Rough Algebras and Automated Deduction

Anita Wasilewska[1] *and Laurent Vigneron*[2]

[1] Department of Computer Science, State University of New York, Stony Brook, NY 11794-4400, USA. E-mail: anita@cs.sunysb.edu

[2] Loria – Université Nancy 2, B.P. 239, 54506 Vandœuvre-lès-Nancy Cedex, France. E-mail: vigneron@loria.fr

Abstract: The notion of *rough equality* was introduced by Pawlak in [17]. It was extensively examined in [14], [4], [16], and the most recently in [2], [27], and [28]. The rough R5 and R4 algebras investigated here are particular cases of topological rough algebras introduced in [26]. We examine and discuss here some of their most interesting properties, their relationship with each other, and with the topological Boolean S4 and S5 algebras, which are algebraic models for modal logics S4 and S5, respectively. The presented properties were chosen out of over seven hundred theorems which were discovered and proved automatically by the theorem prover daTac (*Déduction Automatique dans des Théories Associatives et Commutatives*). This prover was developed at Loria, Nancy (France), by the second author.

1 Introduction

It is difficult to establish who was the first to use the algebraic methods. The investigations in logic of Boole himself led to the notion which we now call Boolean algebra, but one of the turning points in the algebraic study of logic was the introduction by Lindenbaum and in a slight different form by Tarski (in [24]) of the method of treating formulas, or equivalence classes of formulas as elements of an abstract algebra, called now the *Lindenbaum-Tarski algebra*.

In our work we use the algebraic logic techniques to link rough set theory with logic, abstract algebras and topology. In particular, we have shown in [26] that the notion of *rough equality* of sets leads, via logic and a Lindenbaum-Tarski like construction of an algebra of formulas, to a definition of new classes of algebras, called here *topological quasi-Boolean algebras* and *topological rough algebras*. These algebras are a non-classical (quasi-Boolean instead of Boolean) version of topological Boolean algebras. The topological Boolean algebras were introduced in [10], [11] under the name of closure algebras. They were first (algebraic) models for modal logics, as opposed to Kripke models invented some 20 years later [7].

This paper is a continuation of investigations of [2], [27], [28], and [26]. The organization of the paper is as follows.

In Section 2 we introduce some basic definitions and facts in order to make the paper self contained. We also give a short overview of the work by Banerjee and Chakraborty [2] and Wasilewska [26].
In Section 3 we introduce and investigate two of the topological Boolean algebras, named S4 and S5 because they are models for modal logics S4 and S5, respectively. The topological rough algebras considered here are called, accordingly, R4 and R5 algebras, where "R" stands for their rough equality origins.
In Section 4 we examine the properties and relationship between the R4 and R5 algebras and in Section 4 we discuss the relationship between the rough R4 and R5 algebras and their Boolean S4 and S5 counterparts.

All presented properties were discovered and proved automatically by a theorem prover daTac (*Déduction Automatique dans des Théories Associatives et Commutatives*). We give a short presentation of the deduction techniques implemented in daTac in Section 5.

2 Topological Boolean and Topological Rough Algebras

To make our paper self contained we first review in this section some basic definitions and facts.

Approximation space. Let U be a non-empty set called a universe, and let R be an equivalence relation on U. The triple $(U, \emptyset, R)$ is called an approximation space.

Lower, upper approximations. Let $(U, \emptyset, R)$ and $A \subset U$. We denote by $[u]$ an equivalence class of R. The sets

$$IA = \bigcup\{[u] \in A/R : \ [u] \subset A\},$$

$$CA = \bigcup\{[u] \in A/R : \ [u] \cap A \neq \emptyset\}$$

are called *lower* and *upper* approximations of A, respectively. We use here a topological notation for the lower and upper approximations because of their topological interpretation and future considerations.

Rough equality. Given an approximation space $(U, \emptyset, R)$ and any $A, B \subset U$. We say that the sets A and B are *roughly equal* and denote it by $A \sim_R B$ if and only if $IA = IB$ and $CA = CB$.

Boolean algebra. An abstract algebra $(A, 1, \cap, \cup, \neg)$ with unit element 1 is said to be a Boolean algebra if it is a distributive lattice and every element $a \in A$ has a complement $\neg a \in A$.

