A lower bound on the length of the shortest superpattern

Anonymous 4chan Poster, Michael Albert, Robin Houston, Jay Pantone, and Vince Vatter

November 5, 2018

This proof is inspired by that posted anonymously at

http://mathsci.wikia.com/wiki/The_Haruhi_Problem

which itself was taken from a 4chan discussion archived at

https://warosu.org/sci/thread/S3751105#p3751197

In this new version, after we prove the lower bound, we show that Greg Egan's construction is optimal if you forbid edges of weight 3 or more.

1 The graph

We (mostly) use the notation from Greg Egan's page at

http://www.gregegan.net/SCIENCE/Superpermutations/Superpermutations.html

In particular, we use the same graph as he does. The nodes of our graph are all of the permutations of length n. There is a directed edge joining every permutation to every other permutation, and its *weight* is the least number of symbols we need to add to the first permutation so that the last n symbols are equal to the second permutation. We also need to eliminate so-called improper edges: an edge is *improper* if performing the associated transformation of our word would visit another permutation on the way. So, for example, the edge of weight 2 from 12345 to 34512 is improper, because in performing this transformation we would pass through the permutation 23451. We will assume henceforth that the improper edges of the graph have been removed.

A *Hamiltonian path* through a graph is a path that visits every vertex exactly once. It is tempting to think that superpermutations correspond to Hamiltonian paths in the graph we have described, but there is nothing preventing a superpermutation from containing more than one copy of a permutation. Thus what we really want is a *walk* $\pi_1, \ldots, \pi_m$ in this graph that visits every vertex at least once. If there is an edge from π to σ, we denote the weight of this edge by $\mathrm{wt}(\pi, \sigma)$, and we define the weight of the walk $\pi_1, \ldots, \pi_m$ to be

$$\mathrm{wt}(\pi_1, \ldots, \pi_m) = \sum_{i=1}^{m-1} \mathrm{wt}(\pi_i, \pi_{i+1}).$$

Such a walk corresponds to a superpermutation, and the length of this corresponding superpermutation is precisely n greater than the weight of the walk (this n accounts for the number of symbols required for the first permutation π_1).

2 Cycles in the graph

Every vertex of our graph has precisely one edge of weight 1 leading out of it. This edge leads from the permutation π to its cyclic rotation $\pi(2)\cdots\pi(n)\pi(1)$. Thus if we follow $n-1$ consecutive weight 1 edges, one at a time, we visit the *cyclic class* of π, which we define to be the set of all cyclic rotations of π. The set of all permutations of length n is naturally partitioned in $(n-1)!$ of these cyclic classes. In any walk that visits every permutation, we obviously must visit each of these classes and each of their members.

We also need a slightly more complicated construction. Note that every vertex of our graph also has precisely one edge of weight 2 leading out of it (because we have removed the improper edges). This edge leads from the permutation π to the permutation $\pi(3)\cdots\pi(n)\pi(2)\pi(1)$.

The 2-*loop generated by* π is defined as the set of vertices visited by the walk that starts at π, follows $n-1$ consecutive edges of weight 1, then follows the edge of weight 2, and then repeats these steps $n-2$ more times. For example, in the graph on permutations of length 5, the 2-loop generated by 12345 is shown below. In this picture, the 1-steps are represented by black arrows and the 2-steps are represented by orange arrows.

$$
\begin{tikzcd}
\mathbf{12345} \arrow[r] & 23451 \arrow[r] & 34512 \arrow[r] & 45123 \arrow[r] & 51234 \arrow[d] \\
52341 \arrow[d] & 15234 \arrow[l] & 41523 \arrow[l] & 34152 \arrow[l] & \mathbf{23415} \arrow[l] \\
\mathbf{34125} \arrow[r] & 41253 \arrow[r] & 12534 \arrow[r] & 25341 \arrow[r] & 53412 \arrow[d] \\
54123 \arrow[uuu, bend left] & 35412 \arrow[l] & 23541 \arrow[l] & 12354 \arrow[l] & \mathbf{41235} \arrow[l]
\end{tikzcd}
$$

In this example, inspection reveals that the 2-loop generated by 12345 is also generated by 23415, 34125, and 41235. More generally, we have the following result.

