Relation Algebra with Binders

Maarten Marx

Applied Logic Lab, ILLC, Universiteit van Amsterdam,
Plantage Muidergracht 24, 1018 TV, Amsterdam, The Netherlands
Email: marx@science.uva.nl

Abstract

The language of relation algebras is expanded with variables denoting individual elements in the domain and with the $\downarrow$ binder from hybrid logic. Every elementary property of binary relations is expressible in the resulting language, something which fails for the relation algebraic language. That the new language is natural for speaking about binary relations is indicated by the fact that both Craig's Interpolation, and Beth's Definability theorems hold for its set of validities. The paper contains a number of worked out examples.

Keywords: binary relations, relation algebras, modal logic, hybrid logic, fork algebras, interpolation, Beth definability.

1 Introduction and Motivation

Tarski wrote

> [...] the calculus of relations has an intrinsic charm and beauty which makes it a source of intellectual delight to all who become acquainted with it. [10: p89]

Still, first-order logic is the language universally understood in our field. We write $\forall x \forall y(R(x,y) \rightarrow R(y,x))$ instead of $R \subseteq R^{-1}$, we write $\forall x \forall y \forall z((R(x,y) \wedge R(y,z)) \rightarrow R(x,z))$ instead of $R \circ R \subseteq R$, and $\forall x \forall y \forall z((R(x,y) \wedge R(x,z)) \rightarrow \exists w(R(y,w) \wedge R(z,w)))$ when $R^{-1} \circ R \subseteq R \circ R^{-1}$ expresses the same fact.

I do not want to go into the reasons for the supremacy of first-order logic over relation algebras. I conjecture such research would be both speculative and necessarily interdisciplinary. It might even be that matters boil down to the same reasons as why we are still typing on QWERTY keyboards. Whatever the reasons, I think Tarski himself was not a great advocate of his beloved system.

There is though an obvious reason for preferring first-order logic: not every elementary statement about individuals and binary relations can be expressed in relation algebraic terms. The first counterexample —"the domain contains at least four elements"— was already discovered in 1915 (see Example 3 below). The question whether a given elementary statement can be equivalently expressed in relation algebraic terms is in general undecidable [7, 11]. On the other hand, the set of relation algebraic terms can be characterized as a fragment of first-order logic: every relation algebraic term is equivalent to a first-order formula with at most two free variables and in which at most 3 variables may occur and vice versa [11].

Tarski knew when he wrote his delightful sentence quoted above that under some mild conditions (the most important is logicality of the operators) no finite expansion of the relation algebraic language could provide full first-order expressivity. A finite expansion with a non-logical operator (called fork) was made in fork algebras (cf. the survey containing an extensive bibliography [5]). The fork operation takes two binary relations and produces a new binary relation. The fork operation ∇ is defined assuming that the universe is closed under taking pairs of elements: for R, S binary relations and x, y elements of the domain

$$xR\nabla Sy \iff \exists u \exists v(y = \langle u, v\rangle \wedge xRu \wedge xSv).$$

In this paper, I will define a much simpler but infinite expansion instead. A number of examples show that the language is natural and easy to use. The expansion has full first-order expressivity. As an immediate result of this the language has interpolation and definability theorems. Both features are lacked by the relation algebraic language. These two properties are powerful indicators of the naturalness of a language, in particular of a good fit between syntax and semantics.

The expansion proposed here is not new. The same expansion has been applied in modal logic under the name of hybrid logic. As hybrid logic is well motivated in this issue and elsewhere [3], a brief explanation suffices. Hybrid logic $\mathcal{H}(@, \downarrow)$ adds three features to the standard uni-modal propositional language:

1. A set VAR of new atomic formulas. Elements of this set are variables ranging over the set of states of a Kripke model. The variables obtain a value by means of an assignment function. A variable x holds at a state s in Kripke model $\mathfrak{M}$ under an assignment g (notation: $\mathfrak{M}, s, g \Vdash x$) iff g assigns the state s to x.

