The Journal of Symbolic Logic
Volume 87, Number 3, September 2022

FIRST-ORDER AXIOMATISATIONS OF REPRESENTABLE RELATION ALGEBRAS NEED FORMULAS OF UNBOUNDED QUANTIFIER DEPTH

ROB EGROT AND ROBIN HIRSCH

Abstract. Using a variation of the rainbow construction and various pebble and colouring games, we prove that RRA, the class of all *representable relation algebras*, cannot be axiomatised by any first-order relation algebra theory of bounded quantifier depth. We also prove that the class At(RRA) of atom structures of representable, atomic relation algebras cannot be defined by any set of sentences in the language of RA atom structures that uses only a finite number of variables.

§1. Introduction. Relation algebras were introduced by Tarski and various coworkers during the 1940s, with the modern definition of a relation algebra being given in [2, 17]. Tarski provided some motivation for this project in the earlier [22], where he noted that though the theory of binary relations was by this time of 'universally recognized' significance, it was then no more developed than it had been at the end of the 19th century. Much more on the historical development of the calculus of relations and relation algebras can be found in [20]. Relation algebras are of theoretical interest as they provide an elegant formalism for the calculus of relations, which is itself an adequate foundational framework for set theory, and thus for mathematics itself (see [24] for the extensive details, and [6, 7] for readable summaries). Relation algebraic methods have also been useful in proving metamathematical results, for example that for all $n \geq 3$ there are sentences involving only three variables whose formal proofs require n variables [13]. In addition, relation algebras and their generalisations have numerous practical applications in computer science, for example in verification [8], computation tasks involving finite topologies [1], and navigation of XML documents [5], to name just a few.

A relation algebra as per Tarski's formulation is an algebraic structure axiomatised by a certain finite set of equations (see the next section for details). In other words, the class of relation algebras is a finitely based variety, which is denoted RA. As discussed above, RA emerged as an attempt to capture properties of binary relations, so it is natural to ask whether this attempt was successful. The answer turns out to be 'not entirely', because there are algebras in RA which do not correspond to concrete systems of binary relations [18]. These algebras are said to not be *representable*, and we provide a formal account of what exactly this means in the next section.

Received August 5, 2020.
2020 *Mathematics Subject Classification*. 03G15, 05C90.
Key words and phrases. RRA, representable relation algebras, finite variable axiomatisation, bounded quantifier depth axiomatisation.

© The Author(s), 2021. Published by Cambridge University Press on behalf of The Association for Symbolic Logic
0022-4812/22/8703-0019
DOI:10.1017/jsl.2021.88

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

So Tarski's axiomatisation of RA was in a sense a failure, but it also turns out to be a remarkable success, for the following reason. Relation algebra equations correspond exactly with first-order statements about binary relations that can be stated using at most three variables (see [24, Theorems 3.9(viii)(ix)] for a proof), and, moreover, a relation algebra equation is valid in RA if and only if the corresponding first-order sentence is *provable* (in classical proof systems) using at most four variables [19, Theorem 24]. So Tarski's axioms neatly capture an intuitively meaningful fragment of the first-order theory of binary relations. Furthermore, an axiomatisation such that the '4' in the statement above could be replaced by '5' would require an infinite number of additional axioms (this result seems to lack a precise statement in the literature, but it can be pieced together from the material in e.g., [11, Section 6]).

Nevertheless, from the perspective of capturing the true properties of binary relations, the class of main interest is that of the representable relation algebras (RRA). Unfortunately, RRA is a difficult class to axiomatise. While it is a variety [23], and is even definable by a recursively enumerable equational theory (see [12, Theorem 8.4] for an example of such an equational theory), it is known that no finite set of relation algebra formulas can define it [21]. Indeed, no set (even infinite) of equations with finitely many variables can define it [16, Theorem 3.5.6], and any axiomatisation of RRA must involve infinitely many non-canonical formulas [14].

In this paper, we extend Monk and Jónsson's negative results by showing that there is no upper bound on the depth of quantifiers needed in an axiomatisation of RRA. The strategy of the paper is based on the well known fact that an atomic relation algebra can be defined from its set of atoms by describing how they interact with each other (this is known as providing an *atom structure*). Since complicated relation algebras can be generated from relatively simple atom structures, this is often an efficient way to construct relation algebras with desirable properties.

Here our goals are achieved by constructing, for each $n < \omega$, two relation algebras, $\mathcal{A}_n$ and $\mathcal{B}_n$, with one representable but not the other, that cannot be distinguished by any relation algebra formula of quantifier-depth at most n. For this construction we use a novel variation of the 'rainbow construction' introduced in [9] (another version of this construction is discussed in [12, Chapter 16]). Background on atom structures can be found in the next section, and the definition of the rainbow construction to be used here is given in Section 3.

Another advantage of atom structures is that games played on what are known as *atomic networks* can be used to show the associated atomic relation algebra is completely representable. We take advantage of this in Section 3 to describe the circumstances in which the algebra constructed from the modified rainbow construction is representable (Theorem 3.2). As an immediate corollary to this we are able to prove that the class of atom structures of representable atomic relation algebras cannot be defined by any set of sentences in the language of RA atom structures that uses only a finite set of variables (Corollary 3.3).

Later, in Section 4, for $n \in \omega$ we find simple sufficient conditions for two algebras constructed using our rainbow construction to be equivalent with respect to first-order sentences of quantifier depth n. The methods here also use games. First a straightforward adaptation of the well known Ehrenfeucht–Fraïssé pebble game to relation algebras, and also a new colouring game which we call the *Seurat game*

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

for sets, as it is a special case of the Seurat games for directed graphs and binary relational structures introduced in [3, 4].

As mentioned previously, the main result (Theorem 4.6) is then obtained simply by, for each $n \in \omega$, describing the algebras $\mathcal{A}_n$ and $\mathcal{B}_n$. The definition of the modified rainbow construction together with the characterisation of representability provided by Theorem 3.2 and the sufficient condition for quantifier depth n equivalence provided by Corollary 4.5 make checking that the algebras have the required properties straightforward.

We had originally hoped to include in this paper a proof that no theory involving only finitely many variables (but with potentially unbounded quantifier depth) could define RRA. An obvious approach would again be to construct (for $n < \omega$) two relation algebras, one in RRA one not, that cannot be distinguished by any n-variable formula. However, many relation algebra properties may be expressed with a finitely bounded number of variables. For example, with just two reusable variables we can write a formula that holds precisely on those relation algebras that have a given finite cardinality (see the discussion at the end of Section 3). The algebras $\mathcal{A}_n, \mathcal{B}_n$ used in this paper for the quantifier depth result as described above have different finite cardinalities, and so they are distinguishable with just two variables, making them unsuitable for the task.

It remains unknown whether a finite variable axiomatisation of RRA exists. A suggestion for how to approach and perhaps solve this problem was given in [12, Problem 1], however the suggested approach was faulty. An early draft of this paper contained a correction to the strategy from [12]. This correction involves a more complex rainbow construction, a version of the Seurat game for general binary relational structures, and a significantly more complicated version of Theorem 3.2. Since the additional technical overhead obscures the argument for the main results here, and since there are independent reasons to believe the corrected strategy might be very difficult, if not impossible, we have moved this material and more detailed discussion to [3].

§2. Technical background.

2.1. Relation algebra, representation, complete representation. A *relation algebra* $\mathcal{A} = (A, 0, 1, +, -, 1', \smile, ;)$ consists of a set A with elements $0, 1, 1'$, unary functions $-, \smile$ and binary functions $+, ;$ over A, such that

(i) $(A, 0, 1, +, -)$ is a boolean algebra,
(ii) $(A, 1', \smile, ;)$ is a convoluted monoid,
(iii) $\smile$ and $;$ are normal operators, and
(iv) the Piercean law $a\smile; (- (a; b)) \leq - b$ holds.

Since these axioms are all equations, the class **RA** of all relation algebras is a finitely based equational variety.

Let E be any equivalence relation over base X. The *proper relation algebra* $\mathcal{P}(E) = (\wp(E), \emptyset, E, \cup, \setminus_E, Id_X, \smile, ;)$ consists of all subsets of E, with the identity $Id_X = \{(x, x) : x \in X\}$, converse defined by $a\smile = \{(y, x) : (x, y) \in a\}$, and composition defined by $a; b = \{(x, y) : \exists z\ ((x, z) \in a \wedge (z, y) \in b\}$, where $a, b \subseteq E$. The axioms of relation algebra are mostly easy to verify in a proper relation algebra, perhaps the Peircean law would take slightly longer.

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

An embedding $\theta : \mathcal{A} \to \mathcal{P}(E)$ from an abstract relation algebra $\mathcal{A}$ into a proper relation algebra $\mathcal{P}(E)$ is called a *representation* of $\mathcal{A}$. For $a \in \mathcal{A}$ we write a^θ for the image of a under θ, which is a binary relation. The class of all representable relation algebras (RRA) turns out to be a proper subclass of RA (not every relation algebra is representable). If a representation θ preserves arbitrary suprema wherever they exist it is called a *complete representation*. For finite relation algebras, every representation is trivially complete, as there are no non-finite suprema, but this is not true in general.