Orlowska has shown in [15] that propositional aspects of rough set theory are adequately captured by the modal system S5. In this case a Kripke model gives the approximation space $(A, \emptyset, R)$ in which the well formed formulas are interpreted as *rough sets.*

Following Orlowska result, Banerjee and Chakraborty introduced in [2] a new binary connective $\sim$ in S5, the intended interpretation of which is the notion of the *rough equality.* I.e, they added to the standard set $\{\cup, \cap, \rightarrow, \Leftrightarrow, \neg, \Box, \Diamond\}$ of propositional modal connectives a new binary connective $\sim$ defined in terms of standard connectives as follows: for any formulas A, B (of the modal S5 language), we write $A \sim B$ for the formula $((\Box A \Leftrightarrow \Box B) \cap (\Diamond A \Leftrightarrow \Diamond B))$. In the next step they have used this new connective to define a construction similar to the construction of Lindenbaum-Tarski algebra on the set of all formulas of $S5$ with additional connective $\sim$. Before describing their construction leading to the definition of the rough algebra, we include below a description of a standard construction of a Lindenbaum-Tarski algebra for a given logic.

Lindenbaum-Tarski construction. Given a propositional logic with a set $\mathcal{F}$ of formulas. We define first two binary relations $\leq$ and $\approx$ in the algebra $\mathcal{F}$ of formulas of the given logic as follows. For any $A, B \in F$,

$A \leq B$ if and only if $\vdash (A \Rightarrow B)$, and
$A \approx B$ if and only if $\vdash (A \Rightarrow B)$ and $\vdash (B \Rightarrow A)$.

Then we use the set of axioms and rules of inference of the given logic to prove all facts listed below.

The relation $\leq$ is a quasi-ordering in F.
The relation $\approx$ is an equivalence relation in F. We denote the equivalence class containing a formula A by $[A]$.
The quasi-ordering $\leq$ on F induces an ordering relation on $F/\approx$ defined as follows: $[A] \leq [B]$ if and only if $A \leq B$.
The equivalence relation $\approx$ on F is a congruence with respect to all logical connectives.
The resulting algebra with universe $F/\approx$ is called a Lindenbaum-Tarski algebra.

Example 1. The Lindenbaum-Tarski algebra for *classical propositional logic* with the set of connectives $\{\cup, \cap, \Rightarrow, \neg\}$ is the following.

$$\mathcal{LT} = (F/\approx, \cup, \cap, \Rightarrow, \neg),$$

where the operations $\cup$, $\cap$, $\Rightarrow$ and $\neg$ are determined by the congruence relation $\approx$, i.e. $[A] \cup [B] = [(A \cup B)]$, $[A] \cap [B] = [(A \cap B)]$, $[A] \Rightarrow [B] = [(A \Rightarrow B)]$, $\neg[A] = [\neg A]$.

We prove, in this case (see [19]) that the Lindenbaum-Tarski algebra is a Boolean algebra with a unit element V. Moreover, for any formula A, $\vdash A$ if and only if $[A] = V$.

Example 2. The Lindenbaum-Tarski algebra for modal logic $S4$ or $S5$ with the set of connectives $\{\cup, \cap, \Rightarrow, \neg, \Box, \Diamond\}$ is the following.

$$\mathcal{LT} = (F/\approx, \cup, \cap, \Rightarrow, \neg, I, C),$$

where the operations $\cup$, $\cap$, $\Rightarrow$, $\neg$, I and C are determined by the congruence relation $\approx$, i.e. $[A] \cup [B] = [(A \cup B)]$, $[A] \cap [B] = [(A \cap B)]$, $[A] \Rightarrow [B] = [(A \Rightarrow B)]$, $\neg[A] = [\neg A]$, $IA = [\Box A]$, and $CA = [\Diamond A]$.

In the case of modal logic $S4$ the Lindenbaum-Tarski algebra (see [9], [10], [19]) is *a topological Boolean algebra* and in the case of $S5$ it is topological Boolean algebra such that every open element is closed and every closed element is open. Moreover, in both cases, for any formula A, $\vdash A$ if and only if $[A] = V$.

