

# Second-Order Properties of Undirected Graphs

Walter Guttman

Department of Computer Science and Software Engineering,  
University of Canterbury, Christchurch, New Zealand  
walter.guttman@canterbury.ac.nz

**Abstract.** We study second-order formalisations of graph properties expressed as first-order formulas in relation algebras extended with a Kleene star. The formulas quantify over relations while still avoiding quantification over elements of the base set. We formalise the property of undirected graphs being acyclic this way. This involves a study of various kinds of orientation of graphs. We also verify basic algorithms to constructively prove several second-order properties.

## 1 Introduction

Binary relations and relational operations provide convenient abstractions for expressing various kinds of logical specification in concise ways as the following examples demonstrate:

- Relation  $R$  is transitive if  $RR \subseteq R$  (using relational composition), which is logically equivalent to  $\forall x : \forall y : \forall z : (x, y) \in R \wedge (y, z) \in R \Rightarrow (x, z) \in R$ .
- Point  $Q$  is reachable from point  $P$  in graph  $R$  if  $P \subseteq R^*Q$  (using reflexive-transitive closure  $*$ ), which is equivalent to: there is a number  $n$  and a sequence of vertices  $x_0, \dots, x_n$  with  $\forall i : 0 \leq i < n \Rightarrow (x_i, x_{i+1}) \in R$ , where  $x_0$  and  $x_n$  correspond to  $P$  and  $Q$ , respectively. See Section 2 for the relational specification of points.
- Directed graph  $R$  is acyclic if  $R^+ \subseteq \bar{I}$  (using transitive closure  $^+$  and the complement of the identity relation  $\bar{I}$ ), which is equivalent to: there is no number  $n$  and sequence of vertices  $x_0, \dots, x_n$  with  $\forall i : 0 \leq i < n \Rightarrow (x_i, x_{i+1}) \in R$  and  $(x_n, x_0) \in R$ .

In these examples, conciseness is gained by eliminating quantifiers from logical specifications. The resulting expressions facilitate equational reasoning about entire relations rather than point-wise arguments involving elements of the base set.

The above logical formulas quantify over elements of the base set of the relation. Sometimes quantification over relations is used:

- A relation algebra is pair-dense if

$$\forall R : \mathbf{0} \neq R \subseteq \mathbf{1} \Rightarrow \exists Q : \mathbf{0} \neq Q \subseteq R \wedge Q\bar{I}Q\bar{I}Q \subseteq \mathbf{1}$$

(using the empty relation  $\mathbf{0}$ ) [19].

- The intermediate point theorem states

$$P \subseteq RSQ \Leftrightarrow \exists X : X \text{ is a point} \wedge P \subseteq RX \wedge X \subseteq SQ$$

for any relations  $R$  and  $S$  and any points  $P$  and  $Q$  [30].

- Two characterisations of difunctional relations are

$$R = RR^\top R \Leftrightarrow \exists P : \exists Q : P^\top P \subseteq I \subseteq PP^\top \wedge Q^\top Q \subseteq I \subseteq QQ^\top \wedge R = PQ^\top$$

(using relational converse  $^\top$ ). The formula specifies that  $P$  and  $Q$  are mappings, that is, univalent and total relations; see Section 2. The above equivalence is from [28] which also characterises various types of orders by quantifying over relations.

- A form of the axiom of choice can be expressed as

$$\forall R : I \subseteq RR^\top \Rightarrow \exists Q : Q \subseteq R \wedge Q^\top Q \subseteq I \subseteq QQ^\top$$

This considers the set of  $R$ -image sets of each element of the base set, and selects one element from each according to choice function  $Q$ . The formula specifies that  $R$  is total and  $Q$  is a mapping.

Of course, already the axioms of relation algebras universally quantify over relations, but the above properties also use existential quantification. We call properties that quantify over relations ‘second-order’ to distinguish them from logical formulas that quantify over elements of the base set. We express these properties in the language of relation algebras extended with a Kleene star, which abstracts from elements of the base set. Hence, in this language, we can use first-order formulas with variables ranging over the elements of a relation algebra.

In this paper we study second-order properties that are useful for the application area of graphs. One of the motivations for this work is that while  $R^+ \subseteq \bar{I}$  concisely states that directed graph  $R$  is acyclic, no similarly compact formalisation of acyclicity is known for *undirected* graphs represented by symmetric relations. This complicates the formalisation of graph algorithms and their verification [13, 14]. The focus of this paper is to present a number of second-order properties and study their relationships; future work will use these properties to simplify relational correctness proofs of graph algorithms.

Relation algebras are frequently associated with the aim of eliminating logical quantifiers and thereby enabling point-free equational reasoning at a higher abstraction level. The present work does not contradict this aim by reintroducing quantifiers. The quantifiers in our formulas are second-order, that is, they quantify over relations not elements of the base set. For comparison, consider the map-fusion law for lists in functional programming [4]. Its point-wise form uses functions  $f$  and  $g$  and a list  $xs$ :

$$\text{map } f (\text{map } g \text{ } xs) = \text{map } (f \circ g) \text{ } xs$$

Its point-free form eliminates the list argument  $xs$ :

$$\text{map } f \circ \text{map } g = \text{map } (f \circ g)$$

It still involves implicit quantification over functions  $f$  and  $g$ , but the law can now be understood as talking about functions rather than lists. The variables  $f$  and  $g$  are ‘higher-order points’ and not usually eliminated from this law, though they could be removed in formalisms like combinatory logic [7, 31].

The contributions of this paper are:

- We study and compare various notions of orientability of undirected graphs in Section 3. They serve as a basis for formalising more specific properties.
- We introduce several second-order formalisations of the property that an undirected graph is acyclic in Section 4. We prove a number of relationships between these formulas and give counterexamples in cases where formulas are not equivalent in relation algebras extended with a Kleene star.
- We give several equivalent formalisations of general and specific transitively orientable graphs in Section 5. We also formalise the property that an undirected graph contains only simple paths.
- We verify the correctness of several basic algorithms in Section 6 to constructively prove a number of axioms used throughout this paper.

Moreover, all results in this paper except the counterexamples have been formally verified in Isabelle/HOL [25]. The corresponding proofs are omitted and can be found in the Archive of Formal Proofs [15].

