stimulated, in part, by its relationship to graph isomorphism. If one can
find canonical forms, then one can test graph isomorphism. This result opens a gap
in parallel computation between isomorphism-testing and finding canonical forms for
vertex-colored graphs with bounded color classes.

2 Definitions and Preliminaries

We assume familiarity with the complexity classes P, NP, and NC. We refer the reader
to any standard text, e.g.,[3], for basic facts about groups. For permutation group
concepts we refer to [5]. The group of all permutations of an n-element set \( \Omega \) is denoted \( \text{Sym}(\Omega) \), and we write \( H \leq G \) if \( H \) is a subgroup of \( G \). A standard tool for permutation group computation is a strong generating set (SGS). The following definitions are due to Sims and may be found in [4].

A base for \( G \leq \text{Sym}(\Omega) \) is a sequence of points \( B = b_1, b_2, \ldots, b_k, b_i \in \Omega, \) such that the only element in \( G \) fixing all of the \( b_i \) is the identity. The tower of subgroups \( G = G^0 \geq G^1 \geq \cdots \geq G^k = \{1\} \) where \( G^i = G_{\{b_i, \ldots, b_k\}} \), \( 1 \leq i \leq k \) is the chain of stabilizers of \( G \) relative to \( B \). An SGS for \( G \) relative to \( B \) is a subset \( Z \) of \( G \) such that \( G^i \) is generated by \( Z \cap G^i \), \( 0 \leq i \leq k - 1 \). Unless otherwise stated we shall assume throughout the paper that the SGS is the union of sets \( U_i \) of coset representatives for \( G_{i-1} \mod G_i \). Thus, one can sift any \( g \in G \) through the SGS to find the unique factorization \( g = u_1 u_2 \cdots u_k \) with \( u_i \in U_i \).

2.1 The Problems

It was shown in [1] that given generators for \( G \leq \text{Sym}(\Omega) \) one could find, in NC, a base and SGS for \( G \). However, the question remained open as to whether or not one can sift, in NC, an element \( g \in G \) using the SGS. To be precise we state the sifting problem as given in [1].

SIPT Instance: An SGS, \( S = \bigcup_{i=1}^{k} U_i \), for \( G \leq \text{Sym}(\Omega) \) relative to a base \( B = b_1, b_2, \ldots, b_k \). An element \( g \in G \), and \( u \in U_k \).

Question: Does \( g = u_1 u_2 \cdots u_k \) where \( u_i \in U_i \) and \( u_k = u \)?

The other algebraic problem we consider is canonical forms of vertex-colored graphs with bounded color classes (CFBCC). Let \( C = \{C_1, C_2, \ldots, C_m\} \) be the set of colors for graph \( \Gamma(V, E) \), and let \( V(C_i) \) be the set of all vertices with color \( C_i \). We will assume that \( \Gamma(V, E) \) is color regular (i.e., all nodes in \( V(C_i) \) have the same \( C_i \)-valance).

For any pair of colors \( C_i, C_j \in C \) we order the pairs by,

\[ C_1 C_2, C_1 C_3, C_2 C_3, C_1 C_4, C_2 C_4, C_3 C_4, C_1 C_5, \ldots, C_{m-1} C_m. \]

Let \( \Gamma_{C_i, C_j} \) denote the induced bipartite graph on \( V(C_i) \) and \( V(C_j) \), and let \( \Delta_{ij} \) be the set of all bipartite graphs on \( V(C_i) \) and \( V(C_j) \). Since the color classes are bounded, \( \Delta_{ij} \) is bounded. If \( \Delta = \bigcup_{i,j} \Delta_{ij} \), then the vertex-colored graph \( \Gamma(V, E) \) can be viewed as a sequence of points from \( \Delta \).

Let \( G = \text{Sym}(V(C_1)) \times \cdots \times \text{Sym}(V(C_m)) \), then \( G \) acts naturally on \( \Delta \) and two vertex-colored graphs \( \Gamma_1 \) and \( \Gamma_2 \) are isomorphic if and only if there exists \( g \in G \) such that \( \Gamma_1^g = \Gamma_2 \). Using the \( V(C_i) \), \( V(C_j) \) blocks in the adjacency matrix of \( \Gamma \) we will associate each graph with a binary string. The string will be comprised of the blocks

\[ V(C_1) V(C_2), V(C_1) V(C_3), V(C_2) V(C_3), V(C_1) V(C_4), \ldots, V(C_{m-1}) V(C_m) \]

and within each block the elements are ordered using the columns of the block (i.e., all elements in column one come first then column two and so on). The canonical labeling for \( \Gamma \) will be the lexicographical largest graph under the action of \( G \). We now state formally the CFBCC problem.
CFBCC Instance: A vertex-colored graph \( \Gamma(V,E) \) with bounded color classes, and a position \((v,w)\) specified in the adjacency matrix \(M\).