Banerjee, Chakraborty construction. We define a new binary relation $\approx$ on the set F of formulas of the modal S5 logic as follows. For any $A, B \in F$,

$A \approx B$ if and only if $A \sim B$, i.e.
$A \approx B$ if and only if $\vdash ((\Box A \Leftrightarrow \Box B) \cap (\Diamond A \Leftrightarrow \Diamond B))$.
We prove that the above relation $\approx$, corresponding to the notion of rough equality is *an equivalence relation* on the set F of formulas of $S5$.
We define a binary relation $\leq$ on $F/{\approx}$ as follows.
$[A] \leq [B]$ if and only if $\vdash ((\Box A \Rightarrow \Box B) \cap (\Diamond A \Rightarrow \Diamond B))$.
We prove that $\leq$ is an order relation on $F/{\approx}$ with the greatest element $1 = [A]$, for any formula A, such that $\vdash A$, and with the least element $0 = [B]$, such that $\vdash \neg B$.
We prove that $\approx$ *is a congruence relation* with respect to the logical connectives $\neg$, $\Box$, $\Diamond$, but *is not a congruence relation* with respect to $\Rightarrow$, $\cap$ and $\cup$.
We introduce two new operations $\sqcup$ and $\sqcap$ in $F/{\approx}$ as follows.
$[A] \sqcup [B] = [(A \cup B) \cap (A \cup \Box A \cup \Box B \cup \neg\Box(A \cup B))]$,
$[A] \sqcap [B] = [(A \cap B) \cup (A \cap \Diamond A \cap \Diamond B \cap \neg\Diamond(A \cap B))]$.
We call the resulting structure a *rough algebra (of formulas of logic S5)* or S5 rough algebras, for short.

The formal definition of the S5 rough algebra is hence the following.

S5 Rough algebra. An abstract algebra

$$\mathcal{R} = (F/{\approx}, \sqcup, \sqcap, \neg, I, C, 0, 1),$$

such that the operations $\sqcup$, $\sqcap$ are defined above and the operations $\neg$, I, C are induced, as in the Lindenbaum-Tarski algebra, by the relation $\approx$, i.e. $\neg[A] = [\neg A]$, $IA = [\Box A]$, and $CA = [\Diamond A]$, is called the S5 rough algebra.

In [1], many important properties of the S5 rough algebra were were proved. They were also reported in [2]. We cite here only those which are relevant to our future investigations.

P1 $(F/{\approx}, \leq, \sqcup, \sqcap, 0, 1)$ is a distributive lattice with 0 and 1.
P2 For any $[A], [B] \in F/{\approx}$, $\neg([A] \sqcup [B]) = (\neg[A] \sqcap \neg[B])$,
P3 For any $[A] \in F/{\approx}$, $\neg\neg[A] = [A]$.
P4 The rough algebra is not a Boolean algebra, i.e. there is a formula A of a modal logic $S5$, such that $\neg[A] \sqcap [A] \neq 0$ and $\neg[A] \sqcup [A] \neq 1$.
P5 For any $[A], [B] \in F/{\approx}$, $I([A] \sqcap [B]) = (I[A] \sqcap I[B])$, $I[A] \leq [A]$, $II[A] = I[A]$, $I1 = 1$, and $CI[A] = I[A]$, where $C[A] = \neg I \neg[A]$.

The above, and other properties of the rough algebra lead to some natural questions and observations.

By the property **P4**, the rough algebra's complement operation ($\neg$) is not a Boolean complement. Let's call it *a rough complement*. We can see that the rough complement is pretty close to the Boolean complement because the other de Morgan law $\neg([A] \sqcap [B]) = (\neg[A] \sqcap \neg[B])$ holds in the rough algebra, as well as the very Boolean laws $\neg 1 = 0$ and $\neg 1 = 0$. So what kind of a complement is the rough complement? The rough algebra *is not*, by **P4**, a Boolean algebra, so which kind of algebra is it? Has such an algebra been discovered and investigated before?

