Minimal Generating Sets in Magmas*

Mikoláš Janota[1], António Morgado[2], and Petr Vojtěchovský[3]

[1] Czech Technical University, Prague
[2] INESC-ID, Lisboa
[3] University of Denver

A subset S of an algebra A *generates* A if the smallest subalgebra $\langle S\rangle$ of A that contains S is all of A. A generating subset of A of smallest possible cardinality is called a *minimal generating set*. In other words, calculating the closure of S, i.e., applying exhaustively the multiplication operation of A, S generates all the elements of A. The *rank* of an algebra is the cardinality of its minimal generating set.

Finding small and minimal generating sets is of importance in algebra, both theoretically and for the purposes of computations. For instance, a vector space is completely characterized by it rank (that is, dimension) and the underlying field. Groups with a single generator are very easy to understand while groups with two generators can be in some sense arbitrarily complicated. Alternative algebras with two generators are associative and hence relatively easy to understand compared to general alternative algebras. In computational group theory, the efficiency of algorithms often depends heavily on the number of generators given.

We present a method for calculating minimal generating sets in *magmas* (sets with a single binary operation) by means of SAT solvers and integer linear programs. This method cannot compete with specialized algorithms in highly structured magmas, such as groups, but it appears to be efficient in the general case. We focus on *loops*, that is, magmas M with identity element in which all translations $y \mapsto yx$ and $y \mapsto xy$ are bijections.

The main idea is as follows: Let M be a finite magma of size n and let S be any nonempty subset of M. If $\langle S\rangle = M$ then S is a generating set. Otherwise $\langle S\rangle < M$ and *every* generating set of M must contain an element from the complement $M \setminus \langle S\rangle$. (Indeed, if A is a generating set of M such that $A \cap (M \setminus \langle S\rangle) = \emptyset$ then $A \subseteq \langle S\rangle$ and $\langle A\rangle \leq \langle S\rangle < M$, a contradiction.)

Every subset $S \subseteq M$ with $\langle S\rangle < M$ therefore yields a restriction $t(S)$ on every generating set of M, in particular on every minimal generating set of M. This restriction can be expressed as a condition suitable for SAT solvers, namely $t(S) = \bigvee_{x\in M\setminus\langle S\rangle} x$, and can be readily translated into a constraint of an integer linear program (see below).

Given a collection $\{S_i : i \in I\}$ of subsets of M, any generating set must satisfy the conjunction $t(I) = \bigwedge_{i\in I} t(S_i)$. Finding a candidate for a minimal generating set is equivalent to solving the corresponding minimal hitting set problem (which is in general NP-complete).

To prove that M has rank larger than k, it suffices to find a collection $\{S_i : i \in I\}$ of subsets of M for which the following integer linear program is infeasible (unsatisfiable).

$$x \in \{0,1\} \text{ for every } x \in M, \tag{1}$$

$$\sum_{x\in M-\langle S_i\rangle} x \geq 1 \text{ for every } i \in I, \tag{2}$$

$$\sum_{x\in M} x \leq k. \tag{3}$$

*The results were supported by the Ministry of Education, Youth and Sports within the dedicated program ERC CZ under the project POSTMAN no. LL1902. This work was supported by national funds through FCT, Fundação para a Ciência e a Tecnologia, under project UIDB/50021/2020, the project INFOCOS with reference PTDC/CCI-COM/32378/2017.

If the above formulation becomes unsatisfiable, k must be increased. If the formulation is satisfiable, we obtain a set of elements that represent a *candidate* $S = \{x \in M : x = 1\}$ for the generator. If $\langle S\rangle = M$, we are done because the candidate is an actual generator; because k was increased only when needed, it is also guaranteed that this generator is minimal. Otherwise, if $\langle S\rangle < M$, we add S to our collection of subsets of M. Effectively, this means adding the restriction $\sum_{x\in M\smallsetminus\langle S\rangle} x \geq 1$ to Equation 1.

This approach can be seen as an instantiation of the framework proposed by Saikko et al. [4], which shows that a class of problems can be tackled by iterative generation of the minimal hitting set problem. We remark that the problem can be directly encoded as a single SAT problem; this formulation is cubic, which has proven prohibitive. Applying SAT technology on the minimal hitting set instances obtained from this process (Equation 1) exhibited poor results. The integer linear programming solver gurobi [2] proved to be far more adequate.

We ran experiments on groups and on *Moufang loops*. Moufang loops are loops satisfying the identity $x(y(xz)) = ((xy)x)z$ and are closely related to the alternative algebras mentioned above.

`GAP` [1] contains extensive libraries of groups. There exist very efficient algorithm for $r(G)$ if G is a solvable group. We calculated $r(G)$ for all nonsolvable groups of order less than 2048. In all cases, we verified the rank $r(G)$ as posted in `GAP`. (In some instances the generating set of G stored in `GAP` is larger than $r(G)$ but then our $r(G)$ can be verified heuristically by methods of `GAP`.)

The package `LOOPS` [3] of `GAP` contains all nonassociative Moufang loops of order $n \leq 64$ and of orders $n = 81$ and $n = 243$. For instance, there are 4262 such loops of order 64 and 5 of order 81. No efficient methods for calculating the rank of Moufang loops are known. We calculated $r(M)$ for all Moufang loops M of the form $M = A \times G$, where A is a Moufang loop from the library of `LOOPS` and G is the cyclic group of order 8. Here, $r(A) > 2$ due to Moufang theorem and $r(G) = 1$. In future experiments we want to include products with all groups of order 8 (there are 5 of them).

All the instances were solved with the average time of 1.75 s. The calculated r remains small for the considered instances, typically 3, 4, 5. Interestingly, the largest considered loops of order $1944 = 243 \times 8$ have all rank 3; only several loops of order $512 = 256 \times 8$ have the maximal rank found 6. The number of iterations needed, i.e., size of Equation 1, is also typically small, in the range of hundreds.

The following remarks are specific to generating sets in (Moufang) loops and groups and play a role in the search.

- If $S \leq M$ and M is a finite Moufang loop then $|S|$ divides $|M|$. (This is false in general loops.)
- If $S \leq M$ and $x, y \in M$ then the cosets xS, yS might interest nontrivially (that is, $xS \cap yS \neq \emptyset$ and $xS \neq yS$) but the cosets xS, S either coincide or are disjoint.
- If $S < M$ then $|S| \leq |M|/2$. Consequently, $|M \setminus S| \geq n/2$ and the number of variables in every term $t(S)$ is large, resulting in a difficult hitting set problem that SAT solvers struggle with.
- The rank $r(M)$ of M is at most $\lfloor \log_2(|M|)\rfloor$.
- If A, B are finite loops then $r(A \times B) \leq r(A) + r(B)$. It is not well understood when the equality holds.

References

[1] The GAP Group. *GAP – Groups, Algorithms, and Programming, Version 4.11.1*, 2021.

[2] LLC Gurobi Optimization. Gurobi optimizer reference manual, 2021.

[3] GP Nagy and P Vojtěchovský. LOOPS, a package for gap 4.3. *Download GAP at https://www.gap-system.org/ GAP.*, 2006.

[4] Paul Saikko, Johannes Peter Wallner, and Matti Järvisalo. Implicit hitting set algorithms for reasoning beyond NP. In Chitta Baral, James P. Delgrande, and Frank Wolter, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016*, pages 104–113. AAAI Press, 2016.