Journal of Integer Sequences, Vol. 21 (2018),
Article 18.1.7

Generation of Union-Closed Sets and Moore Families

Gunnar Brinkmann and Robin Deklerck
Applied Mathematics, Computer Science and Statistics
Ghent University
Krijgslaan 281 S9
B9000 Ghent
Belgium
gunnar.brinkmann@ugent.be
robindeklerck@skynet.be.

Abstract

We describe an algorithm to constructively enumerate non-isomorphic union-closed sets and Moore sets. We confirm the number of isomorphism classes of union-closed sets and Moore sets on $n \leq 6$ elements presented by other authors and give the number of isomorphism classes of union-closed sets and Moore sets on 7 elements. Due to the enormous growth of the number of isomorphism classes, it seems unlikely that constructive enumeration for 8 or more elements will be possible in the foreseeable future.

1 Introduction

All sets in this article are finite. A *union-closed set* is a set $\mathcal{U}$ of sets with the property that for all $A, B \in \mathcal{U}$ we have $A \cup B \in \mathcal{U}$. We call $\Omega_{\mathcal{U}} = \bigcup_{A \in \mathcal{U}} A$ the *universe* of $\mathcal{U}$. Two union-closed sets with universe $\Omega_{\mathcal{U}}$, resp., $\Omega_{\mathcal{U}'}$ are defined to be *isomorphic* if there is a bijective mapping $\Omega_{\mathcal{U}} \to \Omega_{\mathcal{U}'}$ inducing a bijection between the union-closed sets.

As we are mainly interested in isomorphism classes, we may assume $\Omega_U = \Omega_n = \{1, \ldots, n\}$ for some n. While the whole universe Ω_U is by definition an element of a union-closed set $\mathcal{U}$, this is not the case for the empty set. As the empty set has no impact on the structure of a union-closed set, one often either requires the empty set to be an element of each union-closed set or forbids it to be an element. We choose for the first convention, so our union-closed sets contain Ω_n as well as the empty set. We denote a set containing one representative of each isomorphism class of union-closed sets with universe Ω_n as $\mathcal{R}_n$.

The famous *union-closed sets conjecture* (or *Frankl's conjecture*) states that for each union-closed set there is an element that is present in at least half of the sets. Unfortunately, the number of union-closed sets grows so quickly that complete enumeration is not a suitable approach for testing this conjecture. A lot is known about the structure of possible counterexamples to the union-closed sets conjecture (see [2] for a survey), so any approach to extend the knowledge on the smallest size of a possible counterexample by constructive enumeration must focus on the subset of union-closed sets with those additional structural properties (e.g., with small average size of the sets, without some subconfigurations like singletons, etc.).

Union-closed sets are closely related to *Moore families*. A Moore family for a universe Ω_n is a set of subsets of Ω_n that is closed under intersection and contains Ω_n. It is easy to see that for a union-closed set $\mathcal{U}$ the set $\mathcal{U}^c = \{\Omega_n \setminus A | A \in \mathcal{U}\}$ is a Moore family. For a Moore family $\mathcal{M}$ the set $\mathcal{M}^c = \{\Omega_n \setminus A | A \in \mathcal{M}\}$ is closed under union, but as the empty set is not necessarily contained in $\mathcal{M}$, it is a union-closed set for $\Omega_n \setminus \bigcap_{A \in \mathcal{M}} A$, which is isomorphic to a union-closed set for some $\Omega_{n'}$ with $n' \leq n$.

A set $\mathcal{M}_n$ of representatives of Moore families (with the canonical definition of isomorphism) for the universe Ω_n can be obtained from sets $\mathcal{R}_0, \ldots, \mathcal{R}_n$ of representatives of union-closed sets containing the empty set as follows: $\mathcal{M}_n = \bigcup_{i=0}^{n} \{\mathcal{U}^c | \mathcal{U} \in \mathcal{R}_i\}$ if the complement is in each case taken in the universe Ω_n.

So union-closed sets correspond to Moore sets, which again characterize closure operators, and have many applications in topology, algebra and logic.