Observation. A complement operation with similar properties to the *rough complement* was introduced in 1935 by Moisil [12] and lead to a definition of a notion of *de Morgan Lattices*. De Morgan lattices are distributive lattices satisfying the conditions **P2** and **P3**. In 1957 Białynicki-Birula and Rasiowa have used the de Morgan lattices to introduce a notion of *a quasi-Boolean algebra*. They defined (in [3]) the quasi-Boolean algebras as de Morgan lattices with *unit* element 1. The above led, in [26] to the following definition and observation.

Definition 1 Topological quasi-Boolean algebra. An algebra $(A, \cap, \cup, \sim, 1, I)$ is called *a topological quasi-Boolean algebra* if $(A, \cup, \cap, \sim, 1)$ is a quasi-Boolean algebra and for any $a, b \in A$, $I(a \cap b) = Ia \cap Ib$, $Ia \cap a = Ia$, $IIa = Ia$, and $I1 = 1$.

The element Ia is called a *quasi-interior of a*. The element $\sim I \sim a$ is called *quasi-closure of a*. It allows us to define in A an unary operation C such that $Ca = \sim I \sim a$. We can hence represent the topological quasi-Boolean algebra as an algebra $(A, \cap, \cup, \sim, I, C, 0, 1)$ *similar* to the rough algebra $(F/\approx, \sqcup, \sqcap, \neg, I, C, 0, 1)$. From **P4** we immediately get the following.

Fact 2. *A rough algebra* $\mathcal{R} = (F/\approx, \sqcup, \sqcap, \neg, I, C, 0, 1)$ *is a* topological quasi-Boolean algebra.

Moreover, the property **P5** of the rough algebra tells us also that the operations I and C fulfill an additional property: for any $[A] \in F/\approx$, $CI[A] = I[A]$. This justifies the following definition.

Definition 3 Topological rough algebra. A topological quasi-Boolean algebra $(A, \cap, \cup, \sim, I, C, 0, 1)$ such that for any $a \in A$, $CIa = Ia$, is called a *topological rough algebra*.

3 R5 and R4 Algebras

The R5 and R4 algebras are particular cases of the topological rough algebras [26]. They are not purely mathematical invention. The S5 rough algebra developed and examined in [2] is an example the R5 algebra. The R4 algebra is a quasi-Boolean correspondent of the topological Boolean algebra.

The R5 algebra is a quasi-Boolean version of the topological Boolean algebras such that each open element is closed and each closed element is open.
We adopt here the following formal definition of R4 and R5 algebras.

Definition 4 R4 and R5 algebras. An abstract algebra $(A, 1, \cup, \cap, \sim, I, C)$ is called *a R4 algebra* if it is *a distributive lattice with unit element* 1 and additionally for all $a, b \in A$ the following equations are satisfied:

q1	$\sim\sim a \;=\; a,$
q2	$\sim(a \cup b) \;=\; \sim a \cap \sim b,$
t1	$I(a \cap b) \;=\; Ia \cap Ib,$
t3	$Ia \cap a \;=\; Ia,$
t4	$IIa \;=\; a,$
t5	$I1 \;=\; 1,$
t6	$Ca \;=\; \sim I \sim a.$

The algebra obtained from the R4 algebra by adding the following axiom:

CI $\quad CIa \;=\; Ia$

is called a R5 algebra.

Axioms **q1**, **q2** say that R4 is a quasi-Boolean algebra, axioms **t1** – **t5** are the axioms of a topological space, **t6** defines the rough closure operation, and axiom **CI** says that every (roughly) open element is closed.

A natural set theoretical interpretation of the properties of the topological Boolean algebras is established by the representation theorem. For example, $a \cap Ia = Ia$ means that any set A contains its interior **I**A. The representation theorem provides an intuitive motivation for new properties and is an useful source of counter-examples.

The case of R4 and R5 algebras is more complicated and much less intuitive. While the operations $\cup$ and $\cap$ are represented as set theoretical union and intersection, the operation $\sim$ cannot be represented as a set theoretical complementation. The set theoretical interpretation of the rough complement depends on the mapping $g : A \longrightarrow A$ such that for all $a \in A$, $g(g(a)) = a$, called *involution*. The representation theorem for R4 or R5 algebras states that their properties have to hold in R4, R5 algebras (fields) of sets. For example, the set theoretical meaning of a R4 algebra property $a \cap \sim 1 = \sim 1$ is the following. Given any non empty set X, given any involution g on X, for any $A \subset X$, $A \cap (X - g(X)) = (X - g(X))$. This property is intuitively obvious, because any involution has to map the set X onto itself.