## 2 Relational and Algebraic Basics

This section recalls algebras we will use for reasoning about properties of directed and undirected graphs in the remainder of the paper. In particular we discuss Boolean algebras, relation algebras and Kleene relation algebras. We also recall basic relational definitions and give a number of general results.

A *Boolean algebra* [9] is a structure  $(S, \sqcup, \sqcap, \bar{\phantom{x}}, \perp, \top)$  such that

$$\begin{array}{ll}
 x \sqcup (y \sqcup z) = (x \sqcup y) \sqcup z & x \sqcap (y \sqcap z) = (x \sqcap y) \sqcap z \\
 x \sqcup y = y \sqcup x & x \sqcap y = y \sqcap x \\
 x \sqcup x = x & x \sqcap x = x \\
 x \sqcup \perp = x & x \sqcap \top = x \\
 x \sqcup \top = \top & x \sqcap \perp = \perp \\
 x \sqcup (x \sqcap y) = x & x \sqcap (x \sqcup y) = x \\
 x \sqcup (y \sqcap z) = (x \sqcup y) \sqcap (x \sqcup z) & x \sqcap (y \sqcup z) = (x \sqcap y) \sqcup (x \sqcap z) \\
 x \sqcup \bar{x} = \top & x \sqcap \bar{x} = \perp
 \end{array}$$

for each  $x, y, z \in S$ . The axioms specify that the operations  $\sqcup$  and  $\sqcap$  are associative, commutative and idempotent, have units  $\perp$  and  $\top$ , have zeros  $\top$  and  $\perp$ , absorb each other, distribute over each other and are complementary.

The lattice order is obtained by  $x \sqsubseteq y \Leftrightarrow x \sqcup y = y$  or the equivalent  $x \sqsubseteq y \Leftrightarrow x \sqcap y = x$ . The join  $x \sqcup y$  is the  $\sqsubseteq$ -least upper bound of  $x$  and  $y$ ; their

meet or  $\sqsubseteq$ -greatest lower bound is  $x \sqcap y$ . The  $\sqsubseteq$ -least element is  $\perp$ ; the  $\sqsubseteq$ -greatest element is  $\top$ . The element  $\bar{x}$  is the complement of  $x$ .

A *relation algebra* [33] is a structure  $(S, \sqcup, \sqcap, \cdot, \bar{\cdot}, \top, \perp, \top, 1)$  such that the reduct  $(S, \sqcup, \sqcap, \bar{\cdot}, \perp, \top)$  is a Boolean algebra and

$$\begin{aligned} x \cdot (y \cdot z) &= (x \cdot y) \cdot z & (x \cdot y)^\top &= y^\top \cdot x^\top \\ x \cdot 1 &= x & (x^\top)^\top &= x \\ (x \sqcup y) \cdot z &= (x \cdot z) \sqcup (y \cdot z) & (x \sqcup y)^\top &= x^\top \sqcup y^\top \\ & & x^\top \cdot \bar{x \cdot y} &\sqsubseteq \bar{y} \end{aligned}$$

for each  $x, y, z \in S$ . It follows that composition  $\cdot$  is a monoid with identity 1 and distributes over join, transpose  $\top$  is involutive, antidistributes over composition and distributes over join and meet, and De Morgan's Theorem K holds. We abbreviate  $x \cdot y$  by  $xy$ .

An element  $x$  of a relation algebra is reflexive if  $1 \sqsubseteq x$ , irreflexive if  $x \sqsubseteq \bar{1}$ , symmetric if  $x^\top = x$ , asymmetric if  $x \sqcap x^\top = \perp$ , antisymmetric if  $x \sqcap x^\top \sqsubseteq 1$ , transitive if  $xx \sqsubseteq x$ , a partial order if  $x$  is reflexive and antisymmetric and transitive, a strict order if  $x$  is irreflexive and transitive, a total order if  $x \sqcup x^\top = \top$  and  $x$  is a partial order, a strict total order if  $x \sqcup x^\top = \bar{1}$  and  $x$  is a strict order, univalent if  $x^\top x \sqsubseteq 1$ , injective if  $xx^\top \sqsubseteq 1$ , total if  $1 \sqsubseteq xx^\top$ , surjective if  $1 \sqsubseteq x^\top x$ , bijective if  $x$  is injective and surjective, a vector if  $x^\top = x$ , a point if  $x$  is a bijective vector, and an arc if  $x^\top$  and  $x^\top x^\top$  are points. The symmetric closure of  $x$  is  $x \sqcup x^\top$ . See [30] for further details about these properties.

A *Kleene relation algebra* is a structure  $(S, \sqcup, \sqcap, \cdot, \bar{\cdot}, \top, *, \perp, \top, 1)$  such that the reduct  $(S, \sqcup, \sqcap, \cdot, \bar{\cdot}, \top, \perp, \top, 1)$  is a relation algebra and

$$\begin{aligned} 1 \sqcup xx^* &\sqsubseteq x^* & xy \sqsubseteq y &\Rightarrow x^*y \sqsubseteq y \\ 1 \sqcup x^*x &\sqsubseteq x^* & yx \sqsubseteq y &\Rightarrow yx^* \sqsubseteq y \end{aligned}$$

for each  $x, y \in S$ . It follows that  $x^*y$  is the  $\sqsubseteq$ -least fixpoint of  $\lambda z. xz \sqcup y$  and  $yx^*$  is the  $\sqsubseteq$ -least fixpoint of  $\lambda z. zx \sqcup y$ . The above unfold and induction axioms for the Kleene star  $*$  are from [17]. The transitive closure of  $x$  is  $x^+ = xx^*$  and  $x^*$  models the reflexive-transitive closure of relations. Relation algebras with transitive closure have been studied in [22].

An element  $x$  of a Kleene relation algebra is acyclic if  $x^+$  is irreflexive, and a forest if  $x$  is injective and acyclic.

The following theorem states a number of general results in (Kleene) relation algebras. Theorems 1.2 and 1.3 appear in [28, 30].