For each $(x, y) \in 1^\theta$, the set $\{a \in \mathcal{A} : (x, y) \in a^\theta\}$ is an ultrafilter of the boolean part of $\mathcal{A}$. An ultrafilter is *principal* if it includes its infimum, and is *non-principal* otherwise. In the former case the infimum must be an *atom* (minimal non-zero element of $\mathcal{A}$), in the latter case the infimum can be shown to be zero. A representation θ is *atomic* if for all $(x, y) \in 1^\theta$ there is an atom a such that $(x, y) \in a^\theta$. By [12, Theorem 2.21] a representation is complete if and only if it is atomic. So, if θ is a complete representation and $(x, y) \in 1^\theta$, there is a (necessarily unique) atom a such that $(x, y) \in a^\theta$. We denote this atom by $(x, y)_\theta$, so $(x, y) \in ((x, y)_\theta)^\theta$, for $(x, y) \in 1^\theta$.

2.2. Atom structure, Peircean transforms. A boolean algebra (possibly with additional operations, e.g., a relation algebra) is *atomic* if every non-zero element is above an atom. This is equivalent to saying that every element is the supremum of the set of atoms below it. Every finite boolean algebra is atomic. Using the relation algebra axioms, it can be shown that that the operators $^\smile, ;$ are *completely additive*, i.e., if the supremum $\bigvee S$ of a subset S of a relation algebra exists then $(\bigvee S)^\smile$ is the supremum of the set $\{s^\smile : s \in S\}$, and for any a in the algebra, $a; \bigvee S$ is the supremum of $\{a; s : s \in S\}$ and $\bigvee S; a$ is the supremum of $\{s; a : s \in S\}$ (see e.g., [12, p. 109] for a proof).

It follows easily from completeness of $^\smile$ and relation algebra condition (ii) that $x^\smile$ is an atom if and only if x is. Hence, for an atomic relation algebra $\mathcal{A}$, each operator is determined by its restriction to atoms. This information is conveyed by its *atom structure*, which consists of the set of atoms, the set of atoms below the identity, the set of pairs $(a, a^\smile)$ where a ranges over atoms, and the set of *consistent* triples of atoms (a, b, c) where $a; b \geq c$. (Beware, [12] uses the consistency condition $a; b \geq c^\smile$, so (a, b, c) is consistent according to our definition here if and only if $(a, b, c^\smile)$ is consistent by the [12] definition.) Instead of giving the set of consistent triples of atoms it is often convenient to specify the *forbidden* triples of atoms (a, b, c)—those where $a; b \cdot c = 0$.

A consequence of the relation algebra axioms, in particular the Peircean law, is that if (a, b, c) is a forbidden triple of atoms then all six so-called *Peircean transforms* of (a, b, c), i.e.,

$$(a, b, c), (a^\smile, c, b), (c, b^\smile, a), (b, c^\smile, a^\smile), (c^\smile, a, b^\smile), (b^\smile, a^\smile, c^\smile)$$

are also are forbidden.

Every atomic relation algebra embeds into the *complex algebra* of its atom structure, a uniquely determined relation algebra whose elements are arbitrary sets of atoms (see [12, Section 2.7.2]), and a finite relation algebra is isomorphic to the complex algebra of its atom structure. It is often convenient to define a relation

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

algebra by giving an atom structure, i.e., a set of atoms, the subset of atoms below the identity, the consistent/forbidden triples and the converses. This is what we will do in Section 3, for example.

2.3. Atomic networks, games. Given an atomic relation algebra $\mathcal{A}$ with atoms $At(\mathcal{A})$, an atomic labelling $N = (nodes(N), \lambda)$ consists of a set of nodes and an edge labelling function $\lambda : nodes(N) \times nodes(N) \to At(\mathcal{A})$. If

(i) $\lambda(x, x) \leq 1'$,
(ii) $\lambda(y, x) = (\lambda(x, y))^{\smile}$, and
(iii) $(\lambda(x, y), \lambda(y, z), \lambda(x, z))$ is not forbidden, for all $x, y, z \in nodes(N)$,

the atomic labelling is *coherent* and we call it an *atomic network*. Given atomic labellings $N = (nodes(N), \lambda)$, $N' = (nodes(N'), \lambda')$ we say N' is an extension of N, and write $N \subseteq N'$, if $nodes(N) \subseteq nodes(N')$ and the restriction of λ' to $nodes(N) \times nodes(N)$ equals λ.

Henceforth, we may write N for the name of the atomic labelling, its set of nodes and also the edge labelling function, distinguishing by context, e.g., $x \in N$ means x is a node of N and $N(x, y)$ denotes the atom labelling the edge (x, y).

The atomic network game $\mathsf{N}(\mathcal{A})$ is played by two players, $\forall$ and $\exists$, and has ω rounds. In each round an atomic labelling N_i is played $(i < \omega)$. In the initial round, $\forall$ picks any atom $a \in At(\mathcal{A})$ and $\exists$ must respond with an atomic labelling N_0 with nodes x, y where $N_0(x, y) = a$ (the nodes x, y may be distinct or not). In round $i > 0$ suppose N_{i-1} was the last atomic labelling played. Then $\forall$ picks any two nodes $x, y \in N_{i-1}$ and any pair of atoms $a, b \in At(\mathcal{A})$ such that $(a, b, N_{i-1}(x, y))$ is not forbidden. We denote his move (x, y, a, b). In response, $\exists$ must play $N_i \supseteq N_{i-1}$ where there is $z \in N_i$ such that $N_i(x, z) = a$, $N_i(z, y) = b$. If in any round the atomic labelling N_i fails to be coherent (so fails to be an atomic network) then $\forall$ wins the play. If $\forall$ does not win in any of the ω rounds then $\exists$ wins.

If the current atomic labelling is N, an $\forall$-move (x, y, a, b) is *trivial* if there is $z \in N$ with $N(x, z) = a$, $N(z, y) = b$. $\exists$ can always respond to a trivial move by playing $N' = N$. We will assume that $\forall$ avoids trivial moves. More details on atomic networks and games can be found in [12, Chapter 11]. They are important here as, in certain circumstances, the chain of networks created during a play of $\mathsf{N}(\mathcal{A})$ corresponds to a representation of $\mathcal{A}$. This is formalised by the following result.

PROPOSITION 2.1. *Let $\mathcal{A}$ be an atomic relation algebra with at most countably many atoms. Then $\mathcal{A}$ is completely representable if and only if $\exists$ has a winning strategy in the atomic network game* $\mathsf{N}(\mathcal{A})$.

PROOF. See [12, Theorem 11.7]. ⊣

This proposition may be generalised to atomic relation algebras with uncountably many atoms, but the atomic network game has to run transfinitely with as many rounds as atoms, taking limits of atomic networks at rounds indexed by limit ordinals. Working out the details of this is given as [12, Exercise 11.4.3], but we do not not need the general result here.

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

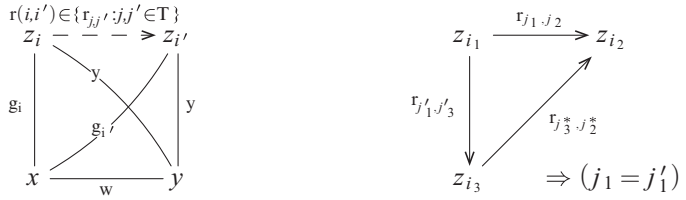


FIGURE 1. Defining an injection from a representation.

§3. The rainbow construction. Given two sets S, T we define an atomic relation algebra $\mathcal{B}_{S,T}$ by defining its atom structure. The atoms are

$$\{1', \mathsf{b}, \mathsf{w}, \mathsf{y}\} \cup \{\mathsf{g}_i : i \in S\} \cup \{\mathsf{r}_{j,j'} : j, j' \in T\}.$$

The non-identity atoms are considered to be black, white, yellow, green or red. All atoms are self-converse, except $\mathsf{r}_{j,j'}^{\smile} = \mathsf{r}_{j',j}$. Forbidden triples of atoms are Peircean transforms of

(I) $(1', a, b)$ where $a \neq b$
(II) $(\mathsf{g}_i, \mathsf{g}_{i'}, \mathsf{g}_{i''}), (\mathsf{g}_i, \mathsf{g}_{i'}, \mathsf{w})$, any $i, i', i'' \in S$
(III) $(\mathsf{y}, \mathsf{y}, \mathsf{y}), (\mathsf{y}, \mathsf{y}, \mathsf{b})$
(IV) $(\mathsf{r}_{j_1,j_2}, \mathsf{r}_{j'_2,j'_3}, \mathsf{r}_{j_1^*,j_3^*})$, unless $j_1 = j_1^*,\ j_2 = j'_2,\ j'_3 = j_3^*$.
(V) $(\mathsf{g}_i, \mathsf{g}_i, \mathsf{r}_{j,j'}),\ (\mathsf{g}_i, \mathsf{g}_{i'}, \mathsf{r}_{j,j})$, any $i, i' \in S,\ j, j' \in T$

DEFINITION 3.1. The relation algebra $\mathcal{B}_{S,T}$ is the complex algebra of this atom structure.