2. For every $x \in VAR$ and for every formula φ, $@_x\varphi$ is a formula too. The formula $@_x\varphi$ is a "holds" predicate: it states that at the state to which x is bound, φ holds. Formally:
$$\mathfrak{M}, s, g \Vdash @_x\varphi \iff \mathfrak{M}, g(x), g \Vdash \varphi.$$

3. For every $x \in VAR$ and formula φ, $\downarrow x.\varphi$ is a formula too. The formula $\downarrow x.\varphi$ binds the variable x to the current state of evaluation:
$$\mathfrak{M}, s, g \Vdash \downarrow x.\varphi \iff \mathfrak{M}, s, g^s_x \Vdash \varphi,$$
in which g^s_x is the assignment which is just like g except that x is sent to s.

   The downarrow binder provides the means to temporarily store the state of evaluation in the value of the variable x. The operator $\downarrow x$ obeys the same binding rules as the quantifiers in first-order logic. The binder $\downarrow x$ is best seen as an explicit substitution operator. This can readily be seen from its standard translation [1] into first-order logic: $ST_x(\downarrow x_i.\varphi) = (ST_x(\varphi))[x_i/x]$. Note that $(ST_x(\varphi))[x_i/x] \equiv \exists x_i(x_i = x \wedge ST_x(\varphi)) \equiv \forall x_i(x_i = x \rightarrow ST_x(\varphi))$.

A nice example of the strength of the formalism is given by the following definition of the until operator. (a) gives the intended meaning of Until and (b) its definition in hybrid logic.

$$\begin{array}{llcll} (a) & \mathfrak{M}, x \Vdash \mathsf{Until}(\varphi, \psi) & \iff & \exists y(xRy \wedge \mathfrak{M}, y \Vdash \varphi & \wedge \quad \forall z(xRzRy \rightarrow \mathfrak{M}, z \Vdash \psi)) \\ (b) & \mathsf{Until}(\varphi, \psi) & := & \downarrow x.(\Diamond \downarrow y.(\varphi & \wedge \quad @_x \Box(\Diamond y \rightarrow \psi))). \end{array}$$

That is, we name the current state x, use $\Diamond$ to move to an accessible φ-state which we name y, and then use $@_x$ to jump back to the state named x. We then insist that ψ holds in all successors of the x named state which precede the state named y.

A second example is the spy-point formula from [1]:

$$\downarrow z.\Box\Box \downarrow x.@_z \Diamond x. \tag{1}$$

If it is true in a model at a state s, then every state in the submodel generated from s is reachable in one R step from s. The state s is "spying" on the model. (1) is an example of a *pure sentence*, a formula without propositional variables and free occurrences of variables. Pure sentences can be used to force properties of the accessibility relation. For instance, the following two statements are equivalent on every Kripke model $\mathfrak{M}$:

1. $\downarrow z.\Box\Box \downarrow x.@_z \Diamond x$ holds at every state in $\mathfrak{M}$,
2. the accessibility relation R in $\mathfrak{M}$ is transitive.

The hybrid logic has a nice characterization as the bounded fragment of first-order logic. Like first-order logic it is undecidable and has interpolation and definability properties [1].

I now turn to the hybridization of relation algebras.

2 The Expansion

Relation algebraic terms are defined from atomic relation symbols and the special relation symbols $\emptyset, \top$ and $1'$ by the following grammar:

$$R \cap S \mid R \cup S \mid -R \mid R^{-1} \mid R \circ S.$$

The inequality relation $-1'$ is abbreviated as $0'$. Models are just first-order models (D, I) consisting of a domain of individuals D and an interpretation I of the atomic relation symbols. As the language contains only binary relations it holds that $I(R) \subseteq D \times D$ for every atomic symbol R. In a model $\mathfrak{M} = (D, I)$, every relation algebraic term denotes a binary relation. $\mathfrak{M} \models a[R]b$ denotes that the pair (a, b) stands in the relation R in model $\mathfrak{M}$. If the model is clear from the context just $a[R]b$ is written. The statement $\mathfrak{M} \models R = S$ expresses that in $\mathfrak{M}$, R and S denote the same relation.