**Theorem 1.** *Let  $S$  be a Kleene relation algebra and let  $x, y \in S$ . Then*

1. *Every acyclic element is asymmetric.*
2. *Every asymmetric element is irreflexive.*
3. *Acyclic, asymmetric and irreflexive are equivalent for transitive elements.*
4.  *$x$  is asymmetric if and only if  $xx$  is irreflexive.*
5.  *$x$  is a strict order if and only if  $x$  is transitive and acyclic.*

6.  $x$  is a strict total order if and only if  $x$  is transitive and  $x \sqcup x^\top = \bar{1}$ .
7.  $x$  is acyclic if and only if  $x$  is irreflexive and  $x^*$  is antisymmetric.
8.  $x$  is acyclic if and only if  $x^+$  is asymmetric.
9.  $(x \sqcup y)^+ = x^+ \sqcup y^+ \sqcup x^+ \sqcup y^+$  if  $xy = \perp$ .
10.  $\top(x \sqcap y) \sqcap \top(x \sqcap \bar{y}) = \perp$  if  $x$  is injective.

### 3 Orientations

In the remainder of this paper we model graphs using Kleene relation algebras. A (directed) graph is just an element of (the carrier set of) such an algebra. Graph  $x$  is undirected if  $x$  is symmetric:  $x^\top = x$ .

An *orientation* of undirected graph  $x$  is a directed graph  $y$  obtained by assigning a direction to each edge of  $x$  [8]. Algebraically this is formalised by

$$y \text{ is an orientation of } x \Leftrightarrow_{\text{def}} y \sqcap y^\top = \perp \wedge y \sqcup y^\top = x$$

expressing that  $y$  is asymmetric and its symmetric closure is  $x$ . Asymmetric requires that  $y$  has at most one directed edge between any two vertices; the second equation ensures  $y$  contains at least one directed edge between any two vertices connected by an edge in  $x$ . Graph  $x$  is *orientable* if it has an orientation  $y$ :

$$x \text{ is orientable} \Leftrightarrow_{\text{def}} \exists y : y \sqcap y^\top = \perp \wedge y \sqcup y^\top = x$$

It follows from this formalisation that every orientable graph is symmetric and irreflexive. We now consider the converse, namely, that every symmetric irreflexive element is orientable:

$$\forall x : x = x^\top \wedge x \sqsubseteq \bar{1} \Rightarrow \exists y : y \sqcap y^\top = \perp \wedge y \sqcup y^\top = x \quad (0)$$

The structure of this formula is similar to that of the axiom of choice given in Section 1; essentially a direction is chosen for each edge.

Formula (0) is independent of the axioms of Kleene relation algebras as witnessed by the following counterexample found by Nitpick [6]. The set  $\{\perp, 1, \bar{1}, \top\}$  of relations over a two-element base set forms a Kleene relation algebra which is a subalgebra of the full algebra of relations. In this subalgebra,  $\bar{1}$  is symmetric and irreflexive but not orientable.

We study two variants of orientations: one that admits loops and one that admits additional edges with an assigned direction.

- $y \sqcap y^\top \sqsubseteq 1 \wedge y \sqcup y^\top = x$  replaces asymmetric with antisymmetric in the definition of an orientation. This allows loops in  $x$ , which then must also occur in the orientation  $y$ . We call this a *loop-orientation*.
- $y \sqcap y^\top = \perp \wedge y \sqcup y^\top \sqsupseteq x$  requires the symmetric closure to contain  $x$  rather than to equal  $x$ . So  $y$  can contain extra edges, but at most one direction of each. We call this a *super-orientation*.
- $y \sqcap y^\top \sqsubseteq 1 \wedge y \sqcup y^\top \sqsupseteq x$  combines the two variants to obtain *loop-super-orientations*.

Definitions of loop-orientable, super-orientable and loop-super-orientable are derived for these variants similarly to orientable. Using these notions, we obtain several formulas equivalent to formula (0) as the following result shows.

**Theorem 2.** *The following eight properties are equivalent:*

1. *Every symmetric irreflexive element is orientable, that is, formula (0) holds.*
2. *Every symmetric element is loop-orientable.*
3. *Every irreflexive element is super-orientable.*
4. *Every element is loop-super-orientable.*
5.  $\forall x : x = x^\top \Rightarrow \exists y : y \sqcap y^\top = x \sqcap 1 \wedge y \sqcup y^\top = x.$
6.  $\forall x : x = x^\top \Rightarrow \exists y : y \sqcap y^\top \sqsubseteq x \sqcap 1 \wedge y \sqcup y^\top = x.$
7.  $\forall x : \exists y : y \sqcap y^\top = x \sqcap 1 \wedge y \sqcup y^\top \sqsupseteq x.$
8.  $\forall x : \exists y : y \sqcap y^\top \sqsubseteq x \sqcap 1 \wedge y \sqcup y^\top \sqsupseteq x.$

Theorems 2.2–2.4 show how the notions of loop-/super-orientation allow the assumptions of irreflexive/symmetric to be removed from formula (0).

The definition of an orientation generalises to the following useful ternary predicate:

$$S(x, y, z) \Leftrightarrow_{\text{def}} y \sqcap y^\top = x \wedge y \sqcup y^\top = z$$

In words, the meet of  $y$  and  $y^\top$  is  $x$  and their join is  $z$ . Both  $x$  and  $z$  need to be symmetric for  $S(x, y, z)$  to hold, and  $x \sqsubseteq y \sqsubseteq z$  follows, too. Hence, the intuitive meaning for undirected graphs  $x$  and  $z$  is:

- If an edge is in  $x$  and in  $z$ , it is also in  $y$ .
- If an edge is not in  $x$  and not in  $z$ , it is also not in  $y$ .
- If an edge is not in  $x$  but in  $z$ , exactly one direction of it is in  $y$ .

The following result gives consequences of this definition.

**Theorem 3.**

1.  $S(\perp, y, x)$  if and only if  $y$  is an orientation of  $x$ .
2.  $S(1, y, x)$  implies that  $y$  is a loop-orientation of  $x$ .
3.  $S(x \sqcap 1, y, x)$  if and only if  $y$  is a loop-orientation of  $x$ .
4.  $S(y \sqcap 1, y, x)$  if and only if  $y$  is a loop-orientation of  $x$ .
5.  $S(x, y, z)$  if and only if  $y \sqcap y^\top = x \sqsubseteq z \wedge (y \sqcap \bar{y}^\top) \sqcup (y \sqcap \bar{y}^\top)^\top = z \sqcap \bar{x}$ .