For $n < 7$, several authors developed enumeration algorithms for Moore families [4, 7, 8]. In the most advanced article, Colomb, Irlande, and Raynaud [4] not only counted Moore families for $n \leq 6$, but they also generated representatives of isomorphism classes. For $n = 7$, the approach is not suitable for generating a set of representatives, and by clever use of the structure of representatives of Moore families for $n = 6$ they only determined the number of labeled Moore families — that is, without considering isomorphisms. In our algorithm we determine the number of labeled union-closed sets resp., labeled Moore families for $n = 7$ from representatives and their automorphism groups of the union-closed sets for $n = 7$, resp., $n \leq 7$. This gives a very good independent test for the implementation in [4] as well as for our implementation. When computing the number of labeled Moore families for Ω_7 from the number of labeled union-closed sets with $n \leq 7$, for those union-closed sets with $n < 7$, one must take into account that for Ω_7 isomorphic copies not only occur for the universe $\{1, \ldots, n\}$, but for each n-element subset of $\{1, \ldots, 7\}$.

2 The algorithm

A subset $A \subseteq \Omega_n$ is represented as a number $b(A)$ given as the binary number $b_{n-1} \cdots b_0$ — possibly with leading zeros — with $b_i = 1$ if $(i+1) \in A$ and $b_i = 0$ otherwise.

We use an ordering of the subsets of Ω_n. For $A, B \subseteq \Omega_n$ we define $A > B$ if $|A| < |B|$ (so sets with more elements are considered smaller in this order) or $|A| = |B|$ and $b(A) > b(B)$. Whenever we refer to *larger* or *smaller* sets, we refer to this ordering.

The construction algorithm generates union-closed sets recursively based on the following easy lemma:

Lemma 1. *If $\mathcal{U} \neq \{\Omega_n\}$ is a union-closed set and A is the largest non-empty element in $\mathcal{U}$, then $\mathcal{U} \setminus \{A\}$ is also a union-closed set.*

This implies that union-closed sets for universe Ω_n can be recursively constructed from the union-closed set $\{\Omega_n, \emptyset\}$ of smallest size by successively adding subsets of Ω_n that are larger than the largest non-empty set already present. Of course it is not assured that adding a smaller set to a union-closed set does not violate the condition that the set must be closed under unions.

In order to turn this into an efficient algorithm, two tests that are (in principle) applied to each structure generated must be very fast:

(i) The test whether the set that has been constructed by adding a new set is closed under union.

(ii) The test for isomorphisms.

We first discuss (i). A straightforward way to test (i) for a union-closed set $\mathcal{U}$ to which a new set A is added would be to form all unions $A \cup B$ with $B \in \mathcal{U}$ and test whether they are in $\mathcal{U} \cup \{A\}$. Although all these steps can be implemented as efficient bit-operations, their number would slow down the program. We make the following definition.

Definition 2. For a union-closed set $\mathcal{U}$ we define the reduced set $\mathrm{red}(\mathcal{U})$ as follows:

$$\mathrm{red}(\mathcal{U}) = \{A \in \mathcal{U} | A \neq \emptyset \text{ and there is no } A_1 \neq \emptyset \text{ in } \mathcal{U}, A_1 \subsetneq A\}.$$

Lemma 3. *Let $\mathcal{U}$ be a union-closed set for a universe Ω_n and let $A \subset \Omega_n$, that is larger than any non-empty set in $\mathcal{U}$. Then $\mathcal{U} \cup \{A\}$ is closed under union if and only if $A \cup B \in \mathcal{U}$ for all $B \in$ red($\mathcal{U}$).*

Proof. First assume that $\mathcal{U} \cup \{A\}$ is closed under union and let $B \in \mathrm{red}(\mathcal{U})$. Then $A \cup B \in (\mathcal{U} \cup \{A\})$ and as A is larger than B, we have $A \cup B \neq A$, so $A \cup B \in \mathcal{U}$.

For the other direction assume that $A \cup B \in \mathcal{U}$ for all $B \in \mathrm{red}(\mathcal{U})$ and let $D \in \mathcal{U}$.