This relation algebra can be obtained from the rainbow algebra of [12, Section 16.2] by regarding S, T as 'binary structures' in a relational language with no predicates except equality, by deleting all white atoms w_X where $X \subseteq S$ has at most two elements, and by forbidding Peircean transforms of $(\mathsf{g}_i, \mathsf{g}_i, \mathsf{r}_{j,j})$ for $i \in S,\ j \in T$.

THEOREM 3.2. *Let S, T be finite sets with $|S| \geq 2$. Then the algebra $\mathcal{B}_{S,T}$ is representable if and only if $|S| \leq |T|$.*

PROOF. Suppose first that $\mathcal{B}_{S,T}$ is representable, and that θ is a representation. Since $\mathcal{B}_{S,T}$ is finite, we know that θ is a complete representation. As w is an atom, and hence non-zero, there must be points x, y in the base of the representation θ such that $(x, y) \in \mathsf{w}^\theta$ (equivalently, $(x, y)_\theta = \mathsf{w}$, recalling the notation from the end of Section 2.1). For each $i \in S$ the triple $(\mathsf{g}_i, \mathsf{y}, \mathsf{w})$ is not forbidden, so $\mathsf{w} \leq (\mathsf{g}_i; \mathsf{y})$, and there must be a point z_i with $(x, z_i) \in \mathsf{g}_i^\theta$ and $(z_i, y) \in \mathsf{y}^\theta$. For distinct $i, i' \in S$ the atom $(z_i, z_{i'})_\theta$ must be red, by forbidden triples (I)–(III). Call this atom $\mathsf{r}(i, i')$, for some element of $\{\mathsf{r}_{j,j'} : j, j' \in T\}$, and see the first part of Figure 1. For distinct $i_1, i_2, i_3 \in S$ the first subscript of the red atom $\mathsf{r}(i_1, i_2)$ is the same as that of $\mathsf{r}(i_1, i_3)$ by forbidden triple (IV), see the second part of Figure 1. Fixing $i \in S$, it follows that each $i' \in S$ corresponds to a point $\iota(i') \in T$ such that $(z_i, z_{i'})_\theta = \mathsf{r}_{\iota(i),\iota(i')}$. As we have just noted, examining the second part of Figure 1 we see that $\iota(i)$ is uniquely specified. Moreover, by the second forbidden triple of (V), all indices $\iota(i')$ (for $i' \in S$) are distinct, hence the map $\iota : S \to T$ is an injection, and so $|S| \leq |T|$.

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

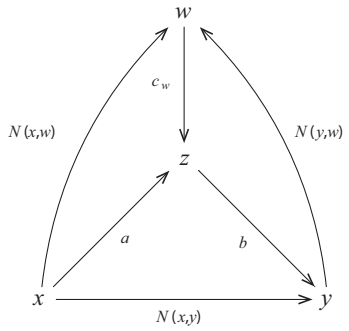


FIGURE 2. For each $w \in N \setminus \{x, y\}$, $\exists$ must assign the label $c_w = N'(w, z)$.

For the converse, suppose $2 \leq |S| \leq |T|$. We aim to show that $\mathcal{B}_{S,T}$ is completely representable. As discussed above, it is sufficient to show that $\exists$ has a winning strategy in the complete representation game for $\mathcal{B}_{S,T}$. We assume that $\forall$ never plays a trivial move, as $\exists$ can safely respond by leaving the current network unchanged. In particular we assume he never plays (x, y, a, b) if $a = 1'$ or $b = 1'$.

Suppose then that in a play of this game the current atomic labelling is N, and $\forall$ has just played the move (x, y, a, b). In her response, $\exists$ adds a new node z to N to create $N' = N \cup \{z\}$, and assigns labels to every node $w \in N \setminus \{x, y\}$ (see Figure 2). Her strategy is as follows:

(a) If $N(w, x)$ and a are not both green, and $N(w, y), b$ are not both green, she sets $N'(w, z) = \mathsf{w}$.

(b) If $N(w, x), a$ are both green but $N(w, y), b$ are not both yellow, or if $N(w, y), b$ are both green but $N(w, x), a$ are not both yellow, she sets $N'(w, z) = \mathsf{b}$.

(c) The remaining case is where, for some $i, i' \in S$, we have $N(x, w) = \mathsf{g}_i$, $a = \mathsf{g}_{i'}$, $N(w, y) = b = \mathsf{y}$ (or similar with x, y swapped). Since we are assuming that $\forall$ makes no trivial moves, we can assume $i \neq i'$. In this case $\exists$ can only use a red label for $N'(w, z)$, as the other colours are precluded by the forbidden triple rules. To complete the definition of the strategy it remains to provide the two subscripts $j, j' \in T$ of the red atom she chooses for $N'(w, z)$.

To help us here we introduce some notation, and make a few preliminary observations. Given an atomic network N for $\mathcal{B}_{S,T}$ and nodes $x, y \in \mathsf{nodes}(N)$ let

$$R_N(x, y) = \{z \in \mathsf{nodes}(N) : N(x, z) \text{ is green and } N(y, z) = \mathsf{y}\}.$$

Observe that $R_N(x, y)$ depends only on the green and yellow edge labels of N. A set of nodes of a network where every edge between distinct nodes has a red label is called a *red clique*, and it follows from the forbidden triple rules that $R_N(x, y)$ must be a red clique.

By definition, part (c) of $\exists$'s strategy is relevant precisely when the new node z is added to non-empty $R_N(x, y)$. The key idea is that every time a red clique of form $R_N(x, y)$ attains size two in the current network, $\exists$ will define

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

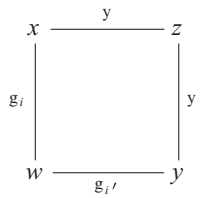


FIGURE 3. Creating a red clique of size 2.

an injection $h_{xy} : S \to T$ and this will be used to guide her choice of labels whenever a new z is added to $R_N(x, y)$ in later rounds. This strategy will be well defined because, as $\exists$ never uses green or yellow labels, the only way the new node z can be added to $R_N(x, y)$ is if a is green and b is yellow (or vice versa), in which case x and y are specified uniquely. Thus, given z, there is only one relevant h_{xy}.

Let N be the current atomic network in a play of the atomic representation game. Assume inductively: if $|R_N(x, y)| \geq 2$ then there is an injection $h_{xy} : S \to T$ such that for all $w \neq w' \in R_N(x, y)$, $i, i' \in S$, we have

$$(N(x, w) = \mathsf{g}_i \wedge N(x, w') = \mathsf{g}_{i'}) \Rightarrow (i \neq i' \wedge N(w, w') = \mathsf{r}_{h_{xy}(i), h_{xy}(i')}). \quad (1)$$

Initially all red cliques $R_{N_0}(x', y')$ are empty so the induction hypothesis holds.

There are three ways to extend the network N to N' so that a new red clique of form $R_{N'}(x, y)$ is created with $|R_{N'}(x, y)| = 2$:

(i) The first is if $|R_N(x, y)| = 1$, say $R_N(x, y) = \{w\}$ where $w \in N$, and $\forall$ plays $(x, y, \mathsf{g}_{i'}, \mathsf{y})$ for some $i' \in S$. In this case, $R_{N'}(x, y) = \{w, z\}$, with $N(x, w) = \mathsf{g}_i$ for some $i \neq i' \in S$. Here $\exists$ defines h_{xy} to be any injection $h_{xy} : S \to T$. Such an injection must exist as we are assuming $|S| \leq |T|$. Then she sets $N'(w, z) = \mathsf{r}_{h_{xy}(i), h_{xy}(i')}$, in accordance with (1).

(ii) The second way such a red clique of size two can be created is where $w \in N$, z is the new node, and $\forall$'s move is $(x, y, \mathsf{y}, \mathsf{y})$, with $N(w, x) = \mathsf{g}_i$ and $N(w, y) = \mathsf{g}_{i'}$ for some $i, i' \in S$ (see Figure 3). Since N is an atomic network and $\forall$'s move is legal, $N(x, y)$ must be red and $i \neq i'$ (using the first forbidden triple of (V)). This creates the red clique $R_{N'}(w, z) = \{x, y\}$ ($\{x, y\}$ was previously a red clique, but not of the form $R_N(x', y')$). Note that x and y can be the only members of $R_{N'}(w, z)$ as $\exists$ never uses yellow labels, so no other network node can have a yellow edge to z. Here the label $N(x, y)$ was defined in N previously, and must be $\mathsf{r}_{j,j'}$ for some $j \neq j' \in T$, as otherwise $\forall$'s move would be illegal. In this case $\exists$ defines h_{wz} to be any injection $S \to T$ with $h_{wz}(i) = j$ and $h_{wz}(i') = j'$.

(iii) The third way is similar to the second, except $\forall$ plays $(x, y, \mathsf{g}_i, \mathsf{g}_{i'})$, and $N(w, x) = N(w, y) = \mathsf{y}$. This is similar to the previous case, and for $\forall$'s move to be legal we must have $N(x, y) = \mathsf{r}_{j,j'}$ for some $j \neq j' \in T$. Again $\exists$ defines h_{wz} to be any injection with $h_{wz}(i) = j$ and $h_{wz}(i') = j'$.

