Algebraic Graphs with Class

Christoph Madlener
`madlener@in.tum.de`

May 31, 2021

Abstract

We give an overview of algebraic graphs as introduced by Mokhov [9]. Algebraic graphs have a small and safe core of construction primitives, alongside a set of axioms, characterizing an algebra of graphs. They can be used to elegantly implement a graph transformation library employing functional programming. We also review the suitability of algebraic graphs for formal verification using Isabelle/HOL.

1 Introduction

Graphs are a fundamental structure studied in depth by both mathematicians and computer scientists alike. One very common definition states that a (directed) graph is a tuple $G = (V, E)$, where V is a set of vertices and $E \subseteq V \times V$ is the set of edges. While this is a perfectly valid and natural mathematical definition it is not necessarily suitable for implementation. This can be illustrated by trying to directly translate this to Haskell:

```haskell
data G a = G { vertices :: Set a, edges :: Set (a,a)}
```

The value `G [1,2,3] [(1,2),(2,3)]` then represents the graph $G = (\{1, 2, 3\}, \{(1, 2), (2, 3)\})$, however `G [1,2] [(2,3)]` does not represent a consistent graph, as the edge refers to a non-existent node. The state-of-the-art `containers` library implements graphs with adjacency lists employing immutable arrays [7]. The *consistency condition* $E \subseteq V \times V$ is not checked statically, which can lead to runtime errors. `fgl`, another popular Haskell graph library uses inductive graphs [4], also exhibiting partial functions and potential for runtime errors due to the violation of consistency.

This led Mokhov to conceive *algebraic graphs* [9], a sound and complete representation for graphs. They abstract away from graph representation details and characterize graphs by a set of axioms. Algebraic graphs have a small safe core of graph construction primitives and are suitable for implementing graph transformations. The four construction primitives are the following:

1. the *empty* graph

2. single *vertex* graphs

3. *overlay*ing two graphs

4. *connect*ing two graphs

Overlay is the union of two graphs while connect also adds edges from all vertices from one graph to all vertices of the other.

Alongside proper definitions for these primitives we will see in section 2 that this is indeed a sound and complete graph representation. We will also cover the algebraic structure (2.2) exhibited by these primitives and elegant graph transformations (2.3) based on a type class for the core. In section 2.4 we explore how we can exploit the algebraic structure in verification using Isabelle/HOL.

In section 3 we will examine the *deep embedding* of the construction primitives into an inductive datatype and its applicability in practice and in verification (3.1).

In section 4 we conclude and give an outlook on possible future work and open questions connected to algebraic graphs as presented.

2 A Type Class for Algebraic Graphs

This section will cover most parts of the original paper by Mokhov [9]. The graph construction primitives of algebraic graphs will be formally defined and their soundness and completeness proven. Afterwards we will cover the algebraic structure exhibited by these primitives when interpreted as graphs. As part of this we will introduce a type class in Haskell which further abstracts from implementation details. We will also see that different classes of graphs can be obtained by modifying the axioms. Instances of these different classes and a graph transformation library in Haskell are presented. To conclude this section we will explore the suitability of this type class for formal verification in Isabelle/HOL.

2.1 The Core

Before defining the four graph construction primitives we define the set of graphs $\mathcal{G}$ over a fixed universe of vertices $\mathbb{V}$. A graph $G \in \mathcal{G}$ can be represented as a tuple $G = (V, E)$ where $V \subseteq \mathbb{V}$. A graph is *consistent* if $E \subseteq V \times V$ holds.

The construction primitives then are defined as follows. The *empty* graph, denoted by ε, is the tuple $(\emptyset, \emptyset)$ which is clearly consistent, hence $\varepsilon \in \mathcal{G}$. Graphs with a single *vertex* $v \in \mathbb{V}$ and no edges, i.e. $(\{v\}, \emptyset)$, denoted by just v, are also trivially consistent, hence $v \in \mathcal{G}$. The binary operations *overlay* and *connect*, denoted by $\oplus$ and $\rightarrow$ respectively, allow constructing larger graphs. Their definitions are

$$(V_1, E_1) \oplus (V_2, E_2) := (V_1 \cup V_2, E_1 \cup E_2)$$
$$(V_1, E_1) \rightarrow (V_2, E_2) := (V_1 \cup V_2, E_1 \cup E_2 \cup V_1 \times V_2)$$