The set theoretical interpretation of the definition of the closure operation in R4 is the following: $CA = X - g(I(X - g(A)))$, for any involution g. One can see that it becomes less intuitive than the "normal" Boolean topological definition of closure as complement of the interior of the complement of the set. The situation becomes even more complex when we think about possible (or impossible) properties. For example, one of the simplest properties of R4 algebras proven by the prover (see Section 3.2) is $a \cap Ib \subseteq C(a \cap Ib) \cap b$. Its set theoretical R4 interpretation is that for any $A, B \subset X$ and for any involution g on X, $A \cap IB \subseteq (X - g(I(X - g(A \cap IB)))) \cap B$.

The above examples show that it is much more difficult to build an intuitive understanding of properties of R4 and R5 algebras, than it is in classical case of topological Boolean algebras. It is not only difficult to prove new properties, it is also difficult to think how they should look like. We have hence used the theorem prover daTac as a tool to generate the R4, R5 algebras' properties (and their proofs). Moreover, we have also used it as a tool for a study of the relationship between both algebras. We strongly believe that such a study would be impossible without the use of the prover.

3.1 Automated Deduction of Properties

The properties of the R4 and R5 algebras presented here are chosen from over seven hundred which were discovered and proved automatically by the theorem prover daTac (*Déduction Automatique dans des Théories Associatives et Commutatives*) developed at Loria, Nancy (France). This software implements a new technique [21] (see Section 5 for a description of this technique) of automated deduction in first-order logic, in presence of associative and commutative operators. This technique combines an ordering strategy [5], a system of deduction rules based on resolution [20] and paramodulation [22] rules, and techniques for the deletion of redundant clauses.

daTac proposes either to prove properties by refutation, or by straightforward deduction from clauses. The refutation technique is a proof existence technique, i.e. we add the negation of a formula we want to prove to the set of initial formulas and we search for a contradiction. In this case, if the proof exists the prover would say: "yes, it is a theorem", i.e. the prover acts as a proof existence checker. Of course, the whole system is, as a classical predicate logic, semi-decidable. In the straightforward deduction, the prover acts as a deductive system, i.e. the end product is a set of properties with their formal proof. We have used here mainly the second technique.

Given the R4 algebra $(A, 1, \cup, \cap, \sim, I, C)$. As the first step we used the prover on a non-topological subset of its axioms, i.e. we used as its input only axioms **q1**, **q2** plus axioms for a distributive lattice. As the fact that the considered operators $\cup$ and $\cap$ are associative and commutative is embeded in the structure of the prover, we did not need to specify that portion of the distributive lattice axioms. The prover has immediately deduced the following properties:

$$\begin{array}{ll} a \cup a = a, & a \cap a = a, \\ a \cap \sim 1 = \sim 1, & a \cup \sim 1 = a, \\ \sim(a \cap b) = \sim a \cup \sim b. & \end{array}$$

We have added them to the set of the initial axioms of R4. We have also added the definition of the C operator, i.e. the following equation.

$$Ca = \sim I \sim a.$$

In the paper we use the above, extended version of the definition of the R4 algebra.

3.2 Properties Common to R4 and R5

The axioms of R4 are strictly included in the set of axioms for R5. Hence all properties we can prove in R4, we can prove in R5 and there are also *pure* R5 properties, i.e. the R5 properties which cannot be proved in R4. Of course in general setting the *pure* R5 properties are the set theoretical difference between all R5 properties and those which are *common* to R4 and R5. In general case there is a countably infinite number of all properties of the R4 and R5 algebras, so we never can find all pure R5 properties. In our case all sets of generated properties are finite and we present here a practical way of finding the *common* and *pure* properties. It is not straightforward because of the nature of the prover and we discuss the results in this section for the common properties and in Section 3.3 for the pure ones.