The meaning of the complex terms is defined inductively as follows. For atomic R, $\mathfrak{M} \models a[R]b$ iff $(a, b) \in I(R)$. Relations constructed with the Boolean connectives get the standard meaning with $\top$ denoting the universal relation $D \times D$. Then $\mathfrak{M} \models a[1']b \iff a = b$, $\mathfrak{M} \models a[R^{-1}]b \iff \mathfrak{M} \models b[R]a$ and $\mathfrak{M} \models a[R \circ S]b$ iff there exists a $c \in D$ such that $\mathfrak{M} \models a[R]c$ and $\mathfrak{M} \models c[S]b$.

An important defined operator is the residual $\backslash$. It can be viewed as a dynamic implication, witness the validity of the inequality $R\circ(R\backslash S) \subseteq S$. The $\backslash$ is term-definable using negation, converse and composition:

$$R\backslash S \equiv -(R^{-1}\circ -S).$$

With this definition it holds that

$$a[R\backslash S]b \iff \text{for all } c \in D, \text{ if } c[R]a, \text{ then } c[S]b.$$

The binding machinery from the hybrid logic $H(\downarrow, @)$ is now added to the relation algebraic language. Let VAR be a countably infinite set of variables. Add to the construction rules of the relation algebraic terms the following two:

- all elements in VAR are terms, and
- if $x \in VAR$ and R is a term, then $\downarrow^0_x.R$ is a term as well.

The newly obtained language is denoted by $RL\!\downarrow$, pronounced as "RL downarrow". Assignment functions are needed to give meaning to the new terms. For $\mathfrak{M} = (D, I)$ a model, let $g : VAR \longrightarrow D$ be an assignment. In the definition of the denotation of a term R in $\mathfrak{M}$ relative to g (in symbols: $\mathfrak{M} \models a[R]_g b$), the old clauses do not change apart from adding the assignment everywhere. The new terms are interpreted as follows:

$$\begin{array}{lcl} \mathfrak{M} \models a[x]_g b & \iff & g(x) = a = b, \text{ for } x \in VAR \\ \mathfrak{M} \models a[\downarrow^0_x.R]_g b & \iff & \mathfrak{M} \models a[R]_{g^a_x} b, \end{array}$$

in which g^a_x is the assignment which is just like g except that x is sent to a.

Note that the term x, when interpreted on a model with assignment g, denotes the binary relation $\{(g(x), g(x))\}$. These variables work just as pronouns in natural language and can be freely combined with other relation symbols. A natural language sentence like *"He loves her"* can be translated as $x \circ \mathsf{love} \circ y$. Now $x \circ \mathsf{love} \circ y$ is true on a model under an assignment g precisely when $(g(x), g(y))$ are in the interpretation of the relation love. *He loves someone who loves him* is naturally translated as $x \circ \mathsf{love} \circ \mathsf{love} \circ x$.

An $RL\!\downarrow$ term is called closed if all variables occurring in it are bound by a downarrow.

In writing formulas, the following abbreviations are sometimes useful:

$$\begin{array}{lcl} \downarrow^1_x.R & \equiv & (\downarrow^0_x.R^{-1})^{-1} \\ @^0_x R & \equiv & \top\circ x\circ R \\ @^1_x R & \equiv & R\circ x\circ \top. \end{array}$$

Writing out the definitions, we obtain that

$$\begin{array}{lcl} \mathfrak{M} \models a[\downarrow^1_x.R]_g b & \iff & \mathfrak{M} \models a[R]_{g^b_x} b \\ \mathfrak{M} \models a[@^0_x.R]_g b & \iff & \mathfrak{M} \models g(x)[R]_g b \\ \mathfrak{M} \models a[@^1_x.R]_g b & \iff & \mathfrak{M} \models a[R]_g g(x). \end{array}$$

The new language looks very much like $\mathcal{H}(@, \downarrow)$. Because terms denote binary relations, there are two different versions of both $\downarrow$ and @, one for each coordinate.