Overlay is just the union of two graphs, while connect also adds an edge from each vertex of its first argument to each vertex of its second argument. It is straightforward to see that if $G_1 \in \mathcal{G}$ and $G_2 \in \mathcal{G}$ are consistent, then so are $G_1 \oplus G_2 \in \mathcal{G}$ and $G_1 \rightarrow G_2 \in \mathcal{G}$. In combination with the consistency of ε and of single vertex graphs this already proves that algebraic graphs are a consistent representation for graphs, i.e. the construction primitives only yield consistent graphs. The completeness of this representation, i.e. for each $G \in \mathcal{G}$ there is an algebraic graph expression representing it, is shown shortly (2.2). Both of these facts were also formally proven in Isabelle/HOL (see section 2.4).

As we will shortly see in section 2.2 overlay and connect are somewhat related to addition and multiplication in a semiring, hence, we give connect a higher precedence than overlay. Let us consider a few examples to develop some intuition for the primitives. In these examples the universe of vertices are the natural numbers and they are illustrated in Figure 1.

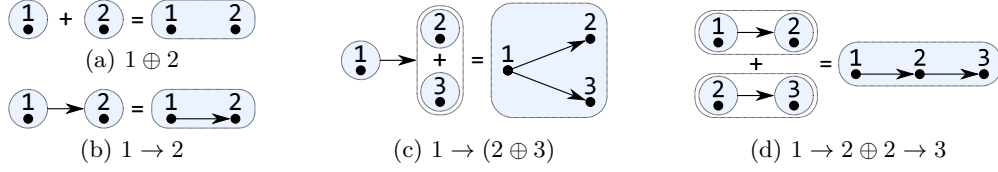(a) $1 \oplus 2 = (\{1, 2\}, \emptyset)$

Figure 1: Examples of graph construction. The overlay and connect operations are denoted by $\oplus$ and $\rightarrow$, respectively. Illustrations from [9].

(b) $1 \rightarrow 2 = (\{1, 2\}, \{(1, 2)\})$

(c) $1 \rightarrow (2 \oplus 3) = (\{1, 2, 3\}, \{(1, 2), (1, 3)\})$

(d) $1 \rightarrow 2 \oplus 2 \rightarrow 3 = (\{1, 2, 3\}, \{(1, 2), (2, 3)\})$

We model the graph construction primitives in Haskell with the following type class:

```haskell
class Graph g where
    type Vertex g
    empty :: g
    vertex :: Vertex g -> g
    overlay :: g -> g -> g
    connect :: g -> g -> g
```

The associated type `Vertex` g represents the universe of vertices $\mathbb{V}$, the remainder of the type class corresponds to ε, single vertex graphs, $\oplus$ and $\rightarrow$ respectively.

At this point we will only introduce some very basic functions for constructing graphs. More advanced constructions and transformations will be given in section 2.3. A graph with a single edge is obtained by simply connecting two vertices:

```haskell
edge :: Graph g => Vertex g -> Vertex g -> g
edge u v = (vertex u) `connect` (vertex v)
```

A graph with only isolated vertices can be constructed from a list of vertices as follows:

```haskell
vertices :: Graph g => [Vertex g] -> g
vertices = foldr overlay empty . map vertex
```

Replacing `overlay` with `connect` in `vertices` leads to a (directed) clique. Later we will also consider undirected graphs. In that context this function actually builds a fully connected graph on the given list of vertices.

```haskell
clique :: Graph g => [Vertex g] -> g
clique = foldr connect empty . map vertex
```

Following and extending this approach leads to a safe and fully polymorphic graph transformation library for any concrete graph representation (e.g. graphs in `containers` or `fgl`) instantiating this class.