Choose any $D' \subset D$ so that $D' \in \mathrm{red}(\mathcal{U})$. Then $A' = A \cup D' \in \mathcal{U}$ and therefore also $A' \cup D \in \mathcal{U}$ as $\mathcal{U}$ is closed under union, but $A' \cup D = A \cup D$ — so $A \cup D \in \mathcal{U} \cup \{A\}$ and $\mathcal{U} \cup \{A\}$ is closed under union. $\square$

It would be inefficient to compute $\mathrm{red}(\mathcal{U})$ each time a new union-closed set is constructed, but as a new union-closed set $\mathcal{U}'$ is constructed by adding a new smallest element A to $\mathcal{U}$, the set $\mathrm{red}(\mathcal{U}')$ can easily be constructed from $\mathrm{red}(\mathcal{U})$ by adding A and removing elements that contain A. Nevertheless the few lines of code testing whether the potential union-closed set is closed under union take more than 50% of the total running time when computing union-closed sets for Ω_6, which is the largest case that can be profiled.

In order to solve problem (ii) efficiently — that is, avoid the generation of isomorphic copies — we use a combination of Read/Faradžev type orderly generation [6, 9] and the homomorphism principle (see, e.g., [1]).

Our first aim is to define a unique representative for each isomorphism class — called the canonical representative — and then only construct union-closed sets that are the canonical representatives of their class. We represent a union-closed set $\mathcal{U}$ with $k+1$ elements $A_1 < A_2 < \cdots < A_k < \emptyset$ as the string $s(\mathcal{U}) = b(A_1), \ldots, b(A_k)$ of numbers. For a given isomorphism class of union-closed sets for a universe Ω_n we choose the union-closed set with the lexicographically smallest string as the representative.

It is *in principle* easy to test whether a given union-closed set $\mathcal{U}$ is the representative of its class by applying all $n!$ possible permutations to $\mathcal{U}$ and comparing the strings. As $n \leq 7$ this would not be extremely expensive, but due to the large number of times that this test has to be applied, it is far too expensive to construct the union-closed sets for Ω_7. In the sequel we describe a way how this can be optimized.

In order to increase the efficiency by making it an orderly algorithm of Read/Faradžev type, we use the canonicity test not only for structures we output, but also during the construction: non-canonical structures are neither output nor used in the construction. This only leads to a correct algorithm if we can prove that canonical representatives are constructed from canonical representatives:

Theorem 4. *Let $\mathcal{U} \neq \{\Omega_n, \emptyset\}$ be a union-closed set for the universe Ω_n that is the canonical representative for its isomorphism class. If $\mathcal{U} = \{A_1, A_2, \ldots, A_k, \emptyset\}$ with $A_1 < A_2 < \cdots < A_k$ and $1 \leq m \leq k$, then $\{A_1, A_2, \ldots, A_m, \emptyset\}$ is also the canonical representative of its class.*

Proof. We prove the result for $m = k-1$. For $k = m$ it is the assumption and for $m < k-1$ it then follows by induction.

Let $s(\mathcal{U}) = (s_1, \ldots, s_k)$. For a permutation Π of Ω_n and a union-closed set $\mathcal{U}$ we write $\Pi(\mathcal{U})$ for the union-closed set obtained by replacing each element of a set in $\mathcal{U}$ by its image under Π. Assume that $\mathcal{U}' = \{A_1, A_2, \ldots, A_{k-1}, \emptyset\}$ is not the canonical representative of its class. So there is a permutation Π of Ω_n with $s(\Pi(\mathcal{U}')) < s(\mathcal{U}')$. Let $s(\Pi(\mathcal{U}')) = (p_1, \ldots, p_{k-1})$ and we have $s(\mathcal{U}') = (s_1, \ldots, s_{k-1})$. Let j be the first position so that $p_j < s_j$. Let us now look at $s(\Pi(\mathcal{U})) = (p'_1, \ldots, p'_k)$ and let r be the position of $\Pi(A_k)$ in this string. If $r > j$ then $p'_i = p_i = s_i$ for $1 \leq i < j$ and $p'_j = p_j < s_j$ — so there is a smaller representative for the isomorphism class of $\mathcal{U}$. If $r \leq j$ then $p'_i = p_i = s_i$ for $1 \leq i < r$ and $p'_r < p_r \leq s_r$ and again we have found a smaller representative contradicting the minimality of $s(\mathcal{U})$.

□

This theorem proves that the algorithm can reject non-canonical union-closed sets and is a correct orderly algorithm, but the cost of the canonicity test would make it impossible to determine the number of union-closed sets for Ω_7.