Theorem 3.5 gives an alternative way to specify the ternary predicate. It requires  $x \sqsubseteq z$ , that  $x$  is the symmetric part of  $y$ , and that the difference  $z \sqcap \bar{x}$  is the symmetric closure of the asymmetric part of  $y$ ; see [21] for a study of the symmetric and asymmetric parts of a relation.

A special case of Theorem 3.1 is that  $S(\perp, y, \bar{1})$  if and only if  $y$  is an orientation of  $\bar{1}$ . An orientation of the complete graph (without loops)  $\bar{1}$  is also known as a *tournament* [8]. The existence of a tournament is equivalent to the conditions in Theorem 2 as the following result shows.

**Theorem 4.** *The following three properties are equivalent:*

1. *Formula (0) holds.*
2.  $\exists y : S(\perp, y, \bar{1})$ .
3.  $\exists y : S(1, y, \top)$ .

There are various ways of strengthening orientability. One is to require the orientation to be injective:

$$x \text{ is injectively orientable} \Leftrightarrow_{\text{def}} \exists y : y \sqcap y^\top = \perp \wedge y \sqcup y^\top = x \wedge yy^\top \sqsubseteq 1$$

Injectively orientable graphs correspond to graphs in which every component has at most one cycle, also known as pseudoforests [11, 27]. They will be used in Section 4. A different strengthening requires orientations to be transitive [28]:

$$x \text{ is transitively orientable} \Leftrightarrow_{\text{def}} \exists y : y \sqcap y^\top = \perp \wedge y \sqcup y^\top = x \wedge yy^\top \sqsubseteq y$$

Transitively orientable graphs, also known as comparability graphs, are the symmetric closures of strict orders. They will be used in Section 5.

## 4 Acyclicity of Undirected Graphs

In this section we discuss various ways to specify that an undirected graph  $x$  is acyclic. When justifying specifications informally, we implicitly assume that  $x$  is symmetric and irreflexive; we explicitly state such assumptions in theorems.

We present the specifications in order of increasing strength, give equivalent characterisations for most of them and study their relationships.

**4.1** Our first specification requires that every orientation of  $x$  is acyclic (in the sense of directed graphs):

$$\forall y : y \sqcap y^\top = \perp \wedge y \sqcup y^\top = x \Rightarrow y^+ \sqsubseteq \bar{1} \tag{1}$$

Intuitively, if  $x$  contained an undirected cycle then this cycle could be oriented and extended to an orientation of  $x$  that would not be acyclic. Conversely, if some orientation of  $x$  contained a cycle then the symmetric closure of this cycle would be an undirected cycle in  $x$ .

The following result shows an equivalent formulation of (1). It replaces  $y^+ \sqsubseteq \bar{1}$  with  $y^* \sqcap y^{\top*} = 1$ , which is equivalent for irreflexive  $y$ .

**Theorem 5.** *The following two properties are equivalent for any  $x$ :*

1.  *$x$  satisfies formula (1).*
2.  $\forall y : y \sqcap y^\top = \perp \wedge y \sqcup y^\top = x \Rightarrow y^* \sqcap y^{\top*} = 1$ .

**4.2** Our second specification weakens the antecedent of formula (1) to asymmetric subsets of  $x$ :

$$\forall y : y \sqcap y^\top = \perp \wedge y \sqsubseteq x \Rightarrow y^+ \sqsubseteq \bar{1} \quad (2)$$

Every orientation of  $x$  clearly satisfies  $y \sqsubseteq x$ , so formula (2) implies formula (1). The converse implication holds for orientable elements according to the following result. It also gives equivalent formulations of (2).

**Theorem 6.** *The following three properties are equivalent for any symmetric  $x$ :*

1.  $x$  satisfies formula (2).
2.  $\forall y : y \sqcap y^\top = \perp \wedge y \sqcup y^\top \sqsubseteq x \Rightarrow y^+ \sqsubseteq \bar{1}$ .
3.  $\forall y : y \sqcap y^\top = \perp \wedge y \sqcup y^\top \sqsubseteq x \Rightarrow y^* \sqcap y^{\top*} = 1$ .

*The last two of the above properties are equivalent for any  $x$ . Moreover,*

4. *Formula (2) implies formula (1) for any  $x$ .*
5. *Formulas (2) and (1) are equivalent for any orientable  $x$ .*

A counterexample shows that formula (1) does not imply formula (2) for all symmetric irreflexive elements. The complex algebra  $\text{Cm}(G)$  of any group  $G$  is a relation algebra; see [12, 19] for construction details. Moreover,  $\text{Cm}(G)$  is a Kleene relation algebra using  $x^* = \bigcup_{i \in \mathbb{N}} x^i$ . Consider  $\text{Cm}(\mathbb{Z}_4)$  where  $\mathbb{Z}_4 = \{0, 1, 2, 3\}$  is the cyclic group of order 4. In  $\text{Cm}(\mathbb{Z}_4)$  the complex  $x = \bar{1} = \{1, 2, 3\}$  satisfies formula (1) since  $x$  has no orientation as it is above symmetric atom  $\{2\}$ . But  $x$  is also above non-symmetric atom  $y = \{1\}$  with  $y^\top = \{3\}$  and  $y^+ = \top = \mathbb{Z}_4$ , whence  $x$  does not satisfy formula (2).

**4.3** Our third specification avoids the reference to acyclic subgraphs. It requires that there is a unique way to sandwich  $x$  between a graph and its reflexive-transitive closure:

$$\forall y : y \sqsubseteq x \sqsubseteq y^* \Rightarrow y = x \quad (3)$$

Intuitively, if  $x$  contained an undirected cycle then one edge of this cycle could be removed without affecting reachability in the graph, so an element strictly below  $x$  would satisfy the antecedent. Conversely, if there was a  $y$  with  $y \sqsubseteq x \sqsubseteq y^*$  then there would be an edge  $e$  in  $x$  that is not in  $y$  but in  $y^*$ , so there would be a path in  $y$  from the start vertex of  $e$  to its end vertex which together with  $e$  would form a cycle. The following result shows equivalent formulations of (3).