2.2 Algebraic Structure

Before turning to more advanced graph constructions we will observe the algebraic structure of the primitives. Overlay and connect form an algebra abiding the following axioms:
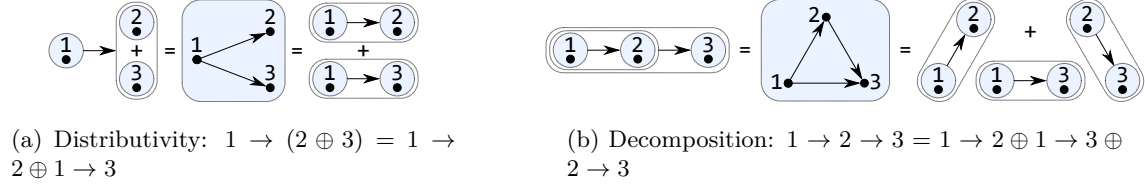
3

(a) Distributivity: $1 \rightarrow (2 \oplus 3) = 1 \rightarrow 2 \oplus 1 \rightarrow 3$

(b) Decomposition: $1 \rightarrow 2 \rightarrow 3 = 1 \rightarrow 2 \oplus 1 \rightarrow 3 \oplus 2 \rightarrow 3$

Figure 2: Two axioms of the algebra of graphs. Illustrations from [9].

- $\oplus$ is commutative and associative
- $(\mathcal{G}, \rightarrow, \varepsilon)$ is a monoid
- $\rightarrow$ distributes over $\oplus$

In addition there is the following *decomposition law*, as illustrated in Figure 2(b).

$$x \rightarrow y \rightarrow z = x \rightarrow y \oplus x \rightarrow z \oplus y \rightarrow z$$

With decomposition we can prove by equational reasoning that $\oplus$ is idempotent and that ε is its identity. Thus, overlay and connect almost form a semiring over graphs:

- $(\mathcal{G}, \oplus, \varepsilon)$ is an idempotent, commutative monoid
- $(\mathcal{G}, \rightarrow, \varepsilon)$ is a monoid
- $\rightarrow$ distributes over $\oplus$

For a proper semiring the annihilating zero for $\rightarrow$ (i.e. $x \rightarrow \varepsilon = \varepsilon = \varepsilon \rightarrow x$) is missing. It is also unusual that the two operations share their identity ε.

According to Mokhov, the stated axioms are a minimal set of axioms *characterizing* algebraic graphs [9]. From the original paper it does not become clear what he means by characterization: Letting $\mathcal{G} = \{1\}$, $\varepsilon = 1$, $1 \oplus 1 = 1$ and $1 \rightarrow 1 = 1$ gives a structure abiding all axioms. This definition of $\mathcal{G}$ corresponds to graphs with an empty universe of vertices (as there is only the empty graph). However, this definition is independent of the underlying universe of vertices. So we can define the same structure for a non-empty universe. In this scenario the primitives don't behave like what we call (isomorphic to) graphs. We don't know which additional axioms are required to obtain this isomorphism. However, the statement $\varepsilon \neq vertex\ u$ for any $u \in \mathbb{V}$ does not follow from the given axioms. Hence, adding an axiom scheme $\forall u \in \mathbb{V} : \varepsilon \neq vertex\ u$ seems sensible. Additionally *vertex* needs to be injective (i.e. $vertex\ u = vertex\ v \implies u = v$).

One can, however, easily verify that the definitions of the primitives given in section 2.1 satisfy these axioms. At this point let us consider the completeness of algebraic graphs. For any graph $G = (V, E)$ the following function builds it using the four construction primitives:

```
graph :: Graph g => [Vertex g] -> [(Vertex g, Vertex g)] -> g
graph vs es = overlay (vertices vs) (edges es)
```

The `edges` function generalizes `edge` to a list of edges:

```
edges :: Graph g => [(Vertex g, Vertex g)] -> g
edges = foldr overlay empty . map (uncurry edge)
```

4

They are a direct translation of the fact that any algebraic graph expression g for a graph $G = (V, E) \in \mathcal{G}$ can be rewritten in the *canonical form*:

$$g = \left(\bigoplus_{v \in V} v \right) \oplus \left(\bigoplus_{(u,v) \in E} u \to v \right)$$