We have used the prover separately for R4 and R5 algebras. All executions have been arbitrary stopped after 5000 deduced clauses. When we have stopped the experiments, the prover had kept only 407 properties for R4 and 294 properties for R5, thanks to the techniques of simplification and deletion used (see Section 5). The answer to the question which properties from $407 + 294 = 701$ are common to both algebras is found in the following way: running a matching program for comparing the properties of R4 and R5, we have found 217 properties belonging to both sets. The 190 $(407 - 217)$ remaining properties in R4 have been found to be either deduced and simplified properties in R5, or properties to be deduced in R5 (not yet deduced because of our arbitrary stop of the execution).
In the 77 remaining properties in R5 $(294 - 217)$, we discovered that 27 would be derived in R4 later, since their proof uses only axioms of R4.

So, in the 701 properties, 651 are R4 (and also R5). We decomposed these properties into two categories. The first-one corresponds to intuitively obvious properties of topological spaces (or modal logics S4 and S5), the second category contains all other properties. The properties of the second category seem to be really not trivial even for topological spaces with normal set theoretical operations.

Remark. As $(A, \cap, \cup)$ is a lattice, we use symbol $\subseteq$ for the natural lattice ordering defined as follows.

$$a \subseteq b \text{ if and only if } a \cup b = b \text{ and } a \cap b = a.$$

The proved has derived immediately some intuitively obvious properties:

$$\begin{array}{rr}
C1 = 1, & Ia \subseteq a, \\
C\sim 1 = \sim 1, & a \subseteq Ca, \\
I\sim 1 = \sim 1, & Ia \subseteq Ca, \\
C(a \cup b) = Ca \cup Cb, & Ia \cap b \subseteq a, \\
CCa = Ca, & a \cap b \subseteq Ca, \\
\sim Ca = I\sim a, & \sim Ia = C\sim a.
\end{array}$$

We list below some, much less intuitive properties derived by the prover.

$$
\begin{aligned}
Ia \cap b \cap c &\subseteq C(Ia \cap b) \cap a, \\
a \cap Ib &\subseteq C(a \cap Ib) \cap b, \\
I(a \cup b) &\subseteq (I(Ca \cup b) \cap a) \cup (I(Ca \cup b) \cap b), \\
Ia \cap Ib \cap Ic &\subseteq a \cap b \cap C(a \cap b \cap c), \\
Ia \cap b &\subseteq I(a \cup c) \cap C(b \cap Ia), \\
a \cap Ib &\subseteq a \cap C(a \cap C(a \cap Ib)) \cap b, \\
b \cap Ia &\subseteq C(C(b \cap Ia) \cap a), \\
Ca \cap I(Ia \cup b) \cap c &\subseteq Ia \cup (Ca \cap b),
\end{aligned}
$$

$$
\begin{aligned}
(I(a \cup b) \cap Ic \cap C(c \cap a)) \cup (I(a \cup b) \cap Ic \cap C(c \cap b)) &= Ic \cap I(a \cup b), \\
(I(a \cup c) \cap I(a \cup b) \cap C(c \cap b)) \cup (I(a \cup c) \cap I(a \cup b) \cap Ca) &= I(a \cup b) \cap I(a \cup c).
\end{aligned}
$$

3.3 Purely R5 Properties

After 5000 properties of R5 deduced, our prover has kept only 294 of them. We subtracted from them the common 217 properties with R4 and the 27 properties deduced by using only R4 axioms. The 50 properties left are *strong candidates* for being *purely* R5 properties, as their proofs used the additional R5 axiom $CIa = Ia$. This does not affirm yet that they are purely R5 properties, because we have not yet proved they do not have other R4 proof. However for some of them, we were able to prove that they are purely R5, using the following process: given a strong candidate P, we have shown that the specific R5 axiom $CIa = Ia$ is a consequence of R4 plus P. Here are some of these *purely R5 properties*:

$$
\begin{aligned}
ICa &= Ca, \\
a &\subseteq I(Ca \cup b), \\
I(Ca \cup Ib) &= Ca \cup Ib, \\
C(Ca \cap Ib) &= Ca \cap Ib.
\end{aligned}
$$