For a given union-closed set $\mathcal{U}$ with universe Ω_n and $1 \leq m \leq n$ we write $\mathcal{U}_m$ for the subset containing all sets of size $k \geq m$ and $\Pi_m(\mathcal{U})$ for all permutations Π of Ω_n with the property that $\Pi(\mathcal{U}_m) = \mathcal{U}_m$.

Lemma 5. *Let $\mathcal{U} \neq \{\Omega_n, \emptyset\}$ be a non-canonical union-closed set for the universe Ω_n with sets $A_1 < A_2 < \cdots < A_k < \emptyset$, so that $\{A_1, A_2, \ldots, A_{k-1}, \emptyset\}$ is canonical and $|A_k| = m$. Then all permutations Π with $s(\Pi(\mathcal{U})) < s(\mathcal{U})$ are in $\Pi_{m+1}(\mathcal{U})$.*

Proof. Any permutation Π not in $\Pi_{m+1}(\mathcal{U})$ would by definition give an isomorphic but different union-closed set $\Pi(\mathcal{U}_{m+1})$. As due to Theorem 4 $s(\mathcal{U}_{m+1})$ is minimal, $s(\Pi(\mathcal{U}_{m+1}))$ would be larger and therefore also the part of the string of $s(\Pi(\mathcal{U}))$ describing sets of size at least $m+1$ would imply $s(\Pi(\mathcal{U})) > s(\mathcal{U})$. □

Lemma 5 speeds up the canonicity test dramatically: We start with a list of all $n!$ permutations as $\Pi_n(\mathcal{U})$. When testing canonicity of a union-closed set with the smallest set of size $k < n$, we only apply permutations from $\Pi_{k+1}(\mathcal{U})$. During this application, we can already compute $\Pi_k(\mathcal{U})$ by simply adding exactly those permutations to the initially empty set $\Pi_k(\mathcal{U})$ that fix $\mathcal{U}$. As we work with small sets, it is no problem to store and use the set of all group elements instead of just a set of generators.

The impact is immediately clear: the number of permutations that has to be computed is much smaller and as soon as some $\Pi_k(\mathcal{U})$ contains only the identity — which happens very fast — no canonicity tests have to be performed, so that the total time for isomorphism checking is only about 7% of the total running time when computing union-closed sets for Ω_6.

2.1 The implementation

The algorithm was implemented in the computer language *C*. Next to an efficient algorithm, of course also implementation details are of crucial importance to be able to compute the results for Ω_7. We precomputed the action of all permutations on all sets, so that they could be applied very fast and used data structures that allow to check whether a set is contained in a union-closed set in constant time. Special functions were written that add sets with only one element. As no sets of smaller size are added, no updates of the automorphism groups are necessary and it turned out that at this stage it is also not efficient any more to remove sets from red($\mathcal{U}$) that are a superset of another element. Details on the implementation can best be seen in the code which can be obtained from the authors or the *Journal of Integer Sequences.*

2.2 Results

Tables 1 and 2 give the numbers of isomorphism classes of union-closed sets and Moore families as well as the numbers of labeled structures. Up to 5 elements the running times are less than 0.01 seconds. For $n = 6$ it is 8.2 seconds on a Xeon(R) CPU E5-2690 0 running with 2.90 GHz and a high load (which can make a difference for these processors). For $n = 7$ the jobs were run in parallel on different machines and some parts had to be divided further in order to finish, so it is hard to give precise times. Estimating the total running time from those parts that were run on the same machine used for $n = 6$, the total time on this type of machine should be around 10 to 12 CPU-years.

n	union-closed sets	labeled union-closed sets
1	1	1
2	3	4
3	14	45
4	165	2,271
5	14,480	1,373,701
6	108,281,182	75,965,474,236
7	2,796,163,091,470,050	14,087,647,703,920,103,947

Table 1: The number of union-closed sets (sequence A108798) and labeled union-closed sets (sequence A102894).

n	Moore families	labeled Moore families
1	2	2
2	5	7
3	19	61
4	184	2480
5	14,664	1,385,552
6	108,295,846	75,973,751,474
7	2,796,163,199,765,896	14,087,648,235,707,352,472

Table 2: The number of Moore families (sequence A193674) and labeled Moore families (sequence A102896).