Remark 1 A different and maybe more natural way of adding the possibility of naming elements of the domain is to add instead of the $\downarrow^0_x$ operator, the following form of composition:

- if $x \in VAR$ and R and S are terms, then also $R \circ_x S$ is a term.

$R \circ_x S$ gets the following meaning

$$\mathfrak{M} \models a[R \circ_x S]_g b \quad \iff \quad (\exists c \in D) : \mathfrak{M} \models a[R]_{g^c_x} c \text{ and } \mathfrak{M} \models c[S]_{g^c_x} b.$$

Note that $\circ_x$ binds free occurrences of x in $R \circ_x S$ *both* in R and in S. This backwards binding power is solely due to the fact that we write $\circ_x$ as an infix operator. Thus $R \circ_x S$ stores the value of the intermediate point of the composition in the value of x. It does not matter which of $\circ_x$ and $\downarrow^0_x$ is taken as primitive because these two connectives are interdefinable:

$$R \circ_x S = \downarrow^1_x .R \circ \downarrow^0_x .S \text{ and } \downarrow^0_x .R = 1' \circ_x R.$$

I have chosen for the downarrow binders because I found them easier to use.[1]

3 Examples

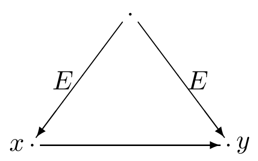
A number of examples are presented. All the properties expressed in $RL\downarrow$ are first-order properties not expressible in relation algebraic terms. All examples except the one about siblings can be found in [11], section 3.6.

Example 1 Express that any two elements have a greatest lower bound (in the order established by E):

$$\forall x \forall y \exists z (zEx \wedge zEy \wedge \forall u((uEx \wedge uEy) \rightarrow uEz)). \tag{2}$$

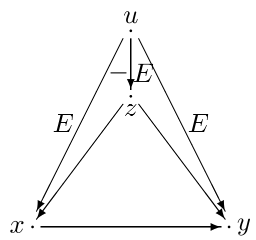
(2) naturally breaks down into two parts:

1. the existential part providing a lower bound. This is easily expressed by $E^{-1} \circ E$. In a picture:



2. and the universal part making it the greatest lower bound. This is more difficult to express. In a picture it says that the following constellation is forbidden if z is the intended greatest lower bound.

[1] A more radical proposal (also suggested by two of the referees) is to have variables referring to *pairs* of elements in the domain, and a corresponding binder. Obviously such an expansion is also term-definably equivalent to the more conservative one developed here. See also Example 4 below in which such a pair-variable can be used.



If we could use the term x as a relation which holds only for the pair (x, x) we can express this as follows: the pair (z, y) should not be in the relation

$$(-E \cap E \circ \mathsf{x} \circ \top)^{-1} \circ E \tag{3}$$

By definition of the residual, (z, y) is not in the relation expressed by (3) iff (z, y) stands in the relation

$$(-E \cap E \circ \mathsf{x} \circ \top) \backslash -E. \tag{4}$$

It now follows that (5) expresses that any two elements have a greatest lower bound

$$\downarrow^{0}_{\mathsf{x}}.(E^{-1} \circ (E \cap (4))) = \top. \tag{5}$$

(Note: in order to distinguish the x used as a name for a point in the drawings and the variable x in the formulas, the variables are written in sans serif script in these examples.)

Example 2 The next example from [11] which is not expressible in ordinary relation algebra is the union axiom:

$$\forall x \forall y \exists u \forall z (zEu \leftrightarrow (zEx \vee zEy)).$$

The intended meaning of zEu is $z \in u$. Graphically the axiom states that given x and y and a choice u for their union, the three constellations in Figure 1 are forbidden. To forbid the situations in (a) and (b), relation algebra suffices. (a) and (b) are expressed by **not** $x[E^{-1} \circ -E]u$ and **not** $y[E^{-1} \circ -E]u$. Using the dynamic implication $\backslash$, we can equivalently express this by $x[E \backslash E]u$ and $y[E \backslash E]u$.

