Relational Approximation of Maximum Independent Sets (Extended Abstract)

Insa Stucke

Institut für Informatik, Christian-Albrechts-Universität zu Kiel, Germany
ist@informatik.uni-kiel.de

1 Introduction

Previous work has shown that relation algebra (as introduced in [14] and further developed in [9,12,13,15], for example) is well suited for computational problems on many discrete structures. In particular, adjacency or incidence relations can be used to model graphs and special relations like vectors and points to represent subsets of vertices or edges, as shown in [12].

We develop and formally verify a relational program for approximating maximum independent sets in undirected and loop-free graphs. Formal program verification means to show with mathematical rigor that the program is correct with respect to a formal problem specification, in our case we comply with the assertion-based Floyd-Hoare-approach.

At the end, the most interesting task is to prove the desired approximation bound. Of course, therefore we need knowledge about cardinalities of relations. In [7] a characterisation of a cardinality operation is developed and further consequences of it were proved. Based on this cardinality operation we not only prove the approximation bound but also facts about the cardinalities of vectors and points in a calculational, algebraic manner only.

2 Relation-Algebraic Preliminaries

In this section we recall the fundamentals of relation algebra based on the heterogeneous approach of [12,13]. Set-theoretic relations form the standard model of relation algebras. We assume the reader to be familar with the basic operations on set-theoretic relations, viz. R^{T} (transposition), $\overline{R}$ (complementation), $R \cup S$ (union), $R \cap S$ (intersection), $R;S$ (composition), the predicates $R \subseteq S$ (inclusion) and $R = S$ (equality), and the special relations O (empty relation), L (universal relation) and I (identity relation). The Boolean operations, the inclusion and the constants O and L form Boolean lattices.

The theoretical framework for this and many other results concerning relations is that of a (heterogeneous) *relation algebra* in the sense of [12,13], with typed relations as elements. Thus, we write $R : X \leftrightarrow Y$ to express that X is the source, Y is the target. The type of R is denoted by $X \leftrightarrow Y$. As constants and operations of a relation algebra we have those of set-theoretic relations, where we frequently overload the symbols O, L and I, i.e., avoid the binding of types to them. If necessary we use indices as e.g., L_{XY} for L with type $X \leftrightarrow Y$. The axioms of a relation algebra are

(1) the axioms of a Boolean lattice for all relations of the same type under the Boolean operations, the inclusion, empty relation and universal relation,
(2) the associativity of composition and that identity relations are neutral elements with respect to composition,
(3) that $Q;R \subseteq S$, $Q^{\mathsf{T}};\overline{S} \subseteq \overline{R}$ and $\overline{S};R^{\mathsf{T}} \subseteq \overline{Q}$ are equivalent, for all relations Q, R and S (with appropriate types),
(4) that $R \neq \mathsf{O}$ implies $\mathsf{L};R;\mathsf{L} = \mathsf{L}$, for all relations R and all universal relations (with appropriate types).

In [12] the laws of (3) are called the *Schröder rules* and (4) is called the *Tarski rule*. In the relation-algebraic proofs of this paper we will mention only applications of (3), (4) and 'non-obvious' consequences of the axioms. Furthermore, we will assume that complementation and transposition bind stronger than composition, composition binds stronger than union and intersection and that all expressions and formulae are well-typed.

Furthermore, we call a relation R *irreflexive* if $R \subseteq \bar{\mathsf{I}}$, *symmetric* if $R = R^{\mathsf{T}}$ and a mapping if R is *univalent*, i.e., $R^{\mathsf{T}};R \subseteq \mathsf{I}$, and *total*, i.e., $R;\mathsf{L} = \mathsf{L}$ (for more details, see e.g., [12,13]). A *vector* is a relation v with $v = v;\mathsf{L}$. For $v : X \leftrightarrow Y$ the condition $v = v;\mathsf{L}$ means that v can be written in the form $v = Z \times Y$ with a subset Z of X. Then we say that *v models the subset Z* of X. Since for this purpose the target of a vector is irrelevant, we use the specific singleton set $\mathbf{1}$ as target. Moreover, a *point* p is a vector with $p;p^{\mathsf{T}} \subseteq \mathsf{I}$ and $\mathsf{L};p = \mathsf{L}$. In the set-theoretic case and if the point $p : X \leftrightarrow Y$ is of the specific form $p = P \times Y$ with $P \subseteq X$ these three conditions mean that p contains exactly one element.

