NOTES ON MAXIMAL SUBSEMIGROUPS ALGORITHMS

WILF WILSON

1. Introduction

A group led by James Mitchell is implementing an algorithm in the Semigroups package for GAP. We are given an arbitrary finite semigroup S, specified by its generators. We wish to return a list of all the maximal subsemigroups of S, each of which will be specified by a generating set.

The paper I mention in this document is called *Maximal Subsemigroups of Finite Semigroups*, by N. Graham, R. Graham, and J. Rhodes.

2. The algorithm

Here is a rough outline and justification for each step of the `MaximalSubsemigroups` algorithm for an arbitrary transformation semigroup.

2.1. **Irredundant Generating Subset.**

Firstly we calculate an irredundant generating set for S, which we call X. We know that any maximal subsemigroup must lack at least one generator of S. For suppose M is a maximal subsemigroup of S containing X. Then $S = <X> \subseteq M \subseteq S$, i.e. $S = M$, a contradiction.

Suppose our generating set for S is given by $X = \{g_1, g_2, ..., g_n\}$. Then for each $i \in \{1, 2, ..., n\}$, since X is irredundant, $<X \setminus \{g_i\}>$ is a proper subsemigroup of S, and hence is contained in some maximal subsemigroup of S. Suppose $<X \setminus \{g_i\}>$ and $<X \setminus \{g_j\}>$ for $i \neq j$ were contained in the same maximal subsemigroup, M. This implies that M contains the generating set of each subsemigroup. But the union of both these sets is precisely X. Therefore M contains a generating set for S, and so $M = S$, a contradiction.

The above argument tells us that each generator of X will give rise to at least one maximal subsemigroup, and maximal subsemigroups arising from removing distinct (single) generators will be distinct. In particular, the number of maximal subsemigroups of S is at least as many as the size of any irredundant generating set for S.

Note that a maximal subsemigroup M of S may lack more than one generator of S (give an example of how).

Part (1) of the proposition tells us that any maximal subsemigroup of S contains all but one $\mathscr{D}$-class of S. Obviously no maximal subsemigroup M of S can arise from removing part of a $\mathscr{D}$-class which contains no generators, as M would still contain all of the generators of S. Therefore we only need to consider those $\mathscr{D}$-classes which contain at least one generator of S. Requiring irredundancy will thus hopefully reduce the number of $\mathscr{D}$-classes to consider.

2.2. **Partial Order.** We thus next wish to know which $\mathscr{D}$-classes contain generators, and also where they are located in terms of the $\mathscr{D}$-class partial order. We use the partial order command to produce a list of such $\mathscr{D}$-classes which are maximal, and a list of such $\mathscr{D}$-classes which are not maximal. We also record which generators are in which $\mathscr{D}$-classes. Note that all maximal $\mathscr{D}$-classes must contain at least one generator, since elements of a maximal $\mathscr{D}$-class can not be realised as products of elements in other $\mathscr{D}$-classes.

2.3. **Maximal $\mathscr{D}$-classes of size 1.**

First, an aside:

A non-regular maximal $\mathscr{D}$-class has size 1. To see why, let $x \in \mathscr{D}x$, a non-regular maximal $\mathscr{D}$-class, and let $y \in S^1$. If $y \in \mathscr{D}x$, then xy is below $\mathscr{D}x$ in the $\mathscr{D}$-class partial order since $\mathscr{D}x$ is non-regular, and if $y \in S \setminus \mathscr{D}x$, xy is again below $\mathscr{D}x$, since $\mathscr{D}x$ is maximal. Therefore only right multiplication by the identity 1 gives a product still in $\mathscr{D}x$, so $\mathscr{R}x = \{x\}$. Similiarly, $\mathscr{L}x = \{x\}$. Hence $\mathscr{D}x = \{x\}$, i.e. $|\mathscr{D}x| = 1$. Alternatively, note that a non-regular $\mathscr{D}$-class can only contain at most one of our irredundant generators. I previously argued that each irredundant generator gives rise to distinct maximal subsemigroups. However the paper tells us that the only maximal subsemigroup a non-regular $\mathscr{D}$-class gives rise to is formed by removing the whole $\mathscr{D}$-class. If a non-regular $\mathscr{D}$-class contained two generators, both would give rise to precisely the same maximal subsemigroup, contradicting my argument.

Since any element of a non-regular maximal D-class must be a generator, we conclude it must be a singleton.

Each maximal $\mathscr{D}$-class $\mathscr{D}$ of size 1 corresponds to precisely one maximal subsemigroup M, since $M := S \setminus \mathscr{D}$ is a semigroup, and any subsemigroup of S properly containing M necessarily contains all the elements of S, and hence equals S. We can specify a generating set for such a subsemigroup by including the other generators for S, and generators for the $\mathscr{D}$-classes directly below $\mathscr{D}$ in the partial order.

2.4. **Maximal $\mathscr{D}$-classes of size $\geq$ 2.** Next we consider the rest of the maximal $\mathscr{D}$-classes; note that they are necessarily regular and larger than size 1 by the above argument.