Observations. The prover has a tendency to generate larger and larger formulas. It hence tries to simplify them into smaller ones. But this is not always possible and the study of these complicated formulas is sometimes interesting. For example, we have found (by direct examination of the 50 R5 formulas) the following 2 variables property:

$$I(Ia \cup Ib) = Ia \cup Ib.$$

We have also found the 3 and 4 variables properties:

$$
\begin{aligned}
I(Ia \cup Ib \cup Ic) &= Ia \cup Ib \cup Ic, \\
I(Ia \cup Ib \cup Ic \cup Id) &= Ia \cup Ib \cup Ic \cup Id.
\end{aligned}
$$

They are in fact the 3 and 4 variables generalizations of the first 2 variables property. It is easy to see that they follow an obvious pattern listed below (where $m > 0$).

$$I(Ia_1 \cup \ldots \cup Ia_m) \ = \ Ia_1 \cup \ldots \cup Ia_m.$$

The proof by mathematical induction that this pattern is an R5 property is straightforward.
There are many other patterns. For example the following formulas

$$\begin{array}{cc}
I(Ca \cup Cb) = Ca \cup Cb, & C(Ca \cap Cb) = Ca \cap Cb, \\
I(Ca \cup Ib) = Ca \cup Ib, & C(Ca \cap Ib) = Ca \cap Ib, \\
I(Ia \cup Ib) = Ia \cup Ib, & C(Ia \cap Ib) = Ia \cap Ib.
\end{array}$$

together with their 3 and 4 variables generalizations can be described by the next two patterns ($n + m > 0$).

$$I(Ca_1 \cup \ldots \cup Ca_n \cup Ib_1 \cup \ldots \cup Ib_m) = Ca_1 \cup \ldots \cup Ca_n \cup Ib_1 \cup \ldots \cup Ib_m,$$
$$C(Ca_1 \cap \ldots \cap Ca_n \cap Ib_1 \cap \ldots \cap Ib_m) = Ca_1 \cap \ldots \cap Ca_n \cap Ib_1 \cap \ldots \cap Ib_m.$$

Below is a method of construction of a formal proof in R5 of the first of these two generalized properties. First we use the following derivation to show how it is possible to add a $m + 1^{\text{st}}$ Ia to the union of already obtained m I's. (In the first deduction, b_1 is chosen equal to $Ia_1 \cup Ia_2$.)

$$\frac{\underbrace{I(Ia_1 \cup Ia_2)} = Ia_1 \cup Ia_2 \qquad I(\underbrace{Ib_1} \cup Ib_2 \cup \ldots Ib_m) = Ib_1 \cup Ib_2 \cup \ldots Ib_m}{I(\overbrace{Ia_1 \cup Ia_2} \cup Ib_2 \cup \ldots Ib_m) = I(Ia_1 \cup Ia_2) \cup Ib_2 \cup \ldots Ib_m}$$

$$\frac{\underbrace{I(Ia_1 \cup Ia_2)} = Ia_1 \cup Ia_2 \qquad I(Ia_1 \cup Ia_2 \cup Ib_2 \cup \ldots Ib_m) = \underbrace{I(Ia_1 \cup Ia_2)} \cup Ib_2 \cup \ldots Ib_m}{I(Ia_1 \cup Ia_2 \cup Ib_2 \cup \ldots Ib_m) = \overbrace{Ia_1 \cup Ia_2} \cup Ib_2 \cup \ldots Ib_m}$$

Secondly, as the below derivation shows, we transform an Ia formula into a Ca using the property $ICa = Ca$ and a substitution of Ca for c_1.

$$\frac{\underbrace{ICa} = Ca \qquad I(Cb_1 \cup \ldots Cb_n \cup \underbrace{Ic_1} \cup Ic_2 \cup \ldots Ic_m) = Cb_1 \cup \ldots Cb_n \cup Ic_1 \cup Ic_2 \cup \ldots Ic_m}{I(Cb_1 \cup \ldots Cb_n \cup \overbrace{Ca} \cup Ic_2 \cup \ldots Ic_m) = Cb_1 \cup \ldots Cb_n \cup ICa \cup Ic_2 \cup \ldots Ic_m}$$

$$\frac{\underbrace{ICa} = Ca \qquad I(Cb