**Theorem 7.** *The following two properties are equivalent for any  $x$ :*

1.  $x$  satisfies formula (3).
2.  $\forall y : y \sqsubseteq x \wedge y^* = x^* \Rightarrow y = x$ .

*The following two properties are equivalent for any  $x$ :*

3.  $\forall y : y \sqsubseteq x \sqsubseteq y^+ \Rightarrow y = x$ .
4.  $\forall y : y \sqsubseteq x \wedge y^+ = x^+ \Rightarrow y = x$ .

*All four of the above properties are equivalent for any irreflexive  $x$ .*

**4.4** Our fourth specification expresses the justification underlying formula (3) more directly:

$$\forall y : y \sqsubseteq x \Rightarrow x \sqcap y^* \sqsubseteq y \quad (4)$$

Intuitively, any edge  $e$  contained in both  $x$  and  $y^*$  must already be in  $y$ , otherwise the path obtained from  $y^*$  together with  $e$  would form a cycle. The following result shows that formulas (4) and (3) are equivalent and stronger than formula (2). It also gives further equivalent formulations of (4).

**Theorem 8.** *The following three properties are equivalent for any  $x$ :*

1.  $x$  satisfies formula (4).
2.  $\forall y : y \sqsubseteq x \Rightarrow x \sqcap y^* = y$ .
3.  $\forall y : y \sqsubseteq x \Rightarrow y \sqcap (x \sqcap \bar{y})^* = \perp$ .
4.  $x$  satisfies formula (3).

Moreover,

5. Formula (4) implies formula (2) for any symmetric  $x$ .
6. Formulas (4) and (2) are equivalent for any symmetric irreflexive  $x$  if the following two axioms hold:

$$\forall u : u \neq \perp \Rightarrow \exists v : v \text{ is an arc} \wedge v \sqsubseteq u$$

$$\forall u : \forall v : u \text{ is an arc} \wedge u \sqsubseteq v^* \Rightarrow \exists w : w \sqsubseteq v \wedge w \sqcap w^\top = \perp \wedge u \sqsubseteq w^*$$

The first of these axioms specifies that every non-empty graph contains an edge, which is similar to the point axiom [19, 29]. The second of these axioms states that if the end vertex of an edge  $u$  is reachable from its start vertex using (directed) edges in  $v$ , then the same holds already in an asymmetric subset  $w$  of  $v$ . Intuitively, the asymmetric subset  $w$  is formed by the edges on the (directed) path from the start vertex of  $u$  to its end vertex.

A counterexample found by Nitpick shows that formula (2) does not imply formula (4) for all symmetric irreflexive elements. The set of symmetric complexes  $\text{SCm}(G) = \{x \in \text{Cm}(G) \mid x = x^\top\}$  of a commutative group  $G$  forms a relation algebra which is a subalgebra of  $\text{Cm}(G)$  [12, 16]. Since  $\text{SCm}(G)$  is closed under Kleene star, it also forms a Kleene relation algebra. In  $\text{SCm}(\mathbb{Z}_4)$  the complex  $x = \bar{1} = \{1, 2, 3\}$  satisfies formula (2) because the only asymmetric complex is  $\perp = \emptyset$ . But  $x$  also contains atom  $y = \{1, 3\}$  with  $y^* = \top = \mathbb{Z}_4$ , whence  $x$  does not satisfy formula (4).

**4.5** Our fifth specification generalises the formulation given in Theorem 8.3. According to the latter formulation there cannot be an edge  $e$  in  $y$  such that there is a path from the source of  $e$  to its target using edges of  $x$  that are not in  $y$ . We now allow the edge  $e$  to be in  $y^*$ :

$$\forall y : y \sqsubseteq x \Rightarrow y^* \sqcap (x \sqcap \bar{y})^* = 1 \quad (5)$$

Intuitively, if there is a path in  $y$ , there cannot be a path from its start vertex to its end vertex using edges of  $x$  that are not in  $y$ , except if the start and end vertices coincide. Namely, if the start and end vertices were different, the two disjoint paths together would form a cycle. The following result shows that formula (5) is stronger than formula (4). It also gives equivalent formulations of (5).

**Theorem 9.** *The following six properties are equivalent for any  $x$ :*

1.  $x$  satisfies formula (5).
2.  $\forall y : y \sqsubseteq x \Rightarrow y^* \sqcap (x \sqcap \bar{y})^+ \sqsubseteq 1$ .
3.  $\forall y : y \sqsubseteq x \Rightarrow y^+ \sqcap (x \sqcap \bar{y})^* \sqsubseteq 1$ .
4.  $\forall y : y \sqsubseteq x \Rightarrow y^+ \sqcap (x \sqcap \bar{y})^+ \sqsubseteq 1$ .
5.  $\forall y : \forall z : y \sqcap z = \perp \wedge y \sqcup z \sqsubseteq x \Rightarrow y^* \sqcap z^* = 1$ .
6.  $\forall y : \forall z : y \sqcap z = \perp \wedge y \sqcup z = x \Rightarrow y^* \sqcap z^* = 1$ .

Moreover,

7. Formula (5) implies formula (4) for any irreflexive  $x$ .

The formulation in Theorem 9.6 is particularly conspicuous. If  $x$  is partitioned into  $y$  and  $z$ , then there cannot be a path from the same start vertex to the same end vertex in both partitions, except for the empty path if the start and end vertices coincide. The formulations in Theorems 9.5 and 9.6 generalise the formulations in Theorems 6.3 and 5.2, respectively, by replacing  $y^T$  with a new variable  $z$ .