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

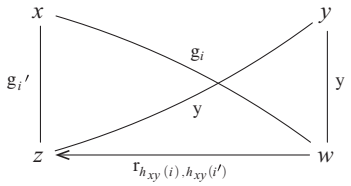


FIGURE 4. The label $N'(w, z)$ in case (c).

The only case where $\exists$ chooses any red labels is when $\forall$'s move is $(x, y, \mathsf{g}_i, \mathsf{y})$ or $(x, y, \mathsf{y}, \mathsf{g}_i)$, for some $i \in S$. So the only way a red clique of form $R_{N'}(x, y)$ with size $k > 2$ can occur is where $|R_N(x, y)| = k - 1$ and $\forall$'s move for the round is $(x, y, \mathsf{g}_i, \mathsf{y})$ for some $i \in S$. Since we are assuming that no trivial moves are played, there is no $w \in R_N(x, y)$ with $N(x, w) = \mathsf{g}_i$. In this case $\exists$ adds a single new node z to the network and includes it in $R_{N'}(x, y)$. Then, according the $\exists$'s strategy, edges (z, w) have red labels if $w \in R_N(x, y)$ but either white or black labels if $w \in N \setminus R_N(x, y)$. Note that $z \in R_{N'}(x', y')$ if and only if $x = x'$ and $y = y'$. Since $k - 1 \geq 2$ we know inductively that there is an injection $h_{xy} : S \to T$, satisfying (1). For $w \in R_N(x, y)$ where $N(x, w) = \mathsf{g}_{i'}$ say, she lets $N(w, z) = \mathsf{r}_{h_{xy}(i'),h_{xy}(i)}$ (see Figure 4). This maintains (1) and completes the definition of $\exists$'s strategy.

We must show that this strategy is a good one by proving that N' is an atomic network, i.e., that the labelling is coherent. We must check that the labeling of each triangle (w, w^*, z) for $w \neq w^* \in N$ is not forbidden. If $\{w, w^*\} = \{x, y\}$ then the triangle is consistent (else $\forall$'s move would be illegal), so without loss of generality we assume that $w \notin \{x, y\}$, and so the label $N'(w, z)$ was just assigned by $\exists$ in this round. Hence $N'(w, z)$ must be either white, black or red, since she only chooses these colours. If $N'(w, z) = \mathsf{w}$, then the only possibility that the triangle (w, w^*, z) could be forbidden comes from (II), but this requires that $N'(w^*, z)$ and $N'(w, w^*)$ be green, and thus that $w^* \in \{x, y\}$. But then the conditions of (a) would not have been met, so $N'(w, z)$ could not be w after all. Similarly, if $N'(w, z) = \mathsf{b}$ then the possibility of violating (III) is ruled out by case (b) conditions.

In the remaining case, $N'(w, z)$ is red, and the only forbidden triples involving red atoms are (IV) and (V). Since $w \notin \{x, y\}$ by assumption, the label $N'(w, z)$ must have been assigned according to part (c) of $\exists$'s strategy, so we assume that $N(x, w) = \mathsf{g}_i$, $a = \mathsf{g}_{i'}$, and $N(y, w) = b = \mathsf{y}$ (the case where x, y are swapped is symmetric). A triangle (w, w^*, z) could only violate forbidden triple (IV) if all three edges were red. For this to happen we must have $N'(x, w^*) = \mathsf{g}_{i^*}$ for some $i^* \in S$, and $N'(y, w^*) = \mathsf{y}$, so $\{w, w^*\} \subseteq R_N(x, y)$. By our induction hypothesis, there is an injection $h_{xy} : S \to T$ satisfying (1) in N. According to part (c) of her strategy, $\exists$ set $N'(w, z) = \mathsf{r}_{h_{xy}(i),h_{xy}(i')}$ and $N'(w^*, z) = \mathsf{r}_{h_{xy}(i^*),h_{xy}(i')}$. It follows that (w, w^*, z) does not violate (IV) (see Figure 5).

For rule(V), the only possible green–green–red triangle incident with z and w is (x, w, z) and the edge labels $(\mathsf{g}_i, \mathsf{g}_{i'}, \mathsf{r}_{h_{xy}(i),h_{xy}(i')})$ do not violate (V), since h is injective. Hence the labelling is coherent, and N' is an atomic network. ⊣

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

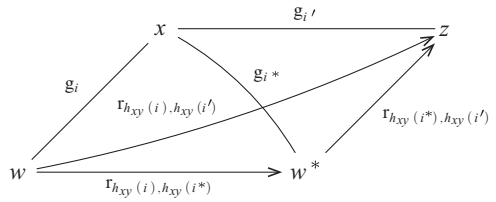


FIGURE 5. The red triangle involving w, w^* and z.

This theorem may be generalised: if S, T are arbitrary sets where T has at least two elements, then $\mathcal{B}_{S,T}$ is completely representable if and only if the cardinality of S is no more than that of T. The left to right implication is proved as above. For the right to left implication we use a transfinitely long atomic network game with as many rounds as $|S| + \omega$, as mentioned above. The relation algebras of interest here are finite, so we leave out a formal proof of the more general result.

The class of atom structures of representable relation algebras is known to be elementary [25]. We can apply the result above to say something about classes of atom structures intermediate between this class and that of the atom structures of completely representable relation algebras.

COROLLARY 3.3. *If $\mathcal{K}$ is a class of relation algebra atom structures including all atom structures of completely representable relation algebras, and contained in the class of atom structures of representable atomic relation algebras, then $\mathcal{K}$ cannot be defined by any theory in the language of RA atom structures using only finitely many atom valued variables.*

PROOF. Given finite sets S, T where $2 \leq |S| < |T|$ consider the algebras $\mathcal{B}_{S,S}$ and $\mathcal{B}_{T,S}$. By Theorem 3.2, the former is completely representable while the latter is not, and, since they are finite, representability is the same as complete representability. Thus the atom structure of $\mathcal{B}_{S,S}$ is in $\mathcal{K}$ and that of $\mathcal{B}_{T,S}$ is not.

However, consider the well-known Ehrenfeucht–Fraïssé game for testing elementary equivalence played over a pair of relational structures using c pebbles. A full description of this game can be found in e.g., [15, Chapter 6], but a brief summary is that, in each round, the player $\forall$ associates a pebble with an element of his choice from one of the structures (again of his choice), and then the player $\exists$ associates the same pebble with an element of her choice from the other structure. This association via pebbles induces a mapping between the two structures, and $\forall$ wins if at any point it fails to be a well-defined isomorphism. The key result is that $\exists$ has a strategy guaranteeing her survival in this game for at least n rounds if and only if the two structures are equivalent with respect to all formulas of quantifier depth at most n using at most c free variables (see e.g., [15, Theorem 6.10]).

In our case, $\exists$ has a winning strategy in the $|S|$-pebble, ω-round Ehrenfeucht Fraïssé-game over the atom structures of $\mathcal{B}_{S,S}$ and $\mathcal{B}_{T,S}$, which we describe now. Whenever $\forall$ picks a non-green atom, $\exists$ picks the corresponding non-green atom in the other atom structure. If $\forall$ places a pebble where another pebble is already placed, then $\exists$ covers the corresponding pebble in the other structure. If $\forall$ picks a green atom

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

not already in play, then $\exists$ picks any green atom in the other atom structure not currently selected. There are always enough green atoms for this, because the game only uses $|S|$ pebbles. This strategy is a winning one because a triple $(\mathsf{g}_i, \mathsf{g}_{i'}, \mathsf{r}_{j,j'})$ is forbidden iff $i = i'$ or $j = j'$, in either atom structure. It follows that the two atom structures agree on all $|S|$-variable formulas. $\dashv$

The corollary shows that the atom structures of $\mathcal{B}_{S,S}$ and $\mathcal{B}_{T,S}$ cannot be distinguished in the language of atom structures restricted to $|S|$ atomic variables. If we use formulas with variables that range over arbitrary elements of a relation algebra, much more can be expressed. Consider, for example, the formula $\phi_k(x)$ with variables x, y of which only x appears free which we will define shortly. It is intended to express that x is above at least k atoms, in an atomic relation algebra. So $\phi_1(x)$ is $\neg(x = 0)$. Recursively, suppose $\phi_k(x)$ and $\phi_k(y)$ have been defined (the variables x and y are swapped throughout in the latter formula), and suppose the formula holds exactly when the free variable denotes an element above at least k atoms. Let $\phi_{k+1}(x)$ be the formula $\exists y(y < x \wedge \phi_k(y))$, where $<$ is the usual shorthand for the formula defining strict order in boolean algebras. If k is the number of atoms in $\mathcal{B}_{T,S}$ then $\exists x \phi_k(x)$ is true in $\mathcal{B}_{T,S}$ but not in $\mathcal{B}_{S,S}$. As well as defining the finite cardinality of an algebra, two variable formulas can express many other properties. Indeed, it is conceivable that any pair of non-isomorphic finite relation algebras can be distinguished by a two variable formula; this remains an open problem.

The proof of Corollary 3.3 is a kind of warm up for the proof of Theorem 4.6, which involves formulas with variables ranging over arbitrary relation algebra elements, and occupies most of the next section.