In the remainder we use the following *point axiom* of [4] which holds for set-theoretic relations, where $\mathcal{P}_v := \{p \mid p \subseteq v \wedge p \text{ is point}\}$ for all vectors v.

Axiom 2.1 *For all sets X we have $\mathsf{L}_{X\mathbf{1}} = \bigcup_{p \in \mathcal{P}_{\mathsf{L}_{X\mathbf{1}}}} p$.*

Additionally we have the following lemma which states that this property can be generalised for arbitrary vectors (see [4]).

Lemma 2.1 *If $v : X \leftrightarrow \mathbf{1}$ is a vector, then $v = \bigcup_{p \in \mathcal{P}_v} p$.* □

3 Cardinality of Relations

In [7] Kawahara discusses the cardinality of set-theoretic relations. The main result is a characterisation of the cardinalities of relations. Considering the properties of this characterisation as axiomatic specification of the cardinality operation $|\cdot|$ this leads to the following definition:

Definition 3.1 *For all relations R we denote its cardinality by $|R|$. The following axioms specify the meaning of the cardinality operation, where Q, R and S are arbitrary relations with appropriate types:*

(C1) If R is finite, then $|R| \in \mathbb{N}$ and $|R| = 0$ iff $R = \mathsf{O}$.
(C2) $|R| = |R^\mathsf{T}|$.
(C3) If R and S are finite, then $|R \cup S| = |R| + |S| - |R \cap S|$.
(C4) If Q is univalent, then $|R \cap Q^\mathsf{T};S| \leq |Q;R \cap S|$ and $|Q \cap S;R^\mathsf{T}| \leq |Q;R \cap S|$.
(C5) $|\mathsf{I}_{\mathbf{11}}| = 1$.

In (C1) and (C3) the relations in question are assumed to be finite so that the cardinality $|R|$ can be regarded as a natural number, in (C2) and (C4) the notation $|R| = |S|$ (respectively $|R| \leq |S|$) is equivalent to the fact that there exists a bijection between R and S (respectively an injection from R to S) and (C5) says that the identity relation on the singleton set $\mathbf{1}$ consists of precisely one pair. Throughout this paper we assume in case of an expression $|R|$ the sets of R's type to be finite and thereby $|R| \in \mathbb{N}$.

Based on the above axioms in [7] a lot of laws for the cardinality operation are derived in a purely calculational manner. For example, from the axioms (C1) and (C3) we get the monotonocity of the cardinality operation, i.e., that $R \subseteq S$ implies $|R| \leq |S|$. Another fact we use in the remainder is following (see [7]):

Lemma 3.1 *If $R : X \leftrightarrow Y$ is univalent and $S : Y \leftrightarrow Z$ is a mapping, then $|R;S| = |R|$.* □

Next, we consider the cardinality of points of type $X \leftrightarrow \mathbf{1}$ and vectors by using only the mentioned cardinality axioms and the presented consequences of them. The next lemma states that a point in deed contains exactly one element.

Lemma 3.2 *If $p : X \leftrightarrow \mathbf{1}$ is a point, then $|p| = 1$.*

Proof. Using cardinality axioms (C2) and (C5) and Lemma 3.1 ($\mathsf{I}_{\mathbf{11}}$ is univalent and $p^\mathsf{T} : \mathbf{1} \leftrightarrow X$ is a mapping), we have the folliwng calculation:

$$|p| = |p^\mathsf{T}| = |\mathsf{I}_{\mathbf{11}};p^\mathsf{T}| = |\mathsf{I}_{\mathbf{11}}| = 1. \qquad \square$$

This lemma allows to show that the cardinality of a vector with target $\mathbf{1}$ equals the cardinality of the set of all points it contains:

Lemma 3.3 *For all $v : X \leftrightarrow \mathbf{1}$ we have $|v| = |\mathcal{P}_v|$.*

Proof. Because of Lemma 2.1, cardinality axioms (C3) and (C1) (the points of $\mathcal{P}_v$ are pair-wise disjoint) and Lemma 3.2 we obtain the claim by

$$|v| = |\bigcup_{p \in \mathcal{P}_v} p| = \sum_{p \in \mathcal{P}_v} |p| = |\mathcal{P}_v|. \qquad \square$$

4 Approximation of Maximum Independent Sets

In this section we use the notions and results of the previous sections to formally verify a relational version of the approximation algorithm of Wei for maximum independent sets (see [16]).

