

The complexity of finding minimum generating sets

Bartosz Walczak

We discuss the following problem: Given a finite algebra A , what is the minimum size of a subset of A which generates A ? It finds a natural application in machine learning, where one can study what the maximum possible compression rate of a knowledge base is if instead of its entire structure, only its generating set is stored.

The problem is NP-complete in general. On the other hand, there are many natural classes of algebras where the problem is polynomial-time solvable. Therefore, a natural question arises what classes of algebras are tractable. One possible direction of answering this question is to study the minimal algebras in the sense of Tame Congruence Theory, i.e. \mathcal{G} -sets, vector spaces, Boolean algebras, distributive lattices, and semilattices. Among these classes only distributive lattices are not known to be tractable. We present an interesting combinatorial characterization of minimum generating sets in distributive lattices, which may help to find a polynomial-time algorithm for this class. If such an algorithm indeed exists, further study of interaction between minimal algebras that preserves tractability will result in broader classes of tractable algebras and will help to understand the nature of the problem.

The presented characterization also provides an elegant solution of the problem in the case of Boolean lattices.