§4. Axiomatisations of RRA. We will use the following minor variation of the classic Ehrenfeucht–Fraïssé game used in the proof of Corollary 3.3. Given $n < \omega$ and two relation algebras $\mathcal{A}, \mathcal{B}$ we define the n-round equivalence game $\Gamma_n(\mathcal{A}, \mathcal{B}, p)$, where p is a sequence of pairs from $\mathcal{A} \times \mathcal{B}$. The idea of the game is to test whether p could define a function extending to an isomorphism from $\mathcal{A}$ to $\mathcal{B}$.

A *play* $(p_0, \dots, p_n)$ of the game $\Gamma_n(\mathcal{A}, \mathcal{B}, p)$ consists of a sequence of sequences of pairs from $\mathcal{A} \times \mathcal{B}$, with $p_0 = p$, and where each p_{i+1} is obtained by appending a pair $(a, b) \in \mathcal{A} \times \mathcal{B}$ to p_i. These sequences are called *positions*, and $p = p_0$ is the *starting position*.

If $n = 0$ then nothing happens, and the result of the game will depend only on the starting position p. For $n \geq 1$, in round $i < n$, the current position is p_i and $\forall$ chooses either $a \in \mathcal{A}$ or $b \in \mathcal{B}$ as he prefers, and $\exists$ chooses the other element. The position p_{i+1} is obtained by appending the pair (a, b) to p_i. The game $\Gamma_n(\mathcal{A}, \mathcal{B}, \emptyset)$ that starts from the empty position is denoted $\Gamma_n(\mathcal{A}, \mathcal{B})$.

Given a position $q = ((a_0, b_0), \dots, (a_k, b_k))$ arising during a play of $\Gamma_n(\mathcal{A}, \mathcal{B}, p)$, define the sequences $\bar{a} = (a_0, \dots, a_k)$ and $\bar{b} = (b_0, \dots, b_k)$. Let $\mathcal{A}_{\bar{a}}$ and $\mathcal{B}_{\bar{b}}$ denote the subalgebras of $\mathcal{A}$ and $\mathcal{B}$ generated by $\bar{a}$ and $\bar{b}$, respectively. We want to use the sequence q to define a map $\langle q \rangle$ from $\mathcal{A}_{\bar{a}}$ to $\mathcal{B}_{\bar{b}}$. To do this note that elements of $\mathcal{A}_{\bar{a}}$ correspond to terms constructed from elements of $\bar{a}$ and relation algebra constants using relation algebra operations, and similar for elements of $\mathcal{B}_{\bar{b}}$. Given such a term t in $\mathcal{A}_{\bar{a}}$ we want to define $\langle q \rangle(t)$ to be the term in $\mathcal{B}_{\bar{a}}$ obtained by replacing each a_i

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

with b_i and preserving relation algebra constants. If $\langle q \rangle$ fails to be an isomorphism, or fails to be well defined at all, then q is a winning position for $\forall$.

If any position in the play is a winning position for $\forall$ then he wins. If none of the positions are winning positions for $\forall$ then $\exists$ wins. Since a winning position for $\forall$ remains a winning position for $\forall$ after further play, it follows that the winner is determined by the final position p_n. The result is a win for $\exists$ if and only if $\langle p_n \rangle$ is a well-defined isomorphism. Note that if during the course of the game $\forall$ chooses an element he already chose in a previous round, $\exists$ can respond by choosing the element she chose in that round, and the outcome of the game is unchanged. So we assume without loss of generality that $\forall$ never plays these redundant moves.

The value of these Ehrenfeucht–Fraïssé games for relation algebras is given by the following definition and lemma.

DEFINITION 4.1. Let $n < \omega$, let $\mathcal{A}$ and $\mathcal{B}$ be relation algebras, and let $\bar{a} = (a_0, \dots, a_{k-1})$ and $\bar{b} = (b_0, \dots, b_{k-1})$ be tuples from $\mathcal{A}, \mathcal{B}$ respectively, of the same length k. We write

$$(\mathcal{A}, \bar{a}) \equiv_n (\mathcal{B}, \bar{b}),$$

if whenever ϕ is a first-order formula of quantifier depth at most n, with free variables from $\{x_i : i < k\}$, in the language of relation algebras, we have

$$\mathcal{A}, \bar{a} \models \phi \iff \mathcal{B}, \bar{b} \models \phi.$$

When $\bar{a}$ and $\bar{b}$ are empty we just write $\mathcal{A} \equiv_n \mathcal{B}$.

LEMMA 4.2. *Let $n < \omega$, let $\bar{a} = (a_0, \dots, a_{k-1})$, $\bar{b} = (b_0, \dots, b_{k-1})$ be tuples from $\mathcal{A}, \mathcal{B}$ respectively and let $p = ((a_0, b_0), \dots, (a_{k-1}, b_{k-1}))$. Then $\exists$ having a winning strategy in $\Gamma_n(\mathcal{A}, \mathcal{B}, p)$ implies $\mathcal{A}, \bar{a} \equiv_n \mathcal{B}, \bar{b}$.*

PROOF. This is half the well known result for relational signatures used in the proof of Corollary 3.3 (see e.g., [15, Theorem 6.10]). Having functions in the signature blocks the proof of the opposite implication. We induct on n. For the base case, if $\exists$ wins $\Gamma_0(\mathcal{A}, \mathcal{B}, p)$ then p induces an isomorphism from $\mathcal{A}_{\bar{a}}$ to $\mathcal{B}_{\bar{b}}$, hence $(\mathcal{A}, \bar{a}) \equiv_0 (\mathcal{B}, \bar{b})$.

For the inductive step, suppose $\exists$ has a winning strategy in $\Gamma_{n+1}(\mathcal{A}, \mathcal{B}, p)$, and let ϕ be a formula of quantifier depth $n+1$. Note that, for all formulas ϕ_1 and ϕ_2, if $\mathcal{A}$ and $\mathcal{B}$ disagree about $\phi_1 \wedge \phi_2$ they must also disagree about either ϕ_1 or ϕ_2, and if they disagree about $\neg\phi_1$ they must also disagree about ϕ_1. So we can assume without loss of generality that $\phi = \exists x_k \psi$, where ψ is a formula of quantifier depth n. If $\mathcal{A}, \bar{a} \models \exists x_k \psi$ then there is $a_k \in \mathcal{A}$ such that $\mathcal{A}, \bar{a}, a_k \models \psi$. If $\forall$ plays a_k in the game, then since $\exists$ has a winning strategy there is $b_k \in \mathcal{B}$ where $\exists$ has a winning strategy in $\Gamma_n(\mathcal{A}, \mathcal{B}, p')$, where p' is p with (a_k, b_k) appended. Inductively, $\mathcal{B}, \bar{b}, b_k \models \psi$, hence $\mathcal{B}, \bar{b} \models \exists x_k \psi$. Since the argument is symmetric, it follows that $(\mathcal{A}, \bar{a})$ agrees with $(\mathcal{B}, \bar{b})$ on all formulas $\exists x_k \psi$ where ψ has quantifier depth at most n, hence they agree on all formulas of quantifier depth at most $n+1$. By induction, the lemma holds for all $n < \omega$. ⊣

We now define a colouring game played by $\forall$ and $\exists$ over a pair of sets (T, T'). Colours are used to colour *subsets* of T and T', rather than the individual elements used in the pebble games of Corollary 3.3 and Lemma 4.2. Let $n < \omega$ and let T, T'

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

be sets. We define the n-round colouring game $\mathbf{G}_n(T, T')$. If $n = 0$ then the game ends immediately with neither player making a move. For $n \geq 1$, a play of the game is a sequence $((T_0, T'_0), \dots, (T_{n-1}, T'_{n-1}))$ where $T_i \subseteq T$, $T'_i \subseteq T'$, for $i < n$. The positions in a play are its initial segments. The initial position p_0 is the empty sequence. For $n > 0$, at the start of round $i < n$ the position is denoted p_i. Then $\forall$ chooses either a subset $T_i \subseteq T$ or a subset $T'_i \subseteq T'$, as he prefers, and $\exists$ chooses the other subset. The position is then updated to p_{i+1}, which is p_i with the pair (T_i, T'_i) appended. Note that the final position p_n is the full play $((T_0, T'_0), \dots, (T_{n-1}, T'_{n-1}))$.