Let $\mathscr{D}$ be a maximal $\mathscr{D}$-class. For a subset $U \subseteq \mathscr{D}$, since $\mathscr{D}$ is maximal, we have that $Y := (S \setminus \mathscr{D}) \cup U$ is a subsemigroup of S if and only if U satisfies:

$$\forall u_1, u_2 \in U : u_1u_2 \in \mathscr{D} \Rightarrow u_1u_2 \in U. \tag{1}$$

i.e. if and only if U is sort of closed.

For a subset T of S, define T^0 be the semigroup $T \cup \{0\}$, where $0 \notin T$, with multiplication defined by:

$$t_1 * t_2 = \begin{cases} t_1t_2, & \text{if } t_1t_2 \in T, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Let $U \subseteq \mathscr{D}$ be a subset of $\mathscr{D}$. Then U satisfies equation (1) if and only if U^0 is a subsemigroup of $\mathscr{D}^0$. Therefore, it's not hard to see that $Y := (S \setminus \mathscr{D}) \cup U$ is a maximal subsemigroup of S if and only if U^0 is a maximal subsemigroup of $\mathscr{D}^0$.

Therefore our task is reduced to finding maximal subsemigroups of $\mathscr{D}^0$. Since $\mathscr{D}^0$ is zero-simple, the Rees theorem says that $\mathscr{D}^0$ is isomorphic to a Rees zero-matrix semigroup. We call this the principal factor of $\mathscr{D}$.

We have a separate algorithm which calculates all maximal subsemigroups of a Rees-zero matrix semigroup, which I will describe another time. The algorithm takes an injection to the principal factor $\mathscr{M}$ of $\mathscr{D}$, and finds the maximal subsemigroups of $\mathscr{M}$. We then adjoin the pre-image of each maximal subsemigroup of $\mathscr{M}$ to $S \setminus \mathscr{D}$, giving us a maximal subsemigroup of S.

2.5. **Non-maximal non-regular $\mathscr{D}$-classes.** Let $\mathscr{D}$ be a non-regular $\mathscr{D}$-class containing at least one generator. We know that removing part of this $\mathscr{D}$-class will give rise to a maximal subsemigroup. However, part (3) of the proposition tells us that we must remove the whole of $\mathscr{D}$ to produce a maximal subsemigroup. We do this by specifying the generators of S not in this $\mathscr{D}$-class, and generators corresponding to the $\mathscr{D}$-classes directly below $\mathscr{D}$ in the partial order. This is sufficient to generate $S \setminus \mathscr{D}$.

2.6. **Non-maximal regular $\mathscr{D}$-classes.** We then look through the non-maximal $\mathscr{D}$-classes which contain generators.

Unfortunately, we can not use the same strategy as with maximal regular $\mathscr{D}$-classes, where we often used the fact that we had a maximal $\mathscr{D}$-class. Indeed, subsets of S corresponding to maximal subsemigroups of the principal factor need not be subsemigroups: This is because a product of two elements above a partially-removed $\mathscr{D}$-class could end up in the removed section.

Therefore, we have to look at the semigroup as a whole as well as our $\mathscr{D}$-class. Let $\mathscr{D}$ be the non-maximal regular $\mathscr{D}$-class we are considering. Currently, the algorithm works roughly like this (recall that X is our irredundant generating set for S):

- Firstly we calculate generators for the $\mathscr{D}$-classes directly below $\mathscr{D}$ in the partial order. Call this set of generators Z. We will need these later to specify maximal subsemigroups arising from $\mathscr{D}$.
- We work through all subsets of $X \cap \mathscr{D}$, from smallest to largest. Call the current such subset Y. We will search for maximal subsemigroups of S lacking the generators in Y. Note that any maximal subsemigroup of S arising from $\mathscr{D}$ must lack some such subset Y.
- We will search through $\mathscr{D}$ to better understand the parts of $\mathscr{D}$ which any maximal subsemigroup of S not containing Y must (and must not) contain. Proceed as follows:
  - Let $U :=< X \setminus Y >$. (In fact we can restrict U to be generated by just the generators in X which are above $\mathscr{D}$ in the partial order - implementation detail. However $<X \setminus Y>$ doesn't hurt).
  - Let $A := \mathscr{D} \setminus (U \cup Y)$.

  - Let $XX := Y$.
- XX is going to be the set of elements of $\mathscr{D} \setminus U$ which, when added back to U, will generate an element of Y, i.e. $XX = \{x \in \mathscr{D} : \ <U, x> \cap Y \neq \emptyset\}$. We calculate XX as follows:
  - Take any element $a \in A$.
  - Let $C :=<U, a>$.
    * If $C \cap XX \neq \emptyset$, adding a generates some element of Y. Therefore add a to XX. Remove a from A since we have now considered it.
    * Otherwise $C \cap XX = \emptyset$. We have that:

      $$\begin{aligned} &\forall a' \in C \cap A : U \cup \{a'\} \subseteq C \\ &\Rightarrow <U, a'> \leq C \\ &\Rightarrow <U, a'> \cap XX = \emptyset \\ &\Rightarrow a' \notin XX. \end{aligned}$$

      Therefore set $A := A \setminus C$, since we have now considered all elements in $C \cap A$.
  - Eventually $A = \emptyset$ and we stop.
  - Redefine $A := \mathscr{D} \setminus (XX \cup U)$ (i.e. now $S = (S \setminus \mathscr{D}) \cup A \cup XX$, all disjoint).