We assume an undirected and loop-free graph $g = (X, E)$ to be given, where the set X of vertices is non-empty and finite. We model g by an *adjacency relation* $R : X \leftrightarrow X$, that is defined by $(x, y) \in R$ iff $\{x, y\} \in E$, for all $x, y \in X$. Due to this definition R is irreflexive and symmetric. The relation R is taken as input for the relational program we want to show as correct. Since the approximation bound depends on the degrees of the vertices, we additionally assume that the maximum degree of g is $k \in \mathbb{N}$. This causes to the conjunction of the following three formulae as pre-condition *Pre*(R, k):

$$R \subseteq \bar{\mathsf{I}} \qquad R = R^{\mathsf{T}} \qquad k = \max\{|R;p| \mid p \in \mathcal{P}_{\mathsf{L}_{X\mathbf{1}}}\}$$

An *independent set* (or *stable set*) of g is a set of vertices S such that $\{x, y\} \notin E$, for all $x, y \in S$. It can be easily derived that a vector $s : X \leftrightarrow \mathbf{1}$ models an independent set with respect to the adjacency relation R iff $R;s \subseteq \overline{s}$. We want to show that our program has approximation bound $k + 1$. So, the post-condition *Post*(R, k, s) is the conjunction of the following two formulae:

$$R;s \subseteq \overline{s} \qquad \forall\, t : X \leftrightarrow \mathbf{1} \bullet R;t \subseteq \overline{t} \Rightarrow |t| \leq |s|(k+1)$$

In the remainder of this section we show that with respect to these specifications the following relational program is totally correct:

```
$s, v := \mathsf{O}, \mathsf{O}_{X\mathbf{1}}$;
while $v \neq \mathsf{L}$ do
    let $p = \mathit{point}(\overline{v})$;
    $s, v := s \cup p, v \cup p \cup R;p$ od          (W)
```

We use the operation *point* that selects deterministically a point such that *point*$(v) \subseteq v$ for all non-empty vectors v. The typing rules of the relational operations in combination with the initialisation of v by $\mathsf{O}_{X\mathbf{1}}$ leads to the typing $s, v, p : X \leftrightarrow \mathbf{1}$ and also $X \leftrightarrow \mathbf{1}$ as type of the constant L of the guard of the loop. The vector v is used to collect the vertices contained in the present indepentend set, that is modeled by the vector s, and also their neighbours.

In the remainder the conjunction of the following two formulae is used as loop invariant *Inv*(R, k, s, v):

$$(1)\ (R \cap v;v^{\mathsf{T}});s \subseteq \overline{s} \qquad (2)\ R;s \cup s = v$$

Here formula (1) is a generalisation of the formula $R;s \subseteq \overline{s}$ of the post-condition *Post*(R, k, s) and formula (2) is simply an auxiliary formula saying that v models the union of the set modeled by s with its neighbours.

We now prove the four proof obligations of assertion-based verification with respect to the above speficied pre- and post-condition. We start with the establishment of the loop invariant by the initialisation of s and v.

Lemma 4.1 *If $R : X \leftrightarrow X$ and $k \in \mathbb{N}$ with Pre(R, k), then Inv$(R, k, \mathsf{O}, \mathsf{O})$.* □

We omit the trivial proof. With the next lemma we prove the maintainence of the loop invariant.

Lemma 4.2 *Given $R : X \leftrightarrow X$, $s, v : X \leftrightarrow \mathbf{1}$ and $k \in \mathbb{N}$ such that* $\mathsf{Inv}(R, k, s, v)$ *and* $v \neq \mathsf{L}$, *we have* $\mathsf{Inv}(R, k, s \cup p, v \cup p \cup R;p)$, *for all $p \in \mathcal{P}_{\overline{v}}$.*