The numbers $i < n$ denote colours, and we can think of $\forall$ and $\exists$ taking it in turns to paint subsets of T and T' with different colours. Continuing the painting analogy, a *palette* π is a subset of $\{i : i < n\}$, in other words, a choice of colours. When $n = 0$ the only palette is $\emptyset$. For $0 < i < n$, at position $p_i = ((T_0, T'_0), \dots, (T_{i-1}, T'_{i-1}))$ we may interpret a palette π in T and in T', by

$$\pi^{p_i}_T = \{x \in T : \forall j < i (x \in T_j \iff j \in \pi)\},$$
$$\pi^{p_i}_{T'} = \{x \in T' : \forall j < i (x \in T'_j \iff j \in \pi)\}.$$

Intuitively, $\pi^{p_i}_T$ is the set of elements of T coloured with exactly the combination of colours defined by π at position p_i. We define $\pi^{p_0}_T = T$ and $\pi^{p_0}_{T'} = T'$ for all palettes π. Note also that for every position p, the set of vertices of T is the disjoint union of the sets π^p_T, as π ranges over palettes, and similar for T'. A position p is a win for $\forall$ if there is a palette $\pi \subseteq \{i : i < n\}$ where

$$|\pi^p_T| \neq |\pi^p_{T'}| \text{ and either } |\pi^p_T| < 2 \text{ or } |\pi^p_{T'}| < 2. \tag{2}$$

If for any $i \leq n$ position p_i is a win for $\forall$ then $\forall$ wins the play, but if no position in the play is a win for $\forall$ then $\exists$ wins. Since the play is finite, $\forall$ wins if and only if the final position p_n is a win for him.

We call this game the *Seurat game for sets* played over T and T'. This is a variant of the Seurat games for digraphs and general binary structures defined in [3, 4], respectively, but with no restriction to the number of colours and no winning condition for $\forall$ relating to edges or binary relations. The Seurat game for sets over T, T' can be thought of as a special case of a Seurat game for digraphs by thinking of T and T' as being complete graphs.

LEMMA 4.3. *Let $n < \omega$, and let T and T' be sets of size at least 2^{n+1}. Then $\exists$ has a winning strategy in $\mathbf{G}_n(T, T')$.*

PROOF. If $|T|, |T'| \geq 2$ then $\exists$ wins $\mathbf{G}_0(T, T')$ as $\pi^{p_0}_T = T$ and $\pi^{p_0}_{T'} = T'$, so assume $n > 0$. Suppose $\exists$ plays according to the following principle: If the position is p_r, for every palette $\pi \subseteq \{i : i < n\}$, she ensures that

$$(|\pi^{p_r}_T| < 2^{n+1-r} \vee |\pi^{p_r}_{T'}| < 2^{n+1-r}) \to |\pi^{p_r}_T| = |\pi^{p_r}_{T'}|. \tag{$\dagger_r$}$$

If $\exists$ can maintain $(\dagger_r)$ while $r \leq n$ she will survive all rounds. At the end of the final round n, $(\dagger_n)$ ensures that (2) holds, but note that if the game were to continue for another round then she might lose, because $(\dagger_{n+1})$ does not insure her against violating (2).

We now prove by induction that $\exists$ can indeed always play so as to ensure $(\dagger_r)$ holds up to and including $r = n$. The base case is $r = 0$. The position here is p_0,

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

and by definition $\pi_T^{p_0} = T$ and $\pi_{T'}^{p_0} = T'$ for all palettes π, and so $(\dagger_0)$ holds as by assumption $|T|, |T'| \geq 2^{n+1}$.

Suppose now that $0 \leq r < n$, that the position is p_r, and that $(\dagger_r)$ holds. Suppose without loss of generality that $\forall$ picks $T_r \subseteq T$ (the case where he chooses a subset of T' is similar). For $\exists$'s response $T'_r \subseteq T'$, she will pick disjoint subsets $T'_\pi \subseteq \pi_{T'}^{p_r}$ (for each palette π) and then she will define $T'_r = \bigcup_{\pi \subseteq \{i : i < n\}} T'_\pi$. How she chooses each T'_π is explained next. Note that if π is a palette, then $\pi_T^{p_{r+1}}$ will either be $\pi_T^{p_r} \cap T_r$ or $\pi_T^{p_r} \setminus T_r$, depending on whether or not $r \in \pi$, and similar for $\pi_{T'}^{p_{r+1}}$. So maintaining $(\dagger_{r+1})$ comes down to ensuring these sets have the right cardinalities.

- If $|\pi_T^{p_r} \cap T_r| < 2^{(n+1)-(r+1)} = 2^{n-r}$ and $|\pi_T^{p_r} \setminus T_r| < 2^{n-r}$, then $|\pi_T^{p_r}| < 2^{n+1-r}$, so by the inductive assumption $(\dagger_r)$ we have $|\pi_T^{p_r}| = |\pi_{T'}^{p_r}|$. Here she lets T'_π be any subset of $\pi_{T'}^{p_r}$ of size $|\pi_T^{p_r} \cap T_r|$. It follows that $\pi_{T'}^{p_r} \setminus T'_\pi$ has the same size as $\pi_T^{p_r} \setminus T_r$.
- If $|\pi_T^{p_r} \cap T_r| < 2^{n-r}$ but $|\pi_T^{p_r} \setminus T_r| \geq 2^{n-r}$ then she lets T'_π be any subset of $\pi_{T'}^{p_r}$ of the same size as $\pi_T^{p_r} \cap T_r$. It follows that $\pi_{T'}^{p_r} \setminus T'_\pi$ will have size at least 2^{n-r}, because otherwise $|\pi_{T'}^{p_r}| < 2^{n+1-r}$, and so by $(\dagger_r)$ we would have $|\pi_T^{p_r}| = |\pi_{T'}^{p_r}|$, and thus $2^{n-r} \leq |\pi_T^{p_r} \setminus T_r| = |\pi_{T'}^{p_r} \setminus T'_\pi| < 2^{n-r}$, which would be a contradiction.
- Similarly, if $|\pi_T^{p_r} \setminus T_r| < 2^{n-r}$ but $|\pi_T^{p_r} \cap T_r| \geq 2^{n-r}$ she chooses T'_π so that $\pi_{T'}^{p_r} \setminus T'_\pi$ has the same size as $\pi_T^{p_r} \setminus T_r$.
- Finally, if both $\pi_T^{p_r} \cap T_r$ and $\pi_T^{p_r} \setminus T_r$ have size at least 2^{n-r} then $|\pi_T^{p_r}| \geq 2^{n+1-r}$, so inductively $|\pi_{T'}^{p_r}| \geq 2^{n+1-r}$. She lets $T'_\pi \subseteq \pi_{T'}^{p_r}$ be any subset of size 2^{n-r}, and so $|\pi_{T'}^{p_r} \setminus T'_\pi| \geq 2^{n-r}$.

So

$$\pi_T^{p_{r+1}} = \begin{cases} \pi_T^{p_r} \setminus T_r \text{ if } r \notin \pi, \\ \pi^{p_r} \cap T_r \text{ if } r \in \pi, \end{cases}$$

$$\pi_{T'}^{p_{r+1}} = \begin{cases} \pi_{T'}^{p_r} \setminus T'_r = \pi_{T'}^{p_r} \setminus T'_\pi \text{ if } r \notin \pi, \\ \pi^{p_r} \cap T'_r \text{ if } r \in \pi. \end{cases}$$

By the definition by cases given above, the cardinalities of these sets agree when necessary. Thus $(\dagger_{r+1})$ is established for the new position p_{r+1}. ⊣

Let T, T', S be finite sets. We are interested in the game $\Gamma_n(\mathcal{B}_{T,S}, \mathcal{B}_{T',S})$. Recall that $\mathcal{B}_{T,S}$ and $\mathcal{B}_{T',S}$ are complex algebras generated by the 'rainbow' atom structures defined in Section 3. Note that $\mathcal{B}_{T,S}$ and $\mathcal{B}_{T',S}$ are both finite (and thus atomic), and differ only with respect to their sets of green atoms (we identify the non-green atoms between each algebra in the obvious way). Given an element $x \in \mathcal{B}_{T,S}$, we define the subset of elements of T indexing green atoms below x in the standard boolean ordering of $\mathcal{B}_{T,S}$ to be T_x. In addition, given an element $x \in \mathcal{B}_{T,S}$ we define the *green part* of x to be the join of the set of green atoms below x, and the non-green part to be the join of all the other atoms below x. We say x is *green*, if it is a sum of green atoms. We make analogous definitions for $y \in \mathcal{B}_{T',S}$.

PROPOSITION 4.4. *Let T, T', S be non-empty finite sets, and let $n < \omega$. If $\exists$ has a winning strategy in $\mathbf{G}_n(T, T')$, then $\exists$ has a winning strategy in $\Gamma_n(\mathcal{B}_{T,S}, \mathcal{B}_{T',S})$*

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

PROOF. $\exists$'s strategy is to simulate a corresponding play of $\mathbf{G}_n(T, T')$ in which she uses her winning strategy, and to maintain a correspondence between the plays of the games. So, if $\forall$ selects $x \in \mathcal{B}_{T,S}$, then she selects $y \in \mathcal{B}_{T',S}$ whose non-green part is identical to that of x and whose green part is defined by her response to the $\forall$-move T_x in the play of $\mathbf{G}_n(T, T')$, and similar when he picks an element of $\mathcal{B}_{T',S}$.

We assume that $n \geq 1$, as the argument for the $n = 0$ case is essentially the same, but simpler as all that is considered are the empty starting positions in both games. So let $p_n = ((a_0, b_0), \dots, (a_{n-1}, b_{n-1}))$ be the position at the end of the final round in a play of the game $\Gamma_n(\mathcal{B}_{T,S}, \mathcal{B}_{T',S})$ in which $\exists$ uses this strategy, and let $((T_0, T'_0), \dots, (T_{n-1}, T'_{n-1}))$ be the corresponding position at the end of the final round in the corresponding play of $\mathbf{G}_n(T, T')$. Note that p_n is not a winning position for $\forall$ if and only if p_i is not a winning position for him for all $i \leq n$, so it is sufficient for us to prove that $\forall$ does not win at the end of round n.