To forbid the situation in (c) we need a name for the point x. If we assume that the only pair in the relation x is (x, x), then we can express it as **not** $u[E^{-1} \circ (-E \cap (-E \circ \mathsf{x} \circ \top))]y$ or equivalently, and much nicer, as

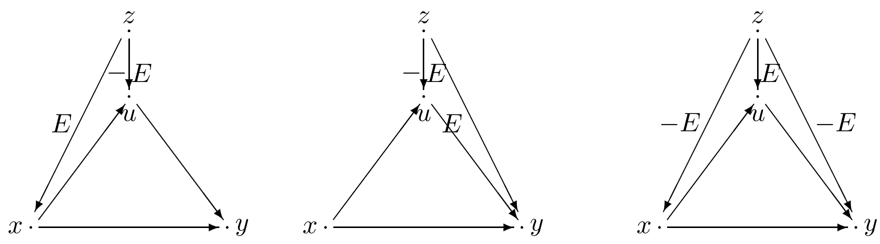
$$u[E \backslash (E \cup E \circ \mathsf{x} \circ \top)]y.$$

(Note that for x denoting a singleton relation $\{(x, x)\}$, $(-E \circ \mathsf{x} \circ \top) = -(E \circ \mathsf{x} \circ \top)$ is valid.) Putting this together, the union axiom is expressed by the equality

$$\downarrow^{0}_{\mathsf{x}}.((E \backslash E) \circ ((E \backslash (E \cup E \circ \mathsf{x} \circ \top)) \cap (E \backslash E)^{-1})) = \top. \tag{6}$$

Example 3 Already in 1915, A. Korselt showed that it is not possible to express by means of a relation algebraic (in)equality that the domain has at least four elements. In the expanded formalism it can be done as in (7).

$$\downarrow^{0}_{\mathsf{x}}.(\top \backslash [(0' \cap \top \circ \mathsf{x} \circ 0') \circ 0']) = \top. \tag{7}$$



(a) (b) (c)

Figure 1: Situations forbidden by the union axiom

(7) is a close translation of the equivalent first-order statement $\forall x\forall y\forall z\exists u(x \neq u \wedge y \neq u \wedge z \neq u)$. Tarski improved on Korselt's example with the following sentence (Tarski was not that impressed by Korselt's sentence because it is a sentence which is, in his words, almost always true):

$$\forall x\forall y\forall z\exists u(xRu \wedge yRu \wedge zRu). \tag{8}$$

This is neatly expressed by

$$\downarrow^{0}_{\mathsf{x}}.(\top\backslash[(R\cap\top\circ\mathsf{x}\circ R)\circ R^{-1}]) = \top. \tag{9}$$

Example 4 Define the following "sibling" relation: two (different) persons are siblings if they share two different parents. In first-order logic, with xPy denoting that x is a parent of y:

$$x\mathsf{Sibling}y \iff x \neq y \wedge \exists p_1\exists p_2(p_1 \neq p_2 \wedge p_1Px \wedge p_2Px \wedge p_1Py \wedge p_2Py).$$

The "dynamic" definition in relation algebraic terms almost shows the two parents in action producing their offspring:[2]

$$\mathsf{Sibling} = 0'\cap \downarrow^{1}_{\mathsf{x}}.[(P\cap[(0'\cap P\circ\mathsf{x}\circ P^{-1})\circ P])^{-1}\circ\top].$$

One might want to add that the set of siblings and the set of parents are disjoint as well. This is obtained by replacing P everywhere by $(P\cap 0')$ and the last $\top$ by $0'$.