Proof. First, we verify that the first formula of the loop invariant holds for the new values of s and v, i.e., that $(R \cap (v \cup p \cup R;p);(v \cup p \cup R;p)^{\mathsf{T}});(s \cup p) \subseteq \overline{s \cup p}$. It is easy to see that showing the following four inclusions is sufficent:

$$(R \cap (v \cup p \cup R;p);(v \cup p \cup R;p)^{\mathsf{T}});s \subseteq \overline{s}$$
$$(R \cap (v \cup p \cup R;p);(v \cup p \cup R;p)^{\mathsf{T}});s \subseteq \overline{p}$$
$$(R \cap (v \cup p \cup R;p);(v \cup p \cup R;p)^{\mathsf{T}});p \subseteq \overline{s}$$
$$(R \cap (v \cup p \cup R;p);(v \cup p \cup R;p)^{\mathsf{T}});p \subseteq \overline{p}.$$

Because of (2) we have $R;s \subseteq v$ and $s \subseteq v$ and, moreover, because of $p \subseteq \overline{v}$ we have $R;s \subseteq \overline{p}$ and $s \subseteq \overline{p}$. Furthermore, we get

$$R;s \subseteq \overline{p} \iff R^{\mathsf{T}};p \subseteq \overline{s} \iff R;p \subseteq \overline{s}$$

using one of the Schröder rules in the first and the second formula of the pre-condition in the second step. With these auxiliary facts the second and third of the above inclusions follow immediately. Since the point p is injective and R is irreflexive due to the first formula of the pre-condition, we obtain $R;p \subseteq \overline{p}$, such that the last of the above inclusions holds. Verifying the first inclusion is more comprehensive since the following three inclusions have to be proved:

$$(R \cap (v \cup p \cup R;p);v^{\mathsf{T}});s \subseteq \overline{s}$$
$$(R \cap (v \cup p \cup R;p);p^{\mathsf{T}});s \subseteq \overline{s}$$
$$(R \cap (v \cup p \cup R;p);p^{\mathsf{T}};R^{\mathsf{T}});s \subseteq \overline{s}$$

We omit the proofs of these inclusions since they are very similar to these of the previous inclusions.

The maintenance of the second formula of the loop invariant is easy to prove, since by using (2) we get $R;(s \cup p) \cup (s \cup p) = R;s \cup R;p \cup s \cup p = v \cup p \cup R;p$. □

For the third proof obligation we verify the error-free termination of the program (W). A consequence of the guard of the loop is that each call of the partial operation *point* is defined. For this reason and the assumed finiteness of the set X, it suffices to show that the loop terminates, i.e., that v is strictly enlarged by each execution of the body of the loop.

Lemma 4.3 *If $v : X \leftrightarrow \mathbf{1}$ with $v \neq \mathsf{L}$, we have $v \subset v \cup p \cup R;p$, for all $p \in \mathcal{P}_{\overline{v}}$.*

Proof. Since $v \subseteq v \cup p \cup R;p$ holds obviously, we show $v \neq v \cup p \cup R;p$ by contradiction. We start with

$$v = v \cup p \cup R;p \iff p \cup R;p \subseteq v \implies p \subseteq v.$$

The last inclusion and the assumption $p \subseteq \overline{v}$ imply $p = \mathsf{O}$, but this contradicts the fact that points are non-empty. □

Finally, we consider the last proof obligation, i.e., that if $v = \mathsf{L}$ holds, then the loop invariant implies the post-condition. Therefore, we also need the precondition, in particular the maximum-degree condition, for the proof.

Lemma 4.4 *Given $R : X \leftrightarrow X$, $k \in \mathbb{N}$ and $s, v : X \leftrightarrow \mathbf{1}$ such that $\mathit{Pre}(R, k)$, $v = \mathsf{L}$ and $\mathit{Inv}(R, k, s, v)$, we have $\mathit{Post}(R, k, s)$.*

Proof. Formula (1) of the loop invariant $\mathit{Inv}(R, k, s, v)$ and $v = \mathsf{L}$ yield $R;s \subseteq \overline{s}$, which is the first formula of $\mathit{Post}(R, k, s)$.