For each $j < n$ we have $T_j = T_{a_j} \subseteq T$, where T_{a_j} is the set of elements of T indexing the green atoms below a_j, and $T'_j = T'_{b_j} \subseteq T'$ similarly. From the assumption that $\exists$ is playing according to a winning strategy in $\mathbf{G}_n(T, T')$ we have

$$\text{either } |\pi_T^{p_n}| = |\pi_{T'}^{p_n}| \text{ or } |\pi_T^{p_n}|, |\pi_{T'}^{p_n}| \geq 2, \text{ for all palettes } \pi, \tag{3}$$

and we have to prove that p_n is not a losing position in the game $\Gamma_n(\mathcal{B}_{T,S}, \mathcal{B}_{T',S})$. If we define $\bar{a} = (a_0, \dots, a_{n-1})$ and $\bar{b} = (b_0, \dots, b_{n-1})$, it is sufficient to show that the map $\bar{a} \mapsto \bar{b}$ induces a relation algebra isomorphism from $(\mathcal{B}_{T,S})_{\bar{a}}$ to $(\mathcal{B}_{T',S})_{\bar{b}}$ (recall that these are the subalgebras generated by $\bar{a}$ and $\bar{b}$ respectively).

Consider the boolean subalgebra $\mathcal{B}$ of $\mathcal{B}_{T,S}$ generated (using boolean operators) by $\{a_0, \dots, a_{n-1}\}$ and all the non-green atoms. Given a palette π, let $\pi^{\mathcal{B}} = \sum_{i \in \pi_T^{p_n}} \mathrm{g}_i \in \mathcal{B}$. If $\pi_T^{p_n} = \emptyset$ then the sum is empty and $\pi^{\mathcal{B}} = 0$, else the sum is non-empty and $\pi^{\mathcal{B}}$ is an atom of $\mathcal{B}$ (though not usually an atom of $\mathcal{B}_{T,S}$). To see that non-zero $\pi^{\mathcal{B}}$ is an atom note that it has zero intersection with every non-green atom, and its intersection with an element a_j is either itself, when $j \in \pi$, or zero otherwise. A little thought reveals that all green atoms of $\mathcal{B}$ arise from palettes in this way. Note that in the $n = 0$ case, the only palette is $\pi = \emptyset$, and $\pi^{\mathcal{B}} = \sum_{i \in T} \mathrm{g}_i$, so there is exactly one green atom in $\mathcal{B}$.

Similarly, let $\mathcal{B}'$ be the boolean sub-algebra of $\mathcal{B}_{T',S}$ generated by the non-green atoms and $\{b_0, \dots, b_{n-1}\}$, and let $\pi^{\mathcal{B}'} = \sum_{i \in \pi_{T'}^{p_n}} \mathrm{g}_i$. Since $\pi_T^{p_n} = \emptyset \iff \pi_{T'}^{p_n} = \emptyset$, for all palettes (by (3)), the map

$$\phi = \{(\pi^{\mathcal{B}}, \pi^{\mathcal{B}'}) : \pi \subseteq \{0, \dots, n-1\}\} \setminus \{(0, 0)\}$$

is a bijection from the green atoms of $\mathcal{B}$ to those of $\mathcal{B}'$ which extends to a unique boolean isomorphism $\hat{\phi} : \mathcal{B} \to \mathcal{B}'$ fixing non-green atoms.

We will show that $\mathcal{B}$ is a relation algebra and that $\hat{\phi}$ is a relation algebra isomorphism. Note first that $\mathcal{B}$ contains the identity and is closed under conversion, since all green elements are self-converse, furthermore both identity and converse

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

are preserved by $\hat{\phi}$. To show that $\mathcal{B}$ is a relation algebra and to show that $\hat{\phi}$ is a relation algebra isomorphism, we must show, for all atoms $x, y, z \in \mathcal{B}$ that

$$\text{either } (\times)(x;y)\cdot z = 0 \text{ or } (\checkmark)x;y \geq z, \text{ and} \tag{4}$$

$$(x;y)\cdot z = 0 \iff (\phi(x);\phi(y))\cdot \phi(z) = 0. \tag{5}$$

In (4), we indicate the two alternatives by $\times$ and $\checkmark$.

The cases where $1' \in \{x, y, z\}$ are easy, so suppose x, y, z are non-identity atoms of $\mathcal{B}$ (so each has a colour). Recall that the only sets of three colours where some but not all triples of atoms of those colours are forbidden, are red–red–red and green–green–red (rules IV and V). In all other cases either all triples of atoms of those colours are forbidden and we have $\times$ (and both sides of (5) are true) or none is forbidden and we have $\checkmark$ (and both sides are false). If z is red we get $\checkmark$ or $\times$ automatically since z is an atom of $\mathcal{B}_{T,S}$. We also have (5) as, if x and y are red, then $\hat{\phi}$ fixes them, since they are non-green. Alternatively, if x and y are green then $x = \pi_1^{\mathcal{B}}$ and $y = \pi_2^{\mathcal{B}}$ for some palettes π_1 and π_2. In this case $(x;y)\cdot z \neq 0$ if and only if there is $i \neq i' \in T$ with $\mathsf{g}_i \leq x$ and $\mathsf{g}_{i'} \leq y$ (by rule (V)), which occurs if and only if $\pi_1^{\mathcal{B}}$ and $\pi_2^{\mathcal{B}}$ are non-empty and are either distinct or have size at least two, if and only if $\pi_1^{\mathcal{B}'}$ and $\pi_2^{\mathcal{B}'}$ are non-empty and are either distinct or have size at least two (by the assumption that $\exists$ is using a winning strategy in $\mathbf{G}_n(T, T')$), if and only if $(\phi(x);\phi(y))\cdot\phi(z) \neq 0$.

So, without loss of generality suppose that $x = \pi_1^{\mathcal{B}}$ and $z = \pi_2^{\mathcal{B}}$ are green and $y = \mathsf{r}_{j',j}$ is red (for some $j, j' \in S$). We are interested in $(\pi_1^{\mathcal{B}};\mathsf{r}_{j',j})\cdot\pi_2^{\mathcal{B}}$. We want to show that either for every green atom g_i (of $\mathcal{B}_{T,S}$) below $\pi_2^{\mathcal{B}}$ there is a green atom $\mathsf{g}_{i'}$ below $\pi_1^{\mathcal{B}}$ such that $(\mathsf{g}_{i'}, \mathsf{r}_{j',j}, \mathsf{g}_i)$ is not forbidden (for $\checkmark$), or that for every green atom g_i (of $\mathcal{B}_{T,S}$) below $\pi_2^{\mathcal{B}}$ and for every green atom $\mathsf{g}_{i'}$ below $\pi_1^{\mathcal{B}}$, the triple $(\mathsf{g}_{i'}, \mathsf{r}_{j',j}, \mathsf{g}_i)$ is forbidden (for $\times$).

Applying the Peircean equivalences, the triple under consideration here is equivalent to $(\mathsf{g}_i, \mathsf{g}_{i'}, \mathsf{r}_{j,j'})$. According to rule (V), this will be forbidden if and only if either $i = i'$, or $j = j'$. If $j = j'$, then $(\mathsf{g}_i, \mathsf{g}_{i'}, \mathsf{r}_{j,j})$ is always forbidden (regardless of i, i'), so we have $\times$, and both sides of (5) are true. Alternatively, suppose $j \neq j'$ and let $\mathsf{g}_i \leq \pi_2^{\mathcal{B}}$. If $x = z$, i.e., if $\pi_1 = \pi_2$, then either $\pi_1^{\mathcal{B}} = \pi_2^{\mathcal{B}} = \mathsf{g}_i$ for some $i \in T$, in which case we have $\times$ as $(\mathsf{g}_i, \mathsf{g}_i, \mathsf{r}_{j,j'})$ is forbidden, or $\pi_1^{\mathcal{B}} = \pi_2^{\mathcal{B}}$ is above at least two distinct green atoms, in which case we have $\checkmark$ as $(\mathsf{g}_i, \mathsf{g}_{i'}, \mathsf{r}_{j,j'})$ is not forbidden when $i \neq i'$ and $j \neq j'$. As before, both sides of (5) are true, or they are both false in each case. Similarly, if $x \neq z$, i.e., if $\pi_1 \neq \pi_2$, then as $\pi_1^{\mathcal{B}} \neq 0$ and palettes are disjoint, for each green atom g_i below $\pi_2^{\mathcal{B}}$, there must be $\mathsf{g}_{i'} \leq \pi_1^{\mathcal{B}}$ distinct from g_i. As $(\mathsf{g}_i, \mathsf{g}_{i'}, \mathsf{r}_{j,j'})$ is not forbidden this proves $\checkmark$, and both sides of (5) are false.