A union-closed set on n elements is called *sparse* if the average number of elements in a set — not counting the empty set — is at most $\frac{n}{2}$. For union-closed sets that are not sparse, the union-closed sets conjecture is trivially true. Table 3 gives the number of sparse union-closed sets. These numbers were computed once by filtering all union-closed sets and once by an independent implementation using special bounding criteria described in [5]. Bruhn and Schaudt [3] give indications that sparse union-closed sets make up only a vanishingly small part of all union-closed sets. The numbers of sparse union-closed sets for $n \leq 7$ support this. Nevertheless also for sparse union closed sets the numbers seem to grow so fast that even with specialized algorithms complete enumeration will not be possible for sizes of the universe that are interesting for the union-closed sets conjecture.

n	sparse union-closed sets
1	0
2	0
3	0
4	2
5	27
6	3,133
7	5,777,931

Table 3: The number of sparse union-closed sets (sequence A299116).

2.3 Testing

The second author [5] developed an independent implementation of the algorithm together with special bounding criteria for sparse union-closed sets. The implementation was used to generate all isomorphism classes of union-closed sets for $\Omega_1, \ldots, \Omega_6$, and — using special bounding criteria — to confirm the numbers of isomorphism classes of sparse union-closed sets for Ω_7.

A further and independent confirmation of the numbers for $\Omega_1, \ldots, \Omega_6$ and also an independent confirmation for Ω_7 was obtained by computing the number of labeled structures corresponding to each representative from the size of the automorphism group. Note that as the size of the automorphism group is known in the algorithm anyway, the additional cost for this test can be neglected. From this we computed the number of (labeled) Moore families and got complete agreement with [4] for $\Omega_1, \ldots, \Omega_7$. Due to the completely different approaches this makes implementation errors leading to the same incorrect results in both cases extremely unlikely.

3 Acknowledgments

We thank Craig Larson for introducing us to these interesting structures and for interesting discussions. Furthermore, we thank Anastasiia Zharkova and Mikhail Abrosimov. The first version of this algorithm was intended as a hands-on example to illustrate isomorphism rejection techniques in a course they attended at Ghent University. Without their visit, this algorithm might not have been developed.

References

[1] G. Brinkmann, Isomorphism rejection in structure generation programs, in P. Hansen, P.W. Fowler, and M. Zheng, editors, *Discrete Mathematical Chemistry*, *DIMACS Series on Discrete Mathematics and Theoretical Computer Science* **51** (2000), 25–38.

[2] H. Bruhn and O. Schaudt, The journey of the union closed sets conjecture. *Graphs and Combinatorics* **31(6)** (2015), 2043–2074.

[3] H. Bruhn and O. Schaudt, The union-closed sets conjecture almost holds for almost all random bipartite graphs. *European J. Combinatorics* **59** (2017), 129–149.

[4] P. Colomb, A. Irlande, and O. Raynaud, Counting of Moore families for $n = 7$, in *Formal Concept Analysis*, Vol. 5986 of *Lecture Notes in Computer Science*, Springer, 2010, pp. 72–87.

[5] R. Deklerck, Een constructief algoritme voor union closed sets, Master's thesis, Ghent University, 2016.

[6] I. A. Faradžev, Constructive enumeration of combinatorial objects, *Colloques Internationaux C.N.R.S. No. 260 — Problèmes Combinatoires et Théorie des Graphes, Orsay*, 1976, pp. 131–135.

[7] M. Habib and L. Nourine, The number of Moore families on $n = 6$, *Discrete Math.* **294** (2005), 291–296.

[8] A. Higuchi, Lattices of closure operators, *Discrete Math.* **179** (1998), 267–272.

[9] R. C. Read, Every one a winner, *Ann. Discrete Math.* **2** (1978), 107–120.

2010 *Mathematics Subject Classification*: Primary 05-04; Secondary 05A15
Keywords: enumeration, Moore set, union-closed set.

(Concerned with sequences A102894, A102896, A108798, A193674, and A299116.)

Received November 20 2017; revised version received February 5 2018. Published in *Journal of Integer Sequences*, February 8 2018.

Return to Journal of Integer Sequences home page.