Mokhov also follows a standard approach of defining a partial order ($\preceq$) for the idempotent overlay operation as $x \preceq y :\Leftrightarrow x \oplus y = y$. This is indeed a partial order under the graph axioms (i.e. it is reflexive, antisymmetric and transitive). In fact this partial order can be used as a definition for the subgraph relation, denoted by $\subseteq$:

$$x \subseteq y :\Leftrightarrow x \preceq y :\Leftrightarrow x \oplus y = y$$

So far we only considered directed graphs. Undirected graphs can be obtained from the very same construction primitives as directed graphs. It suffices to modify the underlying axioms. In the case of undirected graphs this can be achieved by making connect commutative (then denoted by $\leftrightarrow$), i.e. adding the axiom $x \leftrightarrow y = y \leftrightarrow x$. By introducing this axiom we can prove some of the other axioms of directed graphs. Hence, according to Mokhov [9], undirected graphs are characterized by the following minimal set of axioms:

- $\oplus$ is commutative and associative

- $\leftrightarrow$ is commutative and has ε as the identity

- left distributivity: $x \leftrightarrow (y \oplus z) = x \leftrightarrow y \oplus x \leftrightarrow z$

- left decomposition: $(x \leftrightarrow y) \leftrightarrow z = x \leftrightarrow y \oplus x \leftrightarrow z \oplus y \leftrightarrow z$

By choosing different additional axioms, other classes of graphs can be characterized. Some examples are reflexive graphs (i.e. each vertex has a self-loop), transitive graphs (as used in dependency graphs), and the combination thereof; preorders. Combining undirected graphs and reflexive graphs yields equivalence relations. It is even possible to define hypergraphs of arbitrary (but fixed) rank by replacing the decomposition law with another suitable axiom [9]. In all of these the same caveats regarding the term *characterization* from directed graphs apply.

In summary the algebraic approach to graph representation has a small and safe core of construction primitives. Pairing these primitives with suitable sets of axioms leads to a highly flexible representation for graphs.

2.3 Graph Transformation Library

Based on the core type class, Mokhov implements what he refers to as a graph transformation library[1]. In order to understand the following constructions and transformations better, we follow his approach and first define instances for the type class **Graph**. Recall the attempt from the introduction to directly translate the graph representation $G = (V, E)$ into the Haskell datatype **G**. We will now define a **Graph** instance for that datatype, while realizing that more generally it can be used to represent relations on some domain (hence the renaming to **Relation**).

```
data Relation a = R { domain :: Set a, relation :: Set (a, a)} deriving Eq

instance Ord a => Graph (Relation a) where
```

---

[1] Algebraic graphs on hackage: `http://hackage.haskell.org/package/algebraic-graphs`

```
type Vertex (Relation a) = a
empty        = R Set.empty Set.empty
vertex  x    = R (singleton x) Set.empty
overlay x y = R (domain x `union` domain y) (relation x `union` relation y)
connect x y = R (domain x `union` domain y) (relation x `union` relation y `union`
    fromAscList [ (a, b) | a <- elems (domain x), b <- elems (domain y) ])
```

Mokhov also defines a `Num` instance for `Relation` which enables the use of + and ∗ as shortcuts for `overlay` and `connect`.

The given `Graph` instance for `Relation` satisfies the axioms of directed graphs, as was argued in section 2.1. Obtaining `Graph` instances which satisfy the axioms for undirected graphs (or other classes of graphs) is also straightforward: One can simply wrap `Relation` in a `newtype` and define a custom `Eq` instance. An example for undirected graphs (in this case the connection to symmetric directed graphs is very prominent) is given.

```
newtype Symmetric a = S (Relation a) deriving (Graph, Num)

instance Ord a => Eq (Symmetric a) where
  S x == S y = symmetricClosure x == symmetricClosure y
```

It is also sensible to introduce subclasses like `class Graph g => UndirectedGraph g` to reflect the class of graph we are working with in the type system, and thus, increase type safety.

Equipped with these two instances let us now actually consider graph constructions and transformations made possible with the core type class. Simple examples include constructing a *path* from a list of vertices, which then can be easily extended to a *circuit*:

```
path :: Graph g => [Vertex g] -> g
path []  = empty
path [x] = vertex x
path xs = edges $ zip xs (tail xs)

circuit :: Graph g => [Vertex g] -> g
circuit [] = empty
circuit xs = path (xs ++ [head xs])
```

Mokhov also gives functions for constructing *star* graphs (i.e. a center vertex connected to a set of leaves) and for building graphs from `Tree`s and `Forest`s from the `containers` library.