We now might have enough information to specify some maximal subsemigroups.

- If $|XX| = |\mathscr{D}|$, then U does not intersect $\mathscr{D}$ (hence $S \setminus \mathscr{D}$ is a semigroup), and adding in any element of $\mathscr{D}$ will generate one of the generators in Y.
  - If $|Y| = 1$, then clearly $S \setminus \mathscr{D}$ is a maximal subsemigroup of S. Note that this also implies there is only one generator in $\mathscr{D}$.
  - If $|Y| > 1$, then $S \setminus \mathscr{D}$ is not a maximal subsemigroup of S. This is because $S \setminus \mathscr{D}$ would be the only maximal subsemigroup arising from this $\mathscr{D}$-class, and yet I have argued that each generator gives rise to at least one distinct maximal subsemigroup. i.e. We would be contradicting irredundancy of our generating set.
  - If $|Y| > 1$, we have $S \setminus \mathscr{D}$ is a non-maximal subsemigroup of S, but is only contained in subsemigroups containing some of Y (by definition of XX). Therefore the maximal subsemigroups in which $S \setminus \mathscr{D}$ is contained all lack some smaller set Y. We will have encountered and found them at some previous point in our search, since we have already considered all sets Y of smaller size. Therefore stop this search and pick the next set Y.
- We now know $|XX| < |\mathscr{D}|$. If still $A = \emptyset$, then $S \setminus XX =<X \setminus Y, Z>=<U, Z>$ is a subsemigroup of S.
  - If $|Y| = 1$, then clearly $S \setminus XX$ is a maximal subsemigroup of S.
  - If $|Y| > 1$, then adding back any one element of Y will not generate the whole of S since our generating set X is irredundant. So $S \setminus XX$ is not maximal. However let's consider which maximal subsemigroups it *could* be contained in. Any subsemigroup of S properly containing $S \setminus XX$ contains an element

of XX. Hence, for $x \in XX$, if our construction of XX is correct, then the resulting semigroup, $<U, Z, x>$, intersects Y. If this semigroup is a proper subsemigroup of S, we will have already found the maximal subsemigroups it is contained in, as they would lack a smaller subset of generators, Y.
So we won't find any new maximal subsemigroups arising from this subset, Y. Go to the next subset, Y.

- Otherwise $A \neq \emptyset$. If $S \setminus XX$ is a semigroup:
  - If $|Y| = 1$, it is clearly maximal.
  - If $|Y| > 1$, we know that adding in any element of XX will give us an element of Y. However, it is not yet clear to me whether $S \setminus XX$ can be contained in any maximal subsemigroups of S which we have not already seen.
    The question reduces to asking: can $S \setminus XX$ ever itself be maximal? For if $S \setminus XX$ is properly contained in some other maximal subsemigroup M, M must (by definition of XX) not lack all of the generators of Y, and hence we will have seen M before.
    By this argument, $S \setminus XX$ is not maximal if and only if it is contained within some maximal subsemigroup which we have already seen. Therefore, perform this check! In any case, we are now finished with Y.
    Perhaps we can prove that $S \setminus XX$ is never maximal. However, we can check this computationally for now.

Finally, if we have reached this point, we know that $S \setminus XX$ is not a semigroup. We know that any maximal subsemigroup M of S lacking precisely the generators $Y \subset \mathscr{D}$ must be such that $S \setminus \mathscr{D} \subsetneq M \subsetneq (S \setminus \mathscr{D}) \cup A = S \setminus XX$.

We perform a recursive search on A to find out which such subsemigroups M satisfying the above line are maximal. I'll describe this `YannRecursion` soon, once I actually understand what it does.

3. Auxiliary Algorithms

3.1. `IsMaximalSubsemigroup`. This takes two arguments, S and T, and asks whether T is a maximal subsemigroup of S. First the algorithm checks that $T \leq S$ and $T \neq S$. If either of these are false, the algorithm returns false. Otherwise, the algorithm checks that all elements of S not in T, when added to the generating set for T, generate the whole of S. The truth of this condition determines whether T is maximal in S.

4. Notes

4.1. **Groups.** All subsemigroups of finite groups are also groups. Why? If T is a subsemigroup of G, for $a \in T$, the closure of T implies that all powers of a are in T, which includes a^{-1} and 1 (since a has finite order). Therefore if our semigroup S is a group (`IsGroupAsSemigroup`), then the maximal subsemigroups are precisely the maximal subgroups. Take an isomorphism to a group, take the `MaximalSubgroups` of that group, and take their pre-image under the isomorphism.

The algorithm currently fails for groups as semigroups. We only have a single $\mathscr{D}$-class, and in that case the variable `gens2` becomes an empty list at one point, and we call $\texttt{SemigroupIdealByGenerators}(\texttt{S}, \texttt{gens2})$, which will obviously fail.