Theoretical Computer Science ••• (••••) •••–•••

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

An algebraic framework for minimum spanning tree problems

Walter Guttmann

Department of Computer Science and Software Engineering, University of Canterbury, New Zealand

A R T I C L E   I N F O

Article history:
Received 13 April 2017
Received in revised form 7 March 2018
Accepted 9 April 2018
Available online xxxx

Keywords:
Aggregation
Correctness proof
Formal methods
Kleene algebras
Relations
Stone algebras
Weighted graphs

A B S T R A C T

We formally prove that Prim's minimum spanning tree algorithm is correct for various optimisation problems with different aggregation functions. The original minimum weight spanning tree problem and the minimum bottleneck spanning tree problem are special cases, but the framework covers many other problems. To this end we work in new generalisations of relation algebras and Kleene algebras and in new algebraic structures that capture key operations used in Prim's algorithm and its specification. Weighted graphs form instances of these algebraic structures, where edge weights are taken from a linear order with a binary aggregation operation. Many existing results from relation algebras and Kleene algebras generalise from the relation model to the weighted-graph model with no or small changes. The overall structure of the proof uses Hoare logic. All results are formally verified in Isabelle/HOL heavily using its integrated automated theorem provers. This paper is an extended version of [36].

1. Introduction

A well-known algorithm commonly attributed to Prim [59] – and independently discovered by Jarník [41] and Dijkstra [25] – computes a minimum spanning tree in a weighted undirected graph. It starts with an arbitrary root node, and constructs a tree by repeatedly adding an edge that has minimal weight among the edges connecting a node in the tree with a node not in the tree. The iteration stops when there is no such edge, at which stage the constructed tree is a minimum spanning tree of the component of the graph that contains the root (which is the whole graph if it is connected).

The aim of this paper is to demonstrate the applicability of relation-algebraic methods for verifying the correctness of algorithms on weighted graphs. Accordingly, we will use an implementation of Prim's algorithm close to the above abstraction level. Since its discovery many efficient implementations of this and other spanning tree algorithms have been developed; for example, see the two surveys [32,50]. These implementations typically rely on specific data structures, which can be introduced into a high-level algorithm by means of data refinement; for example, see [8]. We do not pursue this in the present paper.

Relation-algebraic methods have been used to develop algorithms for unweighted graphs; for example, see [28,10,9,8]. This works well because such a graph can be directly represented as a relation; an adjacency matrix is a Boolean matrix. Weighted graphs do not have a direct representation as a binary relation. Previous relational approaches to weighted graphs therefore use many-sorted representations such as an incidence matrix and a weight function. In this paper, we directly work with a matrix of weights.

E-mail address: walter.guttmann@canterbury.ac.nz.

https://doi.org/10.1016/j.tcs.2018.04.012
0304-3975/

Please cite this article in press as: W. Guttmann, An algebraic framework for minimum spanning tree problems, Theoret. Comput. Sci. (2018), https://doi.org/10.1016/j.tcs.2018.04.012

In the context of fuzzy systems, relations have been generalised from Boolean matrices to matrices over the real interval $[0, 1]$ or over arbitrary complete distributive lattices [29]. The underlying idea is to extend qualitative to quantitative methods; see [57] for another instance based on automata. We propose to use matrices over lattices to model weighted graphs, in particular in graph algorithms. Previous work based on semirings and Kleene algebras deals well with path problems in graphs [30]. We combine these algebras with generalisations of relation algebras to tackle the minimum spanning tree problem.

Tarski's relation algebras [63], which capture Boolean matrices, have been generalised to Dedekind categories to algebraically capture fuzzy relations [44]; these categories are also known as locally complete division allegories [27]. In the present paper we introduce a new generalisation – Stone relation algebras – which maintains the signature of relation algebras and weakens the underlying Boolean algebra structure to Stone algebras. Stone algebras have a pseudocomplement operation, which satisfies some properties of a complement but not all. This weakening is needed since complementation does not extend from two to more elements in linear orders, in this case, the linear order of edge weights. We show that matrices over bounded linear orders are instances of Stone relation algebras and of Kleene algebras, and can be used to represent weighted graphs.