It follows that $\mathcal{B}$ is closed under all relation algebra operations, and is a sub-relation algebra of $\mathcal{B}_{T,S}$. Since $\{a_0, \ldots, a_{n-1}\} \subseteq \mathcal{B}$ we have the inclusion of relation algebras, $(\mathcal{B}_{T,S})_{\bar{a}} \subseteq \mathcal{B} \subseteq \mathcal{B}_{T,S}$. Similarly, the boolean subalgebra $\mathcal{B}'$ of $\mathcal{B}_{T',S}$ generated by non-green atoms and $\{b_0, \ldots, b_{n-1}\}$ is a sub-relation algebra of $\mathcal{B}_{T',S}$ extending $(\mathcal{B}_{T',S})_{\bar{b}}$. By (5), $\hat{\phi}$ is a relation algebra isomorphism from $\mathcal{B}$ onto $\mathcal{B}'$.

Moreover, for all $i < n$ we have $\hat{\phi}(a_i) = b_i$, as $\exists$'s strategy in $\Gamma_n(\mathcal{B}_{T,S}, \mathcal{B}_{T',S})$ ensures this is true. It follows that the restriction of $\hat{\phi}$ to $(\mathcal{B}_{T,S})_{\bar{a}}$, which, as we have

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

just proved, is an isomorphism onto $(\mathcal{B}_{T',S})_{\bar{b}}$, is generated by $\{(a_i, b_i) : i < n\}$. This proves the result. ⊣

COROLLARY 4.5. *Let $n < \omega$, let T, T', S be finite sets. If $|T|, |T'| \geq 2^{n+1}$, then $\mathcal{B}_{T,S} \equiv_n \mathcal{B}_{T',S}$.*

PROOF. By Lemma 4.3 ∃ has a winning strategy in $\mathbf{G}_n(T, T')$. So, by Proposition 4.4 she has a winning strategy in $\Gamma_n(\mathcal{B}_{T,S}, \mathcal{B}_{T',S})$. It follows by Lemma 4.2 that $\mathcal{B}_{T,S} \equiv_n \mathcal{B}_{T',S}$. ⊣

We can now prove our main result.

THEOREM 4.6. *If Σ is a set of first-order formulas defining RRA, then Σ includes formulas of arbitrary quantifier depth.*

PROOF. Let T, S be sets where $|S| = 2^{n+1}$, $|T| = 2^{n+1} + 1$. Then $\mathcal{B}_{S,S} \in RRA$ but $\mathcal{B}_{T,S} \notin RRA$, by Theorem 3.2, and $\mathcal{B}_{S,S} \equiv_n \mathcal{B}_{T,S}$ by Corollary 4.5. It follows that RRA cannot be axiomatised by any theory consisting of sentences of quantifier depth at most n. ⊣

REMARK 4.7. This suggests that a similar construction could be used to prove that the class of representable cylindric algebras of dimension n cannot be defined by a theory of bounded quantifier depth, however we have not succeeded in demonstrating this. There is a way of constructing a *rainbow cylindric algebra* of dimension $n \geq 4$ from two graphs G, H given in [10, Section 4.3.3]. The atoms of this cylindric algebra are certain labelled hypergraphs on n nodes. The two-dimensional edges of these hypergraphs have green, red and white labels generalising the green, red, yellow and black atoms of the rainbow relation algebra $\mathcal{A}_{G,H}$, but these hypergraphs also have $(n-1)$-ary hyperlabels on some hyperedges in the cylindric version. However, just as we were able to modify the rainbow relation algebra construction, essentially by deleting all atoms w_S for $|S| \leq 2$, we can modify the rainbow cylindric construction by deleting all hyperedges. It follows that $\mathcal{C}^n_{G,H}$ is generated by its *relation algebra reduct*. By considering the graphs K_{m+1} and K_m we can show that $\mathcal{C}^n_{K_{m+1},K_m}$ is not in RCA_n but $\mathcal{C}^n_{K_m,K_m}$ is in RCA_n. The problem is that although the atom structures of these two cylindric algebra agree on all m-variable atom structure formulas, we cannot prove a cylindric version of Proposition 4.4, so we do not know if the two cylindric algebras are equivalent with respect to unrestricted formulas of quantifier depth at most $\log m$. Thus our attempt to extend to various algebras of higher order relations using the known connections between relation algebras and cylindric algebras was not successful.

REFERENCES

[1] R. BERGHAMMER and M. WINTER, *Solving computational tasks on finite topologies by means of relation algebra and the RelView tool*. ***Journal of Logical and Algebraic Methods in Programming***, vol. 88 (2017), pp. 1–25.

[2] L. H. CHIN and A. TARSKI, *Distributive and modular laws in the arithmetic of relation algebras*. ***University of California Publication in Mathematics (New Series)***, vol. 1 (1951), pp. 341–384.

[3] R. EGROT and R. HIRSCH, *A corrected strategy for proving no finite variable axiomatisation exists for RRA*, preprint, 2021, https://arxiv.org/abs/2109.01357.

[4] ———, *Seurat games on Stockmeyer graphs*. ***Journal of Graph Theory***, in press.

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press

[5] G. H. L. Fletcher, M. Gyssens, J. Paredaens, D. Van Gucht, and Y. Wu, *Structural characterizations of the navigational expressiveness of relation algebras on a tree*. ***Journal of Computer and System Sciences***, vol. 82 (2016), no. 2, pp. 229–259.

[6] S. Givant, *Tarski's development of logic and mathematics based on the calculus of relations*, ***Algebraic Logic*** (H. Andréka, J. Monk, and I. Németi, editors), Colloq. Math. Soc. J. Bolyai., vol. 54, North-Holland, Amsterdam, 1991, pp. 189–215.

[7] ———, *The calculus of relations as a foundation for mathematics*. ***Journal of Automated Reasoning***, vol. 37 (2006), no. 4, pp. 277–322.

[8] W. Guttmann, *Verifying minimum spanning tree algorithms with Stone relation algebras*. ***Journal of Logical and Algebraic Methods in Programming***, vol. 101 (2018), pp. 132–150.

[9] R. Hirsch, *Completely representable relation algebras*. ***Bulletin of the interest group in propositional and predicate logics***, vol. 3 (1995), no. 1, pp. 77–92.

[10] R. Hirsch and I. Hodkinson, *Complete representations in algebraic logic*, this Journal, vol. 62 (1997), no. 3, pp. 816–847.

[11] ———, *Relation algebras from cylindric algebras, II*. ***Annals of Pure and Applied Logic***, vol. 112 (2001), pp. 267–297.

[12] ———, ***Relation Algebras by Games***, North-Holland, Elsevier Science, Amsterdam, 2002.

[13] R. Hirsch, I. Hodkinson, and R. Maddux, *Relation algebra reducts of cylindric algebras and an application to proof theory*, this Journal, vol. 67 (2002), no. 1, pp. 197–213.

[14] I. Hodkinson and Y. Venema, *Canonical varieties with no canonical axiomatisation*. ***Transactions of the American Mathematical Society***, vol. 357 (2005), pp. 4579–4605.

[15] N. Immerman, ***Descriptive Complexity***, Graduate Texts in Computer Science, Springer-Verlag, New York, 1999.

[16] B. Jónsson, *The theory of binary relations*. In ***Algebraic Logic*** (H. Andréka, J. Monk, and I. Németi, editors), Colloq. Math. Soc. J. Bolyai., vol. 54, North-Holland, Amsterdam, 1991, pp. 245–292.

[17] B. Jónsson and A. Tarski, *Representation problems for relation algebras*. ***Bulletin of the American Mathematical Society***, vol. 54 (1948), no. 80, p. 1192.

[18] R. Lyndon, *The representation of relational algebras*. ***Annals of Mathematics***, vol. 51 (1950), no. 3, pp. 707–729.

[19] R. Maddux, *Non-finite axiomatizability results for cylindric and relation algebras*, this Journal, vol. 54 (1989), no. 3, pp. 951–974.

[20] ———, *The origin of relation algebras in the development and axiomatization of the calculus of relations*. ***Studia Logica***, vol. 50 (1991), no. 3–4, pp. 421–455.

[21] J. Monk, *On representable relation algebras*. ***The Michigan Mathematical Journal***, vol. 11 (1964), pp. 207–210.

[22] A. Tarski, *On the calculus of relations*, this Journal, vol. 6 (1941), pp. 73–89.

[23] ———, *Contributions to the theory of models, I, II*. ***Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen***, vol. 57 (1954), pp. 572–581 and 582–588 resp.

[24] A. Tarski and S. Givant, ***A Formalization of Set Theory Without Variables***. Colloquium Publications, vol. 41, American Mathematical Society, Providence, 1987.

[25] Y. Venema. Atom structures. In M. Kracht, M. D. Rijke, H. Wansing, and M. Zakharyaschev, editors, ***Advances in Modal Logic '96***, pages 291–305. CSLI Publications, Stanford, 1997.

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
MAHIDOL UNIVERSITY, SALAYA 73170, THAILAND
E-mail: robert.egr@mahidol.ac.th

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY COLLEGE LONDON
LONDON WC1E 6BT, UK
E-mail: r.hirsch@ucl.ac.uk

https://doi.org/10.1017/jsl.2021.88 Published online by Cambridge University Press