4 Full First-Order Expressivity

In this section, it is established that $RL\!\downarrow$ and first-order logic are equally strong in expressive power. Obviously the signature of first-order logic has to be restricted to just binary relation symbols. Call a closed $RL\!\downarrow$ term R and a first-order sentence φ equivalent if for all models $\mathfrak{M}$, φ is true in $\mathfrak{M}$ if and only if $\mathfrak{M} \models R = \top$. We

[2] One of the referees suggested the following translation in which x is a variable which is bound to a *pair* of elements, and $\downarrow_{\mathsf{x}}$ the corresponding binder.

$$\mathsf{Sibling} = 0'\cap\downarrow_{\mathsf{x}}.[P^{-1}\circ(P\cap(0'\circ P\circ\mathsf{x}\circ P^{-1}))].$$

Note that this is precisely the kind of variable ("they") used in the definition in English.

say that the two formalisms are equally strong in expressive power if for any relation algebraic term R there exists an equivalent first-order sentence φ and the other way round. Actually an even stronger form of equivalence holds. There exist recursive translations between the languages providing the equivalent formulation in the other language. Moreover, these translations preserve the atomic relation symbols.

The easy direction can be shown by a translation from $RL{\downarrow}$ terms to first-order formulas. This translation just copies the meaning definition of the $RL{\downarrow}$ terms. Let x, y, z be first-order variables different from the set of $RL{\downarrow}$ variables VAR. Define recursively the translation $ST_{(x,y)}(\cdot)$ from $RL{\downarrow}$ terms to first-order formulas:

$$\begin{array}{lll} ST_{(x,y)}(R) & := & xRy \qquad\qquad\qquad\qquad \text{for all atomic } R \\ ST_{(x,y)}(1') & := & x{=}y \\ ST_{(x,y)}(\mathsf{v}) & := & v{=}x \wedge x{=}y \qquad\qquad\quad \text{for } \mathsf{v} \text{ a variable} \\ \multicolumn{3}{l}{ST_{(x,y)}(\cdot) \text{ commutes with the booleans}} \\ ST_{(x,y)}(R^{-1}) & := & ST_{(y,x)}(R) \\ ST_{(x,y)}(R{\circ}S) & := & \exists z(ST_{(x,z)}(R) \wedge ST_{(z,y)}(S)) \\ ST_{(x,y)}({\downarrow}^0_v.R) & := & \exists v(v = x \wedge ST_{(x,y)}(R)). \end{array}$$

(Note: in order for this translation to work, the variable z used for translating compositions of relations is not occurring free. Obviously this can be taken care of, e.g., by taking a fresh variable every time.) The proof of the following proposition is a straightforward induction on terms.

Proposition 2 For every model $\mathfrak{M}$, for all elements a, b in its domain, and for every closed $RL{\downarrow}$ term R it holds that

$$\begin{array}{rcl} \mathfrak{M} \models a[R]b & \iff & \mathfrak{M} \models ST_{(x,y)}(R)\ [x \mapsto a, y \mapsto b] \\ \mathfrak{M} \models R = \top & \iff & \mathfrak{M} \models \forall x \forall y ST_{(x,y)}(R). \end{array}$$

For the translation in the other direction, the first-order assignments to variables are separated into two parts: a part which is the denotation of a binary relation in (at most) free variables v_0 and v_1, and another part which handles the assignment of values to all other variables. This is because every relation algebraic term denotes a binary relation, and hence can be thought of as a formula in at most two free variables. This inspires a translation $(\cdot)^t$ from first-order formulas φ to $RL{\downarrow}$ terms which satisfies (10) below. For g an assignment to the variables $v_0, v_1, v_2, \ldots$, let $g_{\restriction 3}$ denote the restriction of g to the variables different from v_0 and v_1.

$$\mathfrak{M} \models \varphi\ [g] \iff \mathfrak{M} \models g(v_0)[\varphi^t]_{g_{\restriction 3}} g(v_1). \tag{10}$$