More interesting applications include the implementation of a zero time graph transposition using yet another `newtype` wrapper [9]:

```
newtype Transpose g = T { transpose :: g } deriving Eq

instance Graph g => Graph (Transpose g) where
  type Vertex (Transpose g) = Vertex g
  empty        = T empty
  vertex       = T . vertex
  overlay x y = T $ overlay (transpose x) (transpose y)
  connect x y = T $ connect (transpose y) (transpose x)
```

$\mathcal{O}(1)$ transposition can be achieved by representations which allow to use a flag to treat edges as transposed. For many typical representations (e.g. adjacency lists) the whole graph has to be traversed during transposition, resulting in an $\mathcal{O}(|V| + |E|)$ time operation.

More sophisticated transformations can be implemented using the following `Functor`-like `newtype` wrapper. The function `gmap` takes a function of type `a -> b` and a graph with vertices of type `a` and produces a graph with vertices of type `b`.

```haskell
newtype GraphFunctor a = F { gfor :: forall g. Graph g => (a -> Vertex g) -> g }

instance Graph (GraphFunctor a) where
  type Vertex (GraphFunctor a) = a
  empty       = F $ \_ -> empty
  vertex  x   = F $ \f -> vertex (f x)
  overlay x y = F $ \f -> overlay (gmap f x) (gmap f y)
  connect x y = F $ \f -> connect (gmap f x) (gmap f y)

gmap :: Graph g => (a -> Vertex g) -> GraphFunctor a -> g
gmap = flip gfor
```

Interestingly this can be used to contract vertices fulfilling some predicate `p` by mapping them to a single vertex:

```haskell
mergeVertices :: Graph g => (Vertex g -> Bool)
  -> Vertex g -> GraphFunctor (Vertex g) -> g
mergeVertices p v = gmap $ \u -> if p u then v else u
```

In the original paper `gmap` is then used further to construct rather sophisticated graphs fully polymorphically [9], i.e. independent from any concrete graph representation.

Mokhov also goes on to introduce a `Monad`-like `newtype` wrapper. The `bind` function in the context of graphs allows to replace each vertex with a (possibly empty) subgraph. This can be used to both remove and split vertices.

```haskell
newtype GraphMonad a = M { bind :: forall g. Graph g => (a -> g) -> g }

instance Graph (GraphMonad a) where
  type Vertex (GraphMonad a) = a
  empty       = M $ \_ -> empty
  vertex x    = M $ \f -> f x
  overlay x y = M $ \f -> overlay (bind x f) (bind y f)
  connect x y = M $ \f -> connect (bind x f) (bind y f)
```

Removing a vertex with `bind` can be generalized to a function `induce`, which only keeps vertices satisfying some predicate `p`.

```haskell
induce :: Graph g => (Vertex g -> Bool) -> GraphMonad (Vertex g) -> g
induce p g = bind g $ \v -> if p v then vertex v else empty
```

Splitting a vertex works by replacing that vertex with a graph containing the desired vertices.

```haskell
splitVertex :: (Graph g, Eq (Vertex g)) => Vertex g -> [Vertex g]
  -> GraphMonad (Vertex g) -> g
splitVertex v vs g = bind g $ \u -> if u == v then vertices vs else vertex u
```

Mokhov then defines yet another `newtype` wrapper utilizing `bind` to remove edges. He also showcases the construction capabilities of the presented library by constructing de Bruijn graphs

(a rather sophisticated type of graph occurring frequently in computer engineering and bioinformatics) completely polymorphically. These definitions are omitted here for brevity.

The presented material on the given type class gives an overview of an elegant, reusable and polymorphic graph construction and transformation library. Note that in the available `hackage` library for algebraic graphs the `newtype` wrappers for graph transformation are not implemented as such. Instead there is a higher-kinded type class which is based on `MonadPlus`, but has fewer instances [9].

2.4 Verification

We have now seen how to implement a polymorphic graph transformation library in Haskell using the core type class. Mokhov proposes that algebraic graphs and their axiomatic characterization are suitable for formal verification as they enable equational reasoning. Some simple statements (like `vertices xs` $\subseteq$ `clique xs`) are given accompanied by an Agda formalization[2]. For the examples in that formalization the claim obviously holds.

