Automatic Generation of Classification Theorems for Finite Algebras

Simon Colton[1] Andreas Meier[2*] Volker Sorge[3] Roy McCasland[4]

[1] Department of Computing, Imperial College London, UK, sgc@doc.ic.ac.uk

[2] DFKI GmbH, Saarbrücken, Germany, ameier@ags.uni-sb.de

[3] School of Computer Science, University of Birmingham, UK, V.Sorge@cs.bham.ac.uk

[4] School of Informatics, University of Edinburgh, UK, rmccasla@dai.ed.ac.uk

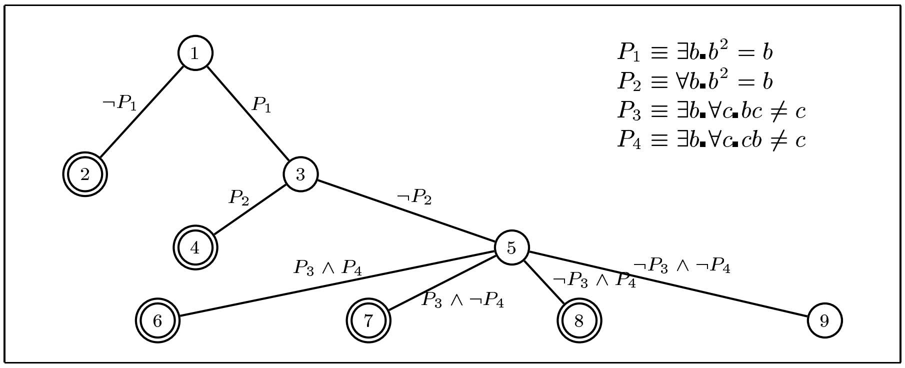
Classifying finite algebraic structures has been a major motivation behind much research in pure mathematics. Automated techniques have aided this process, but this has largely been at a quantitative level, e.g., to prove that there are no quasi-groups of a given type for a given size, or to count the number of groups of a particular order. Classification theorems of a more qualitative nature are often more interesting. For example, Kronecker's classification of finite Abelian groups [1] states that every Abelian group, G, of size n can be expressed as a direct product of cyclic groups, $G = C_{s_1} \times \cdots \times C_{s_m}$, where $n = s_1.s_2 \ldots s_m$ such that each s_{i+1} divides s_i. More recently, in 1980 the group theory community announced that a full classification theorem for finite simple groups had been completed, [2]. This has been described as one of the major intellectual achievements of all time [3].

Progress towards full classification theorems can often be made by classifying algebras of a particular size, because concepts used to classify algebras of one size may be used in the full classification theorem. We have looked at automating the task of generating and fully verifying qualitative classification theorems for algebraic structures of a given size. As a simple example, our system is given the axioms of group theory and told to find a classification theorem for groups of size 6. It returns the following (paraphrased) result: "all groups of size 6 can be classified up to isomorphism as either Abelian or non-Abelian" [where an Abelian group, G, is such that all pairs of elements commute, i.e., $\forall a, b{\in}G_{\blacksquare}\ a \circ b = b \circ a$]. The system generates such results, then proves that they provide valid classifications by showing that each concept is a *classifying property*, i.e., true for all members of exactly *one* isomorphism class.

In our first – semi-automated – approach to generating and verifying classification theorems, the Mace model generator was used to generate representatives of each isomorphism class for the given algebra of given size, then the HR and C4.5 machine learning systems were used to induce a set of classifying properties. Using these concepts, we employed the Gap computer algebra system to produce a set of theorems in terms of generators and factorisations which we then proved using the Spass theorem prover. Taken together, the proved theorems verified that the classification was valid. After experimenting with various schemes in this first approach, we identified some limitations, which enabled us to specify the following requirements for an algorithm in a second, more refined, approach:

- The process should be entirely automatic and bootstrapping, able to produce verified classification theorems starting from the axioms alone, with no human intervention.

*Author's work was supported by EU IHP grant Calculemus HPRN-CT-2000-00102.



- The process should call HR to find classifying concepts for small numbers of isomorphism classes, as HR struggled to solve larger classification problems.
- The process should not require the production of *every* algebraic structure satisfying the axioms. This is because, when using Mace to generate all structures and reduce them using its isomorphism filter, we found that the intermediate files produced could often be unmanageably large (up to 4GB).
- The process should generate the classification theorem as a decision tree. We found that decision trees often involved fewer concept definitions and enabled easier classification of algebras.

Hence, in a second approach, we implemented a fully automated bootstrapping procedure that builds a decision tree which can be used to decide the isomorphism class of a given algebra. A node in the tree represents a conjunction, C, of properties. For each node, the system uses Mace to find two non-isomorphic algebras, m and m', which satisfy C. If two are found, the tree is extended by using HR to find properties that discriminate between m and m', otherwise the node becomes a leaf of the tree and Spass is used to verify that C is a classifying property.

As an example the above figure depicts the decision tree for the classification problem of order 3 quasigroups. The doubly-circled leaf nodes of the tree denote the single isomorphism classes. The property that uniquely characterises an isormorphism class is given as the conjunction of the single properties labelling the edges on the path from the root to a leaf node. Thus the classification theorem given by the decision tree states that a quasigroups of order 3 exhibits exactly one of the following properties: $\neg P_1$ or $P_1 \wedge P_2$ or $P_1 \wedge \neg P_2 \wedge P_3 \wedge P_4$ or $P_1 \wedge \neg P_2 \wedge \neg P_3 \wedge P_4$ or $P_1 \wedge \neg P_2 \wedge P_3 \wedge \neg P_4$.

We used both approaches to generate a number of classification theorems for groups, monoids, quasigroups and loops up to size 6. As an example, which highlights the power of this bootstrapping approach we used it to generate a classification theorem for the 109 isomorphism classes of loops of size 6. Not only does our approach highlight the utility of employing multiple reasoning systems for difficult tasks such as classification, our technique is neither restricted to the algebraic domain nor the isomorphism equivalence relation.

References

[1] L Kronecker. Auseinandersetzung einiger Eigenschaften der Klassenanzahl idealer komplexer Zahlen. *Monatsbericht der Berliner Akademie*, pages 881–889, 1870.

[2] W Feit and J Thompson. A solvability criterion for finite groups and some consequences. *Proc. Nat. Acad. Sci. USA*, 48:968–970, 1962.

[3] J Humphreys. *A Course in Group Theory.* Oxford University Press, 1996.