A counterexample shows that formula (4) does not imply formula (5) for all symmetric irreflexive elements. Consider  $\mathbb{Z}_{12} = \{0, 1, \dots, 10, 11\}$ , the cyclic group of order 12. In  $\text{SCm}(\mathbb{Z}_{12})$ , complex  $x = \{2, 3, 9, 10\} \sqsubseteq \bar{1}$  satisfies formula (4) since only complexes  $\perp$ ,  $y_1 = \{2, 10\}$ ,  $y_2 = \{3, 9\}$  and  $x$  are below  $x$  and

$$\begin{aligned} x \sqcap \perp^* &= x \sqcap 1 \sqsubseteq \bar{1} \sqcap 1 = \perp \\ x \sqcap y_1^* &= x \sqcap \{0, 2, 4, 6, 8, 10\} = y_1 \\ x \sqcap y_2^* &= x \sqcap \{0, 3, 6, 9\} = y_2 \\ x \sqcap x^* &= x \end{aligned}$$

But  $x$  does not satisfy formula (5) since

$$y_1^* \sqcap (x \sqcap \bar{y}_1)^* = y_1^* \sqcap y_2^* = \{0, 2, 4, 6, 8, 10\} \sqcap \{0, 3, 6, 9\} = \{0, 6\} \neq \{0\} = 1$$

**4.6** Our sixth specification asserts the existence of an orientation that is a forest. To this end, we strengthen the property of being injectively orientable, introduced at the end of Section 3, by replacing asymmetric with acyclic:

$$\exists y : y \sqcup y^T = x \wedge y^+ \sqsubseteq \bar{1} \wedge yy^T \sqsubseteq 1 \quad (6)$$

Note that  $y^+ \sqsubseteq \bar{1}$  implies that  $y$  is asymmetric. With  $y \sqcup y^T = x$  it follows that  $y$  is an orientation of  $x$ . The properties acyclic and injective together are

frequently used to specify forests with edges directed away from the roots of the component trees. Overall, the above property requires that there is a (directed) forest whose symmetric closure is  $x$ .

Whereas the previous specifications of acyclic graphs were universally quantified, formulation (6) is existentially quantified. The following result relates formula (6) to both the strongest and the weakest of the previous specifications, namely, formulas (5) and (1).

**Theorem 10.** *The following two properties are equivalent for any  $x$ :*

1.  $x$  satisfies formula (6).
2.  $x$  is injectively orientable and satisfies formula (1).

Moreover,

3. Formula (6) implies formulas (1)–(5) for any  $x$ .

The counterexample showing independence of formula (0) given in Section 3 also shows that formula (5) does not imply formula (6) for all symmetric irreflexive elements. In that algebra,  $x = \bar{1}$  is an atom and satisfies the formulation in Theorem 9.6 since either  $y = \perp$  or  $z = \perp$  for any partition of  $x$ . But since  $x$  is not orientable, it does not satisfy formula (6) by Theorem 10.

We furthermore consider the following weakening of formula (6), which replaces the condition  $y \sqcup y^\top = x$  with two of its consequences  $y \sqsubseteq x$  and  $x \sqsubseteq (y \sqcup y^\top)^*$ :

$$x \text{ is spannable} \Leftrightarrow_{\text{def}} \exists y : y \sqsubseteq x \sqsubseteq (y \sqcup y^\top)^* \wedge y^+ \sqsubseteq \bar{1} \wedge yy^\top \sqsubseteq 1$$

This means that  $y$  no longer needs to contain a direction of every edge of  $x$ , but some edges can be entirely omitted provided their end vertices are still weakly connected in  $y$ . In other words,  $y$  is a spanning forest of  $x$ . A similar formalisation of spanning forests has been used in [14] for verifying the correctness of Kruskal's minimum spanning forest algorithm. The significance of being spannable for the present work is captured in the following result.

**Theorem 11.** *The following two properties are equivalent for any  $x$ :*

1.  $x$  satisfies formula (6).
2.  $x$  is symmetric and spannable and satisfies formula (3).

Moreover,

3.  $\bar{1}$  is spannable if a point exists.

## 5 Transitive Orientations and Simple Paths

In Section 3 we have studied the existence of tournaments, that is, orientations of the complete graph without loops  $\bar{1}$ . In this section we additionally require orientations to be transitive.

Every orientation is asymmetric, and transitive asymmetric relations correspond to strict orders. Hence, the transitively orientable graphs are precisely the graphs of strict orders after ignoring edge directions. Applied to the complete graph without loops  $\bar{1}$  this amounts to the existence of a strict total order.

**Theorem 12.** *The following two properties are equivalent for any  $x$ :*

1.  $x$  is transitively orientable.
2.  $x$  is irreflexive and  $\exists y : y \sqcup y^\top = x \wedge yy \sqsubseteq y$ .

*In particular, the following two properties are equivalent:*

3.  $\bar{1}$  is transitively orientable.
4.  $\exists y : y \sqcup y^\top = \bar{1} \wedge yy \sqsubseteq y$ .

*Moreover, each of the last two properties implies formula (0).*

The following result gives additional equivalent properties.

**Theorem 13.** *The following five properties are equivalent:*

1.  $\bar{1}$  is transitively orientable.
2.  $\exists y : y \sqcap y^\top \sqsubseteq 1 \wedge y \sqcup y^\top = \top \wedge yy \sqsubseteq y$ .
3.  $\exists y : y^+ \sqsubseteq \bar{1} \wedge y^* \sqcup y^{*\top} = \top$ .
4.  $\exists y : S(\perp, y^+, \bar{1})$ .
5.  $\exists y : S(1, y^*, \top)$ .

Theorem 13.2 is a translation of Theorem 12.4 to partial orders. The property in Theorem 13.3 requires that  $y$  is acyclic and unilaterally connected, in other words, between any two vertices there is exactly one (directed) path in  $y$ . Theorems 13.4 and 13.5 express this using the ternary predicate of Section 3.

We finally consider a special case of undirected acyclic graphs, namely those whose maximum degree is at most 2, that is, at most two edges are incident to each vertex. Every component of such a graph is a simple path [2]. To specify this we strengthen formula (6) by additionally requiring  $y$  to be univalent:

$$\exists y : y \sqcup y^\top = x \wedge y^+ \sqsubseteq \bar{1} \wedge yy^\top \sqsubseteq 1 \wedge y^\top y \sqsubseteq 1 \quad (7)$$

Intuitively, if the maximum degree of an acyclic undirected graph is at most 2, it can be oriented by choosing a directed simple path for each of its components. Conversely, if there is a vertex with at least 3 incident edges, any orientation will have at least two incoming or two outgoing edges at that vertex, making the orientation not injective or not univalent. Graphs with maximum degree 2 are not transitively orientable in general, but according to the following result their transitive closure (without loops) is transitively orientable.