To experiment further we first formalized the material from the paper in Isabelle/HOL [11]. This formalization[3] includes several *locale*s. For the purposes of this paper a locale can be thought of as a type class which can be equipped with axioms. The following locale is the basis for our formalization.

```
locale algebraic_pre_graph =
  fixes empty :: 'g ("ε")
    and vertex :: "'v ⇒ 'g"
    and overlay :: "'g ⇒ 'g ⇒ 'g" (infixl ⟨⊕⟩ 75)
    and connect :: "'g ⇒ 'g ⇒ 'g" (infixl ⟨→⟩ 80)
```

Note that it does not have any axioms, as we reuse it to define locales for different classes of graphs (i.e. directed, undirected, etc.). Proving the statements from the paper in Isabelle/HOL works without problems.

Noschinski formalized fundamental aspects of graph theory in Isabelle/HOL [12]. Part of that is essentially the standard representation of a graph as a tuple $G = (V, E)$ (i.e. a locale with one single axiom ensuring consistency). Instantiating a locale in Isabelle/HOL (which can be likened to instantiating a type class in Haskell) requires to prove the axioms attached to it. Using the definitions from section 2.1 for the primitives it is straightforward to instantiate the algebraic graph locale for the tuple representation. We also formally prove that the graph construction primitives only produce consistent graphs. A proof for the completeness of the primitives is given as well (in the sense that every graph represented by a tuple can be obtained using the algebraic graph construction primitives).

As the statements in the original paper were of very limited scope, we define and reason about more *advanced* graph theoretic concepts: We introduce (vertex) walks and the notion of reachability on algebraic graphs. A challenge posed by the approach is that a graph is simply a value of some type `'g`. Thus, we characterize the existence of a vertex or an edge in a given graph by the subgraph relation. A graph g contains a vertex u if and only if $u \subseteq g$ and an edge (u, v) if and only if $u \to v \subseteq g$. The Isabelle definitions for this look as follows:

```
definition has_vertex :: "'g ⇒ 'v ⇒ bool" where
  "has_vertex g u ≡ vertex u ⊑ g"
```

---

[2]available at `https://github.com/snowleopard/alga-theory`
[3]available at `https://github.com/cmadlener/algebraic_graphs`

```
definition has_edge :: "'g ⇒ 'v ⇒ 'v ⇒ bool" where
  "has_edge g u v ≡ edge u v ⊑ g"
```

Where ⊑ is used instead of ⊆ and `edge u v` is defined as `connect (vertex u) (vertex v)`. Equipped with these definitions we can then go on to inductively define when a list of vertices is a walk in a given graph:

```
inductive vwalk where
  vwalk_Nil: "vwalk g []" |
  vwalk_single: "has_vertex g u ⟹ vwalk g [u]" |
  vwalk_Cons: "vwalk g (v#xs) ⟹ has_edge g u v ⟹ vwalk g (u#v#xs)"
```

We define the empty list as well as the list containing only a single vertex of the graph as walks (`vwalk_Nil` and `vwalk_single`). Additionally if a list with head v (note that Haskell's (`:`) is denoted by (`#`) in Isabelle) is a walk and $(u,v) \in E$, then prepending u also yields a walk. Now we can prove various statements about walks (e.g. under which conditions walks can be appended, how walks can be split, etc.). Interestingly, many of these statements follow directly from the inductive definition of walks, i.e. we don't require any of the algebraic graph axioms to prove them. We only need them when a proof relies on the fact that $(u,v) \in E$ implies $u \in V$ as well as $v \in V$ (in other words, the consistency of a graph).

For reachability we took two approaches. First we defined the set of edges of a graph as follows:

```
definition edge_set :: "'g ⇒ ('v × 'v) set" where
  "edge_set g ≡ {(u,v) | u v. has_edge g u v}"
```

Reachability then can be defined via the reflexive, transitive closure of the edge set, while also considering every vertex can reach itself. The second approach is to say that v is reachable from u if and only if there is a walk from u to v. These are, however, only proof of concepts as only transitivity of reachability has been shown for both definitions.

Nonetheless, first impressions from these experiments suggest that it is feasible to base more sophisticated formalizations of graph theory on algebraic graphs. We did not experiment with adding additional axioms as described in section 2.2. We do give an instantiation of the locale with the unit type, illustrating that the current axioms are not sufficient.