To verify the second formula of $\mathit{Post}(R, k, s)$, let $t : X \leftrightarrow \mathbf{1}$ be an arbitrary vector such that $R;t \subseteq \overline{t}$. Then we can calculate as follows:

$$\begin{aligned}
|t| &\le |\mathsf{L}_{X\mathbf{1}}| && t : X \leftrightarrow \mathbf{1}, \text{monotonicity cardinality}\\
&= |v| && \text{since } v = \mathsf{L}_{X\mathbf{1}}\\
&= |R;s \cup s| && \text{formula (2) of } \mathit{Inv}(R, k, s, v)\\
&\le |R;s| + |s| && \text{cardinality axiom (C3)}\\
&= |R;\textstyle\bigcup_{p \in \mathcal{P}_s} p| + |s| && \text{by Lemma 2.1}\\
&= |s| + |\textstyle\bigcup_{p \in \mathcal{P}_s} R;p| &&\\
&\le |s| + \textstyle\sum_{p \in \mathcal{P}_s} |R;p| && \mathcal{P}_s \text{ finite, cardinality axiom (C3)}\\
&\le |s| + \textstyle\sum_{p \in \mathcal{P}_s} k && \text{second formula of } \mathit{Pre}(R, k)\\
&= |s| + k|s| && \text{by Lemma 3.3}\\
&= (k + 1)|s| && \Box
\end{aligned}$$

5 Conclusion and Future Work

By modelling graphs via adjacency relations we developed a relational program based on Wei's algorithm for approximating maximum independent sets in graphs. Therefore, we used vectors and points to model subsets of the vertices. Based on Kawahara's characterisation of the cardinality of relations and further consequences of it we were able to prove facts about their cardinality. In the following, we proved the correctness of the developed program by classical reasoning about the specified pre- and postconditions and loop-invariant. Especially the approximation bound was proved in a purely calculational manner by using Kawahara's and our results about the cardinality of relations.

As future work we plan an exhaustive investigation of the cardinatily operation. We hope to come upon useful laws which can be applied, for example, in the context of correctness proofs of further approximation algorithms. Due to the positive experiences we gained with theorem prover with regard to program verification, we plan an embedding of the cardinality operation in existing libraries for relation algebra as for Isabelle/HOL (see [1]) or Coq (see [10]).

Acknowledgement. I thank P. Höfner and, in particular, R. Berghammer for drawing my attention to this particular research topic and valuable discussions and comments.

References

1. Armstrong, A., Foster, S., Struth, G., Weber, T.: Relation algebra. Archive of Formal Proofs, 2014. `http://afp.sf.net/entries/Relation_Algebra.shtml`
2. Cormen, T.H., Leiserson, C.E., Rivest, R.L.: Introduction to algorithms. The MIT Press (1990)
3. Francez, N.: Program verification. Addison-Wesley (1992)
4. Furusawa, H.: Algebraic formalisations of fuzzy relations and their representation theorems. Ph.D. thesis, Department of Informatics, Kyushu University (1998)
5. Gries, D.: The science of programming. Springer (1981)
6. Höfner, P., Struth, G.: On automating the calculus of relations. In: Armando, A., Baumgartner, P., Dowek, G. (eds.) Automated Reasoning. LNAI, vol. 5195, pp. 50-66. Springer (2008)
7. Kawahara, Y.: On the cardinality of relations. In: Schmidt, R.A. (ed.): Relations and Kleene Algebra in Computer Science. LNCS, vol. 4136, pp. 251-265. Springer (2006)
8. Maddux, R.: Relation algebras. In: Brink, C., Kahl, W., Schmidt, G. (eds.): Relational Methods in Computer Science. Advances in Computing Science, pp. 22-38. Springer (1997)
9. Maddux, R.: Relation algebras. Studies in Logic and the Foundations of Mathematics, vol. 150. Elsevier (2006)
10. Pous, D.: Relation algebra and KAT in Coq. `http://perso.ens-lyon.fr/damien.pous/ra/`
11. Schmidt, G., Ströhlein, T.: Relation algebras: Concept of points and representability. Discrete Mathematics 34, 83-97 (1985)
12. Schmidt, G., Ströhlein, T.: Relations and graphs, Discrete mathematics for computer scientists, EATCS Monographs on Theoretical Computer Science. Springer (1993)
13. Schmidt, G.: Relational mathematics. Encyclopedia of Mathematics and its Applications, vol. 132. Cambridge University Press (2010)
14. Tarski, A.: On the calculus of relations. Journal of Symbolic Logic 6, 73-89 (1941)
15. Tarski, A., Givant, S.: A formalization of set theory without variables. Colloquium Publications 41, American Mathematical Society (1987)
16. Wei, V.K.: A lower bound for the stability number of a simple graph. Bell Lab. Tech. Memor. 81-11217-9 (1981)