Most of the correctness proof of Prim's minimum spanning tree algorithm can be carried out in these general algebras. Therefore, most of our results hold for many instances, not just weighted graphs. A small part of the correctness proof uses operations beyond those available in relation algebras and in Kleene algebras, namely for summing edge weights and identifying minimal edges. We introduce new algebraic structures – m-algebras – that capture essential properties of these operations, which are used in Prim's algorithm and in its specification.

Because the correctness proof of Prim's algorithm relies only on the axioms of m-algebras, we can construct different instances of these algebras to solve various problems. The instances are based on bounded linear orders with a binary aggregation operation. Applied to all entries of a weighted-graph matrix, aggregation describes the value that is minimised by Prim's algorithm. The minimum weight spanning tree problem arises as the instance where aggregation computes the sum of edge weights (say, real numbers). For the minimum bottleneck spanning tree problem aggregation amounts to the maximum of edge weights [18]. Covering these and further optimisation problems, we propose axioms for the aggregation operation based on which Prim's algorithm will work correctly. With this algebraic method, each particular problem arises as a specific instance of the involved algebras. This is analogous to the generalisation of Warshall's transitive closure algorithm [64] and Floyd's all-pairs shortest paths algorithm [26] to a general dynamic-programming algorithm based on semirings, which covers many other instances [2].

With this approach we can apply well-developed methods and concepts of relation algebras and Kleene algebras to reason about weighted graphs in a new, more direct way. The contributions of this paper are:

- Stone relation algebras, a new algebraic structure that generalises relation algebras but maintains their signature (Section 2). Many theorems of relation algebras already hold in these weaker algebras, which we use to represent weighted graphs. We combine them with Kleene algebras to describe reachability in graphs (Section 3). This yields a general yet expressive setting for most of the correctness proof of the minimum spanning tree algorithm.
- M-algebras, a new algebraic structure that extends Stone–Kleene relation algebras by dedicated operations and axioms for finding minimal edges and for computing the total weight of a graph (Section 4).
- Models of the above algebras, including weighted graphs represented by matrices over bounded linear orders (Sections 2, 3 and 5). This includes a formal verification of Conway's automata-based construction for the Kleene star of a matrix.
- Linear aggregation lattices, a new algebraic structure based on bounded linear orders with a binary aggregation operation (Section 5). Matrices over linear aggregation lattices form m-algebras.
- Several classes of instances of linear aggregation lattices covering, in particular, the minimum weight spanning tree and minimum bottleneck spanning tree problems (Section 6). Aggregations based on triangular norms and conorms also arise as special cases.
- A Hoare-logic correctness proof of Prim's minimum spanning tree algorithm entirely based on the above algebras (Section 7).
- Isabelle/HOL theories that formally verify all of the above and all results in and about the algebras stated in the present paper. Proofs are omitted in this paper and can be found in the Isabelle/HOL theory files available at http://www.csse.canterbury.ac.nz/walter.guttmann/algebra/. The Archive of Formal Proofs at https://www.isa-afp.org/ currently holds theories with all results of Sections 2 and 3; the other theories are being prepared for it.

Related work is discussed in Section 8.

The present paper is an extended version of [36]. We have generalised the results of Section 2 from extended real numbers to linear bounded orders. Major changes concern Section 4 and the two new Sections 5 and 6. In Section 4, we have revised Definition 6 and simplified the axioms for s-algebras and m-algebras taking into account new results about Stone relation algebras [37]. We have also added Theorem 4. Section 5 is entirely new. It motivates how different kinds of aggregation amount to solving different optimisation problems. We propose various algebraic structures based on orders with a binary aggregation operation. We show how to extend binary aggregation to finite sums despite the absence of a unit. We define general summation and minimisation operations on matrices based on the binary aggregation, and

Please cite this article in press as: W. Guttmann, An algebraic framework for minimum spanning tree problems, Theoret. Comput. Sci. (2018), https://doi.org/10.1016/j.tcs.2018.04.012