Exploring verification using the locale further is an interesting direction for future work. The Haskell type class forms the basis for a polymorphic graph transformation library. In the same vein, the locale can be seen as a *polymorphic* graph formalization library. Any lemma we can prove using the locale axioms becomes available to any graph representation instantiating the locale. This obviously becomes more interesting the more material is available in the locale. Haslbeck et al. chose a similar approach to implement and verify Kruskal's algorithm independent of the concrete graph representation [6].

3 Deep Embedding

Up until this point we were mainly concerned with exploiting the algebraic structure of graphs for polymorphic graph constructions and transformations, and for formal verification. Deeply embedding the construction primitives in a datatype is interesting in itself. The definition of the datatype is straightforward:

```haskell
data Graph a = Empty
             | Vertex a
             | Overlay (Graph a) (Graph a)
             | Connect (Graph a) (Graph a)
```

An instance for the type class can be defined easily, however, we cannot use the derived `Eq` instance for the datatype, as it would violate the graph axioms: `Overlay Empty Empty` and `Empty` both represent ε but are not deemed equal. One solution is to fold the deep embedding into a concrete graph representation to obtain a law-abiding `Eq` instance:

```haskell
fold :: Graph g => Graph (Vertex g) -> g
fold Empty        = empty
fold (Vertex  x  ) = vertex x
fold (Overlay x y) = overlay (fold x) (fold y)
fold (Connect x y) = connect (fold x) (fold y)


instance Ord a => Eq (Graph a) where
  x == y = fold x == (fold y :: Relation a)
```

This approach is also currently implemented in the library available on `hackage` (although an adjacency map is used instead of a relation). However, this is not ideal for multiple reasons. Coupling the datatype to some other graph representation merely for equality checks is undesirable.

More interestingly algebraic graphs can compactly represent dense graphs. Consider the graph `clique [1..n]`, which has a linear size representation in the deep embedding, while it is quadratic when stored as a `Relation`. Requiring to fully expand graphs to check equality voids this advantage. A promising approach to alleviate this issue is *modular graph decomposition* [8], as it allows to minimize algebraic graph expressions.

More interesting questions for future work come up in connection with the deep embedding. Currently only the transformations presented with the type class, i.e. transpose, as well as anything that can be done using the corresponding versions of `gmap` and `bind`, are implemented directly on the deep embedding. No efficient implementations of fundamental graph algorithms - like depth-first-search - that work directly on the deep embedding are known [9]. The current library implementation therefore resorts to converting algebraic graph expressions to an adjacency map and reuses existing algorithms on this representation.

One application that comes to mind is search. Consider a best case example, where only a small part of a dense graph has to be expanded before finding a target: This allows us to exploit the better space complexity (linear vs. quadratic) while also being competitive regarding time. The neighborhood of a vertex in an algebraic graph can be computed in linear (in the size of the graph expression) time and space.

Benchmarks (on an earlier version of the library) were performed by Moine to compare the performance of the deep embedding to `containers` and `fgl` (and one other representation). The results vary heavily depending on the area of application. In some cases algebraic graphs offer a competitive alternative, while in others, they perform orders of magnitudes worse. For the full results refer to the repository containing the benchmark suite[4].

3.1 Verification

We were also interested in using the deep embedding in formal verification using Isabelle/HOL. Hence, we ported the functions to Isabelle/HOL and then proved their correctness. This turned

---

[4]available at `https://github.com/haskell-perf/graphs`

out to be very convenient, as the inductive datatype definition of the deep embedding enables proofs by induction. Many of the resulting subgoals were then handled automatically by Isabelle.

In the Haskell implementation of the deep embedding we already saw that structural equality does not satisfy the axioms. There we can define a custom `Eq` instance to obtain a law-abiding datatype. In Isabelle/HOL we can follow a similar approach, which results in a `quotient_type`, which can be thought of as an equivalence class construction on the raw datatype. For the moment the equivalence is based on conversion of algebraic graph expressions to the tuple representation by Noschinski mentioned in section 2.4. Again, the minimization of algebraic graph expressions to obtain an equivalence relation on algebraic graphs, decoupled from any other graph representation, is very much desirable.