Proposition 1. *If a* 2*-loop is generated by* π*, then it is generated by all* $n-1$ *permutations obtained by fixing the last entry of* π *and cyclically permuting the other entries, i.e., by* π *and the permutations*

$$\pi(2)\cdots\pi(n-1)\pi(1)\pi(n), \quad \pi(3)\cdots\pi(n-1)\pi(1)\pi(2)\pi(n), \quad \ldots, \quad \pi(n-1)\pi(1)\cdots\pi(n-2)\pi(n).$$

Note that each 2-loop consists of the disjoint union of the cyclic classes of each of its generators. This is one way to see that every 2-loop contains $n(n-1)$ permutations. It is also clear from Proposition 1 that there are $n(n-2)!$ distinct 2-loops (every permutation generates a 2-loop and every 2-loop is generated by $n-1$ permutations).

We say that a walk *enters* the 2-loop generated by π if it follows an edge of weight 2 or more to arrive at π. This has the somewhat odd implication that the notion of which 2-loop we are in is dependent not on the node we are currently at, but on the way we got there. Because each 2-loop contains only $n(n-1)$ permutations, a walk that visits every permutation must enter at least $(n-2)!$ different 2-loops.

3 The proof

Theorem 2. *Every superpermutation for the set of permutations of length n has length at least*

$$n! + (n-1)! + (n-2)! + n - 3.$$

Proof. For a walk $\pi_1, \ldots, \pi_m$ in the graph described above, we define the three parameters

$$
\begin{aligned}
p(\pi_1,\ldots,\pi_m) &= \text{the number of distinct permutations visited,} \\
c(\pi_1,\ldots,\pi_m) &= \text{the number of cyclic classes completed in the walk } \pi_1,\ldots,\pi_{m-1}, \text{ and} \\
v(\pi_1,\ldots,\pi_m) &= \text{the number of 2-loops visited.}
\end{aligned}
$$

In the above parameters, we say that a walk has *completed* a cyclic class if it has visited all of the vertices of that class, and that it has *visited* a 2-loop if it has entered that 2-loop at least once. Our aim is to show that

$$\mathrm{wt}(\pi_1, \ldots, \pi_m) \geq p(\pi_1, \ldots, \pi_m) + c(\pi_1, \ldots, \pi_m) + v(\pi_1, \ldots, \pi_m) - 2. \tag{†}$$

This claim holds in the base case, $m = 1$, because $\mathrm{wt}(\pi_1) = 0$, $p(\pi_1) = 1$, $c(\pi_1) = 0$, and $v(\pi_1) = 1$. Now suppose that the inequality is true for all walks of length m and consider a walk $\pi_1, \ldots, \pi_m, \pi_{m+1}$. Our proof depends on the weight of the edge from π_m to π_{m+1}.

- If $\mathrm{wt}(\pi_m, \pi_{m+1}) = 1$, then π_m and π_{m+1} lie in the same cyclic class, so the value of v cannot increase. If we have visited π_{m+1} before, then the value of p does not increase, and we are done. If we have not visited π_{m+1} before, then since π_m and π_{m+1} are in the same cyclic class, p_m did not complete its cyclic class, and c does not increase. In either case, (†) holds.

- If $\mathrm{wt}(\pi_m, \pi_{m+1}) = 2$ then
$$\pi_{m+1} = \pi_m(3) \cdots \pi_m(n)\pi_m(2)\pi_m(1).$$
We claim that if the value of c increases, then the value of v cannot change. Suppose that the value of c increases, so
$$c(\pi_1, \ldots, \pi_m, \pi_{m+1}) = c(\pi_1, \ldots, \pi_m) + 1.$$
This implies that π_m completed its cyclic class, so we had not previously visited it. Because π_m completes its cyclic class, we must have already visited the permutation we would otherwise get to via a weight 1 edge from π_m,
$$\sigma = \pi_m(2)\pi_m(3) \cdots \pi_m(n)\pi_m(1).$$
However, we didn't visit σ from π_m because we hadn't visited π_m before, and thus we must have taken an edge of weight at least 2 to visit σ. This implies that we have already entered the 2-loop generated by σ. Finally, Proposition 1 shows that σ and π_{m+1} generate the same 2-loop. Thus visiting π_{m+1} does not take us to a new 2-loop, so the value of v does not increase. Having shown that at most one of c or v can increase when traversing an edge of weight 2, (†) is verified in this case.