The translation $(\cdot)^t$ handles the variables v_0 and v_1 in a special way. For a smooth presentation, assume without loss of generality that v_0 and v_1 do not occur as arguments of relation symbols in the first-order formula (using two fresh variables this can always be achieved: e.g., $v_0 R v_3 \equiv \exists v_2(v_0 = v_2 \wedge v_2 R v_3)$). The translation is defined as follows: For atomic formulas,

$$(v_i R v_j)^t := @^0_{v_i} @^1_{v_j} R.$$

For equalities, there is a case distinction:

$$(v_0{=}v_1)^t := (v_1{=}v_0)^t := 1' \text{ and } (v_0{=}v_0)^t := (v_1{=}v_1)^t := \top.$$

For $i \notin \{0, 1\}$,

$$\begin{array}{lclcl} (v_0 = v_i)^t & := & (v_i = v_0)^t & := & v_i \circ \top \\ (v_1 = v_i)^t & := & (v_i = v_1)^t & := & \top \circ v_i. \end{array}$$

For $i, j \notin \{0, 1\}$,

$$(v_i{=}v_j)^t := (@^0_{v_i} v_j) \circ \top.$$

$(\cdot)^t$ commutes of course with the booleans. For the quantifiers, there is again a separation in cases:

$$(\exists v_0 \varphi)^t := \top{\circ}\varphi^t \text{ and } (\exists v_1 \varphi)^t := \varphi^t{\circ}\top,$$

and for $i \notin \{0, 1\}$,

$$(\exists v_i \varphi)^t := \downarrow^0_x.(\top{\circ} \downarrow^0_{v_i}.@^0_x \varphi^t),$$

in which x is a variable which does not occur in φ.

The last clause of the translation describes dynamically what quantification is: store the current point of evaluation in an unoccupied memory place x, then change the point of evaluation at random, then store the value of that new point in v_i, and finally evaluate the formula with this new value for v_i at the original point stored in x. A straightforward induction proves

Proposition 3 For every model $\mathfrak{M}$, for every assignment g, and for every first-order formula φ, it holds that

$$\mathfrak{M} \models \varphi\ [g] \quad \Longleftrightarrow \quad \mathfrak{M} \models g(v_0)[\varphi^t]_{g\restriction 3}\, g(v_1).$$

For sentences φ, $\mathfrak{M} \models \varphi$ iff $\mathfrak{M} \models \varphi^t = \top$.

5 Interpolation and Definability

An immediate consequence of the full first-order expressivity of $RL{\downarrow}$ by language preserving translations is that $RL{\downarrow}$ has an interpolation theorem, that is,

Theorem 4 For all closed $RL{\downarrow}$ terms R, S, the following are equivalent:

1. $R{=}\top \Rightarrow S{=}\top$ is valid.
2. There exists an $RL{\downarrow}$ term I constructed only from atomic symbols occurring both in R and in S for which $R{=}\top \Rightarrow I{=}\top$ and $I{=}\top \Rightarrow S{=}\top$ are both valid.

The interpolation theorem stated in this way fails for both RRA and RA [4]. One can similarly formulate a Beth definability theorem for RRA and RA (restricted to implicitly defined binary and unary relations only). For both these classes this fails as well [9]. Obviously, the thus defined Beth definability theorem holds for $RL{\downarrow}$.

For completeness, a quick simple counterexample to interpolation for RRA is provided together with the required interpolant in $RL{\downarrow}$. Let $p?$ abbreviate $(P \cap 1')$ and similarly for $q?$. Let A be the statement (7) in which the role of the variable x is now played by $p?$:

$$\top \backslash [(0' \cap \top{\circ}p?{\circ}0'){\circ}0'] = \top.$$

If A is true in a model, the model has at least four elements. Let C_1 be the term $\top\circ[q?\cap -(0'\circ q?\circ 0')]\circ\top$. $C_1 = \top$ expresses that there exists precisely one element which stands with itself in the q relation. Let C_2 be $\top\circ[0' \cap (-q)?\circ 0'\circ(-q)?\circ 0'\circ(-q)?]\circ\top$. The equality $C_2 = \top$ expresses that there exist at least three different elements which do not stand with themselves in the q relation.