Question: Does the canonical form \( \Gamma'(V,E) \) have a one in position \((v,w)\)?

To prove that SIFT and CFBCC are P-complete we reduce a restricted version of the P-complete problem, greedy independent set (GIS), to these algebraic problems. For completeness we sketch Cook's logspace reduction of the Monotone Circuit Value Problem (MCVP) to GIS, and point out why a restricted version of GIS remains P-complete. The MCVP is defined as follows:

MCVP Instance: A set of boolean functions \(g_1, g_2, \ldots, g_m\) where \(g_1 = 0, g_2 = 1\) and for \(3 \leq i \leq m, g_i\) is equal to either \(g_j \land g_k\) or \(g_j \lor g_k\), where \(j,k < i\).

Question: Does \(g_m = 1\)?

Let \(\Gamma(V,E)\) be a graph with vertex set \(V\) and edge set \(E\). A subset \(W \subseteq V\) is called an independent set of vertices in \(\Gamma(V,E)\), if for all \(w_1, w_2 \in W, (w_1, w_2) \notin E\).

There is a natural greedy algorithm for constructing a maximal independent set of vertices in \(\Gamma(V,E)\). Given a linear ordering of the vertex set \(V\), the greedy algorithm repeatedly picks the smallest vertex from \(V\) that is not adjacent to a previously selected vertex. The corresponding decision problem, greedy independent set, is defined as follows:

GIS Instance: Graph \(\Gamma(V,E)\) where \(V\) is linearly ordered.

Question: Is the last vertex in the ordering part of the greedy maximal independent set?

**Lemma 2.1** [Cook] The GIS problem is P-complete.

Proof: The GIS problem is clearly in P. To prove the problem is complete we sketch Cook's logspace reduction of MCVP to GIS.

Let \(g_1, g_2, \ldots, g_m\) be an instance of the MCVP. We construct a graph \(\Gamma(V,E)\) with vertex set \(V = \{u_1, u_2, \ldots, u_m\} \cup \{w_1, w_2, \ldots, w_m\}\). We order the vertices so that \(u_i\) and \(w_i\) precede \(u_j\) and \(w_j\), whenever \(i < j\). The ordering of \(u_i\) relative to \(w_i\) is determined by the gate \(g_i\). For any \(i, 3 \leq i \leq m, w_i\) precedes \(u_i\) if \(g_i = g_j \lor g_k\), and \(u_i\) precedes \(w_i\) if \(g_i = g_j \land g_k\). Let \(w_1\) precede \(u_1\) and let \(u_2\) precede \(w_2\). This gives us a linear ordering of the set \(V\).

The edge set, \(E\), is equal to \(E_1 \cup E_2 \cup E_3\), where

\[
E_1 = \{(u_i, u_j) | 1 \leq i \leq m\},
\]

\[
E_2 = \{(w_i, u_j), (w_i, w_k) | 3 \leq i \leq m \text{ and } g_i = g_j \lor g_k\} \text{ and}
\]

\[
E_3 = \{(u_i, w_j), (u_i, w_k) | 3 \leq i \leq m \text{ and } g_i = g_j \land g_k\}.
\]

Note that the construction of \(\Gamma(V,E)\) from the instance of the MCVP can be performed by a logspace algorithm. A simple induction argument shows that \(u_i\) is in the greedy independent set for \(\Gamma(V,E)\) if and only if \(g_i = 1\), and \(w_i\) is in the greedy independent set for \(\Gamma(V,E)\) if and only if \(g_i = 0\). □
Remark 2.2 Let $\Gamma(V, E)$ be an instance of the GIS problem where the linear ordering of $V$ is $v_1 < v_2 < \cdots < v_m$. The GIS problem remains P-complete even if we restrict ourselves to instances in which the following conditions are true. We assume that $(v_1, v_2) \notin E$ and each $v_i$, $3 \leq i \leq m$, is connected to exactly two distinct vertices that are smaller than itself.