The quotient type also allows us to formally prove that it satisfies the axioms, by instantiating the locale corresponding to the core type class. It also paves the way for further formalization efforts on the deep embedding. However, this also requires more work to become truly interesting, i.e. the discovery of algorithms that work directly on the deep embedding or other advanced usages.

4 Conclusion

We saw a novel approach to representing graphs, which is especially suitable for functional programming. It allows for safe, polymorphic and elegant graph constructions and transformations. Algebraic graphs also lend themselves nicely to formal verification.

The alert reader might have noticed that the given primitives can only be used to represent unlabeled graphs. Mokhov extended the algebraic graph library to also represent edge-labeled graphs. There is no published material on this yet (besides the source code). For more information refer to the talk by Mokhov during Haskell eXchange 2018[5].

Algebraic graphs as presented have been successfully used in practice, e.g. for asynchronous circuit design [1] and acceleration of graph processing using FPGAs [10]. In these examples the available transformations are sufficient. An important direction for future research is going beyond those (relatively simple) transformations. That means to find and formulate algorithms either inside the core type class or directly on the deep embedding. Berghammer et. al [2], as well as, Guttmann and Robinson-O'Brien [5] successfully employ algebraic approaches to formally verify graph algorithms. However, they are using an algebra of relations as opposed to the one here defined on graph construction primitives. Dolan [3] uses a semiring on adjacency matrices to find paths in graphs. Mokhov himself postulates that the algebraic approach allows to formulate graph algorithms as systems of equations with unknowns. This may also be an approach for finding novel graph algorithms.

References

[1] Jonathan Beaumont et al. "High-level asynchronous concepts at the interface between analog and digital worlds". In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 37.1 (2017), pp. 61–74.

[2] Rudolf Berghammer et al. "Relational characterisations of paths". In: *Journal of Logical and Algebraic Methods in Programming* 117 (2020), p. 100590.

---

[5]available at `https://skillsmatter.com/skillscasts/12361-labelled-algebraic-graphs`

[3] Stephen Dolan. "Fun with semirings: a functional pearl on the abuse of linear algebra". In: *Proceedings of the 18th ACM SIGPLAN international conference on Functional programming*. 2013, pp. 101–110.

[4] Martin Erwig. "Inductive graphs and functional graph algorithms". In: *Journal of Functional Programming* 11.5 (2001), pp. 467–492.

[5] Walter Guttmann and Nicolas Robinson-O'Brien. "Relational Minimum Spanning Tree Algorithms". In: *Archive of Formal Proofs* (Dec. 2020). `https://isa-afp.org/entries/Relational_Minimum_Spanning_Trees.html`, Formal proof development. ISSN: 2150-914x.

[6] Maximilian P.L. Haslbeck, Peter Lammich, and Julian Biendarra. "Kruskal's Algorithm for Minimum Spanning Forest". In: *Archive of Formal Proofs* (Feb. 2019). `https://isa-afp.org/entries/Kruskal.html`, Formal proof development. ISSN: 2150-914x.

[7] David J. King and John Launchbury. "Structuring Depth-First Search Algorithms in Haskell". In: *Proceedings of the 22nd ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*. POPL '95. San Francisco, California, USA: Association for Computing Machinery, 1995, pp. 344–354. ISBN: 0897916921. DOI: `10.1145/199448.199530`. URL: `https://doi.org/10.1145/199448.199530`.

[8] Ross M McConnell and Fabien De Montgolfier. "Linear-time modular decomposition of directed graphs". In: *Discrete Applied Mathematics* 145.2 (2005), pp. 198–209.

[9] Andrey Mokhov. "Algebraic graphs with class (functional pearl)". In: *ACM SIGPLAN Notices* 52.10 (2017), pp. 2–13.

[10] Andrey Mokhov et al. "Language and hardware acceleration backend for graph processing". In: *Languages, Design Methods, and Tools for Electronic System Design*. Springer, 2019, pp. 71–88.

[11] Tobias Nipkow, Lawrence C. Paulson, and Markus Wenzel. *Isabelle/HOL — A Proof Assistant for Higher-Order Logic*. Vol. 2283. LNCS. Springer, 2002.

[12] Lars Noschinski. "Graph Theory". In: *Archive of Formal Proofs* (Apr. 2013). `https://isa-afp.org/entries/Graph_Theory.html`, Formal proof development. ISSN: 2150-914x.