- If $\mathrm{wt}(\pi_m, \pi_{m+1}) \geq 3$, then since the right-hand side of (†) can increase by at most 3 when traversing a single edge of the graph, (†) holds.

With (†) established, the proof of the theorem follows easily. If the walk $\pi_1, \ldots, \pi_m$ visits every permutation then clearly $p(\pi_1, \ldots, \pi_m) \geq n!$. Also, the walk must complete all $(n-1)!$ cyclic classes, so we must have $c(\pi_1, \ldots, \pi_m) \geq (n-1)! - 1$, and we must visit at least $(n-2)!$ 2-loops. This shows that

$$\mathrm{wt}(\pi_1, \ldots, \pi_m) \geq n! + (n-1)! + (n-2)! - 3.$$

Finally the superpermutation corresponding to this walk has length $n + \mathrm{wt}(\pi_1, \ldots, \pi_m)$, so the length of this superpermutation is at least $n! + (n-1)! + (n-2)! + n - 3$, as desired. □

4 If you forbid edges of weight 3 or more, Greg Egan's construction is best possible

The first observation to make is that every 2-loop consists of $n - 1$ cyclic classes. Another (trivial) observation is that by taking an edge of weight 1, we do not change 2-loops. Then we have the most important observation:

Proposition 3. *If we take an edge of weight 2 from (any node of) one 2-loop to another 2-loop, those two 2-loops must share at least one cyclic class.*

Proof. Let the edge of weight 2 be

$$\pi(n)\pi(n-1)\pi(1)\cdots\pi(n-2) \to \pi.$$

This means that the 2-loop we are leaving contains the cyclic class of $\pi(n)\pi(n-1)\pi(1)\cdots\pi(n-2)$, which is the same as the cyclic class of $\pi(1)\cdots\pi(n-2)\pi(n)\pi(n-1)$.

By Proposition 1, the 2-loop we are entering is generated by π and also by $\pi(n-1)\pi(1)\cdots\pi(n-2)\pi(n)$. Thus our new 2-loop contains the cyclic class of $\pi(n-1)\pi(1)\cdots\pi(n-2)\pi(n)$, which is the same as the cyclic class of $\pi(1)\cdots\pi(n-2)\pi(n)\pi(n-1)$. □

Now consider a walk through our graph which does not use any edges of weight 3 or more (as Greg Egan's construction does). We begin in a 2-loop which has $n-1$ cyclic classes. No matter how we leave this 2-loop, Proposition 3 shows that the next 2-loop we enter has at most $n-2$ *new* cyclic classes. Continuing in this manner, and using the notation of the proof of Theorem 2, we see that the number of cyclic classes we have visited (i.e., seen at least one permutation of) is at most

$$n-1+(v-1)(n-2) = v(n-2)+1.$$

Since we must visit all of the $(n-1)!$ cyclic classes, we must have

$$v(n-2)+1 \geq (n-1)!,$$

which implies that (for $n \geq 3$)

$$v \geq (n-1)(n-3)! - \frac{1}{n-2} = (n-2)! + (n-3)! - \frac{1}{n-2}.$$

Since v is an integer, we conclude that $v \geq (n-2)! + (n-3)!$. Using this lower bound in the proof of Theorem 2, we obtain the following.

Proposition 4. *Suppose $n \geq 3$. If a walk through our graph visits every permutation and traverses no edges of weight 3 or more, than it corresponds to a superpermutation of length at least*

$$n! + (n-1)! + (n-2)! + (n-3)! + n - 3.$$