Proof: It suffices to note that the following changes can be made to the reduction of MCVP to GIS. First, we may assume without loss of generality, that if $g_i = g_j \lor g_k$ (or $g_i = g_j \land g_k$) and $3 \leq i \leq m$, then $j \neq k$. Second, we may eliminate node $v_i$ from the construction of $\Gamma(V, E)$, and we may add edge $(w_1, w_2)$ to $E$. All the nodes in the set $X = \{v_i, w_i | 3 \leq i \leq m\}$ are connected to either 1 or 2 nodes smaller than themselves. For any node $x \in X$ connected to only 1 node smaller than itself, add the edge $(x, w_2)$ to $E$. $\square$

3 The Complexity of the Problems

Lemma 3.1 SIFT is P-complete.

Proof: Since sifting takes $O(nk)$ time [4] it follows that SIFT is in P. To show that SIFT is P-complete we describe a logspace reduction of GIS to SIFT.

Let $\Gamma(V, E)$ be an instance of GIS, with linear ordering $v_1 < v_2 < \cdots < v_m$. By Remark 2.2 we may assume, without loss of generality, that each $v_i$, is connected to at most two vertices less than itself.

Let $R_i$ be the right regular representation of $\mathbb{Z}_m$ with generator $a_i$, $1 \leq i \leq m$. Define $G = \langle a_i | i = 1, 2, \cdots, m \rangle$, then $G < Sym(\Omega)$ where $\Omega = \bigcup_{i=1}^{m} R_i$. Let $c_i = 1, c_{i+1} = c_i a_i, c_{i+2} = a_i c_i a_i, \cdots, c_{m^i}$, where

$$
\epsilon_j = \begin{cases} 
1 & \text{if } (v_i, v_j) \in E \\
0 & \text{otherwise.}
\end{cases}
$$

If $C_i = \{c_1, c_2, c_3\}$ for $i = 1, 2, \cdots, m$ then the $C_i$ are an SGS for $G$ relative to the base $B = a_1, a_2, \cdots, a_m$. To complete the instance of SIFT we let $g = a_1 a_2^2 \cdots a_m^m$ and $u = c_{m^i}$. This instance of SIFT can be constructed from an instance of GIS by an algorithm that uses no more than $O(\log m)$ space.

Let $V'$ be the greedy maximal independent set for $\Gamma(V, E)$, and let $g = u_1 u_2 \cdots u_m$ be the unique factorization of $g$. By Remark 2.2 we know that node $v_i$, $3 \leq i \leq m$, is connected to exactly two nodes, $v_j$ and $v_k$, smaller than itself. Observe that

$$
u_i = \begin{cases} 
c_1 & \text{if either } u_j = c_3 \text{ or } u_k = c_3, \text{ but not both} \\
c_2 & \text{if } u_j = c_3 \text{ and } u_k = c_3 \\
c_3 & \text{if } u_j \neq c_3 \text{ and } u_k \neq c_3.
\end{cases}
$$

A simple proof by induction shows that $v_i \in V'$ if and only if $u_i = c_3$. The hypothesis is clearly true for $i = 1$ and $i = 2$. For $3 \leq i$, $v_i \in V'$ if and only if $v_j \notin V'$ and $v_k \notin V'$. By the induction hypothesis we have $v_i \in V'$ if and only if $u_j \neq c_3$ and $u_k \neq c_3$. Thus, $v_m$ is in the greedy maximal independent set if and only if $u_m = c_{m^i}$. $\square$
Lemma 3.2. The CFBCC problem is P-complete.

Proof: First we sketch an algorithm that finds in polynomial time the canonical labeling for a vertex-colored graph $\Gamma(V, E)$ that has bounded color classes. Let $C = \{C_1, C_2, \ldots, C_m\}$ be the set of colors and let $V(C_i)$ be the set of all vertices with color $C_i$. Recall that $\Delta_{ij}$ is the set of all bipartite graphs on $V(C_i)$ and $V(C_j)$ and $\Delta = \bigcup_{1 \leq i, j \leq n} \Delta_{ij}$. The vertex-colored graph $\Gamma(V, E)$ can be viewed as a sequence of points from $\Delta$.

Let $G = \text{Sym}(V(C_1)) \times \cdots \times \text{Sym}(V(C_m))$, then $G$ acts naturally on $\Delta$ and two vertex-colored graphs $\Gamma_1$ and $\Gamma_2$ are isomorphic if and only if there exists $g \in G$ such that $\Gamma_1^g = \Gamma_2$. Since each $\Delta_{ij}$ is bounded the following algorithm runs in polynomial time.

\begin{align*}
x &:= 1 \\
G &= \text{Sym}(V(C_1)) \times \cdots \times \text{Sym}(V(C_m)) \\
&\text{For } i := 2 \text{ to } m \text{ do} \\
&\quad \text{For } j := 1 \text{ to } i - 1 \text{ do} \\
&\quad \quad \text{Find } xg \text{ in } xG \text{ that maps } \Gamma_{G_i, C_i} \text{ to the lexicographical largest in } \Delta_{ij} \\
&\quad \quad G^g = \text{Stabilizer of } (\Gamma_{G_i, C_i})^g \\
&\quad x := xg
\end{align*}

To show that CLBCC is P-complete it will suffice to describe a logspace reduction of GIS to CLBCC. Given an instance $\Gamma(V, E)$ of GIS, with the linear ordering $v_1, v_2, v_3$ on $V$, we will construct an instance of CLBCC.

Let $G = \langle a, b, c \rangle$ be elementary abelian group of order 8. We construct an instance $\tilde{\Gamma}(V, E)$ of CLBCC with $|\tilde{V}| = 32n$ and all color classes of size 8. For $1 \leq i \leq n$, we construct four replicas $G_{i1}, G_{i2}, G_{i3}, G_{i4}$ of $G$, with $G_{ij}$ assigned color $4(i - 1) + j$.

Then $V = \bigcup_{1 \leq i \leq n, 1 \leq j \leq 4} G_{ij}$.

Making use of the fixed identifications $G \cong G_{ij}$, we can define the edge sets $E(G_{ij}, G_{ij'})$ as subsets of $G \times G$. Unless indicated otherwise below, it is assumed that $E(G_{ij}, G_{ij'}) = \emptyset$.

For $1 \leq i \leq n$:

\begin{align*}
E(G_{i1}, G_{i2}) &= \{(x, y) \in G \times G \mid xy \in \{ab, ac, abc\}\}, \\
E(G_{i1}, G_{i3}) &= \{(x, y) \in G \times G \mid xy \in \{a\}\}, \\
E(G_{i1}, G_{i4}) &= \{(x, y) \in G \times G \mid xy \in \{b, c\}\}, \\
E(G_{i2}, G_{i3}) &= \{(x, y) \in G \times G \mid xy \in \{b, c\}\}, \\
E(G_{i3}, G_{i4}) &= \{(x, y) \in G \times G \mid xy \in \{a, bc\}\}.
\end{align*}

For $1 \leq i \leq n - 1$:

\[E(G_{i1}, G_{i+1, 1}) = \{(x, x) \mid x \in G\}.
\]

For each $(v_i, v_j) \in E$ with $i < j$:

\[E(G_{ia}, G_{jb}) = \begin{cases}
\{(x, y) \in G \times G \mid xy \in \{a, c\}\} & \text{if } i \text{ is minimal} \\
\{(x, y) \in G \times G \mid xy \in \{a, b\}\} & \text{otherwise}.
\end{cases}
\]

123
Finally,

\[ E(G_{1,1}, G_{n,3}) = \{(x, x) \mid x \in G\}. \]

To finish the construction of the FFBCC problem let \( v \) be the first vertex in \( G_{n,2} \) and let \( \bar{v} \) be the first vertex in \( G_{n,3} \). The graph \( \bar{\Gamma}(V, \bar{E}) \) is a vertex-colored graph with bounded color classes and \( (v, \bar{v}) \) is the specified position in the adjacency matrix.

The action of \( x \in G \) on the \( \bar{\Gamma}(V, \bar{E}) \) is by right multiplication by \( x \) in each \( G_{i,j} \cong G \).

To check that \( \bar{E}^x = \bar{E} \), we need only observe that \( xy = xzy \) since \( G \) is elementary abelian.

A tedious but straightforward proof by induction proves that there is an edge between the first vertex in \( G_{i,2} \) and the first vertex in \( G_{i,3} \) in the canonical form for \( \bar{\Gamma}(V, \bar{E}) \) if and only if \( v_i \) is in the GIS. \( \square \)

References