**Theorem 14.**

1. Formula (7) implies formula (6) for any  $x$ .
2.  $x^+ \sqcap \bar{1}$  is transitively orientable if  $x$  satisfies formula (7).

## 6 Axioms and Algorithmic Proofs

In previous sections we have encountered two kinds of property. Properties such as being injectively/transitively orientable or being acyclic hold for some graphs but not for others. In contrast, properties such as formula (0), the two axioms in Theorem 8.6, and  $\bar{1}$  being transitively orientable do not have free variables. Hence, they can serve as axioms that may or may not be assumed to hold in an algebraic setting. In this section we prove that these axioms hold under certain conditions. The conditions are somewhat restrictive from an algebraic perspective but nevertheless satisfied for many practical applications. Our focus is on the proof method which uses constructive algorithms.

For this section we assume that the given Kleene relation algebra is finite and the arc axiom holds, that is, every element except  $\perp$  contains an arc:

$$\forall x : x \neq \perp \Rightarrow \exists y : y \text{ is an arc} \wedge y \sqsubseteq x$$

Finiteness is used to prove that algorithms terminate.

We first show that  $\bar{1}$  is transitively orientable. To this end we use Szpilrajn's algorithm [32]. It constructs a total order that extends a given partial order. By applying this algorithm to the discrete partial order 1 we obtain the desired total order on the base set.

Partial correctness of Szpilrajn's algorithm has been proved in [3] using the automated theorem prover Prover9 [20]. We have transcribed the algorithm to Isabelle/HOL and proved its correctness using a Hoare-logic library [23, 24], which we have extended to total correctness in previous work [14]. From the total-correctness proof we can extract the following result [28].

**Theorem 15.** *For every partial order  $p$  there is a total order  $t$  with  $p \sqsubseteq t$ .*

In particular, by setting  $p = 1$  there exists a total order  $t$ , which is the condition in Theorem 13.2. Hence,  $\bar{1}$  is transitively orientable by Theorem 13. Moreover, formula (0) holds by Theorem 12.

We next show that every symmetric element is spannable. To this end we use Kruskal's algorithm, which constructs a minimum spanning forest of an undirected graph [18]. We modify the algorithm so as to ignore edge weights, in which case it constructs a spanning forest. We have reused an existing specification and correctness proof of this algorithm from previous work [14]. The postcondition of the algorithm implies that the graph is spannable as per the definition in Section 4. The following result is a consequence of this.

**Theorem 16.** *Every symmetric element is spannable.*

We finally establish the second axiom given in Theorem 8.6.

**Theorem 17.** *Let  $x$  be an arc and let  $y$  be such that  $x \sqsubseteq y^*$ . Then there is an asymmetric  $z$  with  $z \sqsubseteq y$  and  $x \sqsubseteq z^*$ .*

To construct the desired element  $z$ , we augment a standard breadth-first search algorithm [1] with a variable  $t$ :

1	input $r, s$	$r =$ directed graph, $s =$ start vertex
2	$t \leftarrow \perp$	$t =$ constructed asymmetric element
3	$q \leftarrow s$	$q =$ vertices reached so far
4	$p \leftarrow \bar{s} \sqcap r^\top s$	$p =$ new vertices reached in the next step
5	while $p \neq \perp$ do	any new vertices?
6	$t \leftarrow t \sqcup (r \sqcap qp^\top)$	include all edges from current to new vertices
7	$q \leftarrow q \sqcup p$	add new vertices to current set
8	$p \leftarrow \bar{q} \sqcap r^\top p$	take another step
9	end	
10	output $q, t$	

As precondition we require that  $s$  is a vector. The loop invariant states that  $q$  is a vector,  $t$  is asymmetric,  $t \sqsubseteq r$  and  $t \sqsubseteq q = t^{\top*}s$  and  $p = \bar{q} \sqcap r^\top q$ . This allows us to prove the postcondition that  $t$  is asymmetric,  $t \sqsubseteq r$  and  $q = t^{\top*}s = r^{\top*}s$ . Termination of the algorithm follows using the number of elements below  $\bar{q} \sqcap r^{\top*}s$  as bound function. This reflects vertices that are reachable from  $s$  in  $r$  but have not been reached so far. From this total-correctness proof we can extract the following result.

**Theorem 18.** *For every element  $r$  and every vector  $s$  there is an asymmetric  $t$  with  $t \sqsubseteq r$  and  $t^{\top*}s = r^{\top*}s$ .*

We can now show Theorem 17. Given an arc  $x$  and an element  $y$  with  $x \sqsubseteq y^*$ , we apply Theorem 18 using  $r = y$  and vector  $s = x^\top$ , obtaining an element  $t$  with the stated properties. The desired  $z$  is obtained as  $z = t \sqcap y$ . It is clearly below  $y$  and asymmetric as it is below the asymmetric  $t$ ; moreover  $x \sqsubseteq z^*$  follows from the assumptions and  $t^{\top*}x^\top = y^{\top*}x^\top$ .

## 7 Conclusion

In this paper we have used second-order properties expressed in relation algebras extended with a Kleene star to formalise that an undirected graph is acyclic in various ways. The formalisations are based on the concept of orientability, which we have therefore studied. We have also verified the correctness of constructive algorithms to validate several of the axioms.

The quantifiers used in second-order properties cause no issues for formal reasoning in Isabelle/HOL whose logic directly supports them as first-order formulas in relation algebras. Sledgehammer [5, 26] can also be applied to such formulas and its integrated provers automatically find proofs in some cases. We should note that the formulas considered in this paper do not have complex nestings of quantifiers. In cases where steps are too big for automated provers, the quantifiers are easy to handle manually as most of them are at the outermost level. The integration with equational reasoning in the proof language Isar [34] is

seamless; it might be further improved by automatically generating bits of boilerplate code to break down quantifiers depending on the structure of a formula. Isabelle/HOL would also support formalising the properties using second-order quantification over concrete relations, but we prefer working in relation algebras.