It is thus clear that $A{=}\top \Rightarrow (C_1 \cap -C_2){=}\top$ is valid. Any interpolant should be written in the empty language and should express that there exist at least four elements. But in the empty language we can not express this, as observed above. In $RL{\downarrow}$, the interpolant is expressible by the equation (7).

Hajnal Andréka observed that such counterexamples can be constructed for any finite number $n > 3$. Thus also for obtaining the interpolation theorem an infinite similarity type seems to be needed. This even holds for much weaker classes of relational type algebras than RA cf., [8].

6 Conclusions and Further Research

Though equally expressive, $RL{\downarrow}$ and the full first-order language differ enormously. Expressing properties which are familiar in first-order logic (like the union axiom) in $RL{\downarrow}$ in an appealing way can be quite hard in the beginning. The sibling example shows that the same relation can have wildly different definitions. The first-order definition somehow forces the reader to consider the submodel consisting of the two siblings and two parents at once. The $RL{\downarrow}$ definition forces the reader to build this submodel during evaluation of the term. Such different modes of evaluation suggest possibly also different algorithms for doing model-checking and theorem proving.

Axiomatics were not discussed. I conjecture that complete axiom systems can rather easily be defined using irreflexivity style rules or something like the Paste rule from [2].

A related research topic is the relation with finite variable fragments of first-order logic and with the proof theoretic approximations to RRA studied by Hirsch, Hodkinson and Maddux [6]. Many questions can be asked here. We mention a few. Is it possible to equationally define a class of $RL{\downarrow}$ algebras whose relation algebra reducts generated the variety RA_n (the class of relation algebras with an n-dimensional relational basis)? How many variables are needed? Is $RL{\downarrow}$ restricted to n variables equally expressive as first-order logic with n variables?

Acknowledgments

The author is supported by NWO grant 612-062-001. Part of this research was carried out during a visit to Patrick Blackburn in Saarbrücken which has been made possible by a grant for international collaboration with the DFKI and the University of Saarbrücken, awarded by the Dutch Research Council NWO to the Dutch research School in Logic. Thanks are due to Jorge Petrucio Viana and Renata Pereira de Freitas and to the anonymous referees for valuable comments.

References

[1] C. Areces, P. Blackburn, and M. Marx. Hybrid logics. Characterization, interpolation and complexity. Technical Report PP-1999-07, Institute for Logic, Language and Computation, University of Amsterdam, 1999. To appear in Journal of Symbolic Logic.

[2] P. Blackburn. Internalizing labelled deduction. *Journal of Logic and Computation*, 10:136–168, 2000.

[3] P. Blackburn. Representation, reasoning, and relational structures: a hybrid logic manifesto. *Logic Journal of the IGPL*, 8(3):339–365, 2000.

[4] S. Comer. Classes without the amalgamation property. *Pacific Journal of Mathematics*, 28:309–318, 1969.

[5] M. Frías, G. Baum, and A. Haeberer. Fork algebras in algebra, logic and computer science. *Fundamenta Informaticae*, 32:1–25, 1997.

[6] R. Hirsch and I. Hodkinson. Relation algebras with n-dimensional relational bases. *Annals of Pure and Applied Logic*, 101:227–274, 2000.

[7] A. Kwatinetz. *Problems of expresibility in finite languages*. PhD thesis, University of California, Berkeley, 1981.

[8] M. Marx. Amalgamation in relation algebras. *Journal of Symbolic Logic*, 63(2):479–484, 1998.

[9] I. Sain. Beth's and Craig's properties via epimorphisms and amalgamation in algebraic logic. In C. H. Bergman, R. D. Maddux, and D. L. Pigozzi, editors, *Algebraic logic and universal algebra in computer science*, volume 425 of *Lecture Notes in Computer Science*, pages 209–226. Springer-Verlag, Berlin, 1990.

[10] A. Tarski. On the calculus of relations. *Journal of Symbolic Logic*, 6:73–89, 1941.

[11] A. Tarski and S. Givant. *A Formalization of Set Theory without Variables*, volume 41. AMS Colloquium publications, Providence, Rhode Island, 1987.