For simplicity we have presented all results in the framework of Kleene relation algebras. Our Isabelle/HOL theory shows that most results hold in more general structures, such as single-object bounded distributive allegories [10], Stone relation algebras and Stone-Kleene relation algebras [13, 14]. A possible exception is the result that formula (3) implies formula (4), which we were able to prove only in Kleene relation algebras. The more general structures are useful for modelling weighted graphs. We will therefore apply the specifications of acyclic undirected graphs introduced in Section 4 to the verification of graph algorithms involving edge weights. Future work will consider the formalisation of further properties of graphs using higher-order formulas.

**Acknowledgement.** I thank Nicolas Robinson-O’Brien and the anonymous referees for helpful comments.

## References

1. Berghammer, R.: Combining relational calculus and the Dijkstra–Gries method for deriving relational programs. *Information Sciences* **119**(3–4), 155–171 (1999)
2. Berghammer, R., Furusawa, H., Guttmann, W., Höfner, P.: Relational characterisations of paths. *Journal of Logical and Algebraic Methods in Programming* **117:100590** (2020)
3. Berghammer, R., Struth, G.: On automated program construction and verification. In: Bolduc, C., Desharnais, J., Ktari, B. (eds.) *Mathematics of Program Construction (MPC 2010)*. LNCS, vol. 6120, pp. 22–41. Springer (2010)
4. Bird, R., Wadler, P.: *Introduction to Functional Programming*. Prentice Hall (1988)
5. Blanchette, J.C., Böhme, S., Paulson, L.C.: Extending Sledgehammer with SMT solvers. *Journal of Automated Reasoning* **51**(1), 109–128 (2013)
6. Blanchette, J.C., Nipkow, T.: Nitpick: A counterexample generator for higher-order logic based on a relational model finder. In: Kaufmann, M., Paulson, L.C. (eds.) *Interactive Theorem Proving (ITP 2010)*. LNCS, vol. 6172, pp. 131–146. Springer (2010)
7. Curry, H.B., Feys, R.: *Combinatory Logic*, vol. 1. North-Holland Publishing Company (1958)
8. Diestel, R.: *Graph Theory*. Springer, third edn. (2005)
9. Dwingier, P.: *Introduction to Boolean Algebras*. Physica-Verlag, second edn. (1971)
10. Freyd, P.J., Ščedrov, A.: *Categories, Allegories*, North-Holland Mathematical Library, vol. 39. Elsevier Science Publishers (1990)
11. Gabow, H.N., Tarjan, R.E.: A linear-time algorithm for finding a minimum spanning pseudoforest. *Inf. Process. Lett.* **27**(5), 259–263 (1988)
12. Givant, S.: *Introduction to Relation Algebras*, *Relation Algebras*, vol. 1. Springer (2017)
13. Guttmann, W.: An algebraic framework for minimum spanning tree problems. *Theor. Comput. Sci.* **744**, 37–55 (2018)

14. Guttman, W.: Verifying minimum spanning tree algorithms with Stone relation algebras. *Journal of Logical and Algebraic Methods in Programming* **101**, 132–150 (2018)
15. Guttman, W.: Relational forests. *Archive of Formal Proofs* (2021), available at [https://www.isa-afp.org/entries/Relational\\_Forests.html](https://www.isa-afp.org/entries/Relational_Forests.html)
16. Jipsen, P., Lukács, E.: Minimal relation algebras. *Algebra Universalis* **32**(2), 189–203 (1994)
17. Kozen, D.: A completeness theorem for Kleene algebras and the algebra of regular events. *Information and Computation* **110**(2), 366–390 (1994)
18. Kruskal, Jr., J.B.: On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society* **7**(1), 48–50 (1956)
19. Maddux, R.D.: Pair-dense relation algebras. *Transactions of the American Mathematical Society* **328**(1), 83–131 (1991)
20. McCune, W.: Prover9 and Mace4 (2005–2010), accessed 10 August 2021 online at <https://www.cs.unm.edu/~mccune/prover9/>
21. Monjardet, B.: Axiomatiques et propriétés des quasi-ordres. *Mathématiques et sciences humaines* **63**, 51–82 (1978)
22. Ng, K.C.: Relation Algebras with Transitive Closure. Ph.D. thesis, University of California, Berkeley (1984)
23. Nipkow, T.: Winskel is (almost) right: Towards a mechanized semantics textbook. *Formal Aspects of Computing* **10**(2), 171–186 (1998)
24. Nipkow, T.: Hoare logics in Isabelle/HOL. In: Schwichtenberg, H., Steinbrüggen, R. (eds.) *Proof and System-Reliability*. pp. 341–367. Kluwer Academic Publishers (2002)
25. Nipkow, T., Paulson, L.C., Wenzel, M.: Isabelle/HOL: A Proof Assistant for Higher-Order Logic, LNCS, vol. 2283. Springer (2002)
26. Paulson, L.C., Blanchette, J.C.: Three years of experience with Sledgehammer, a practical link between automatic and interactive theorem provers. In: Sutcliffe, G., Ternovska, E., Schulz, S. (eds.) *Proceedings of the 8th International Workshop on the Implementation of Logics*. pp. 3–13 (2010)
27. Picard, J.C., Queyranne, M.: A network flow solution to some nonlinear 0-1 programming problems, with applications to graph theory. *Networks* **12**(2), 141–159 (1982)
28. Schmidt, G.: *Relational Mathematics*. Cambridge University Press (2011)
29. Schmidt, G., Ströhlein, T.: Relation algebras: Concept of points and representability. *Discrete Mathematics* **54**(1), 83–92 (1985)
30. Schmidt, G., Ströhlein, T.: *Relationen und Graphen*. Springer (1989)
31. Schönfinkel, M.: Über die Bausteine der mathematischen Logik. *Mathematische Annalen* **92**(3–4), 305–316 (1924)
32. Szpilrajn, E.: Sur l’extension de l’ordre partiel. *Fundamenta Mathematicae* **16**, 386–389 (1930)
33. Tarski, A.: On the calculus of relations. *The Journal of Symbolic Logic* **6**(3), 73–89 (1941)
34. Wenzel, M.: Isar – a generic interpretative approach to readable formal proof documents. In: Bertot, Y., Dowek, G., Hirschowitz, A., Paulin, C., Théry, L. (eds.) *Theorem Proving in Higher Order Logics (TPHOLs 1999)*. LNCS, vol. 1690, pp. 167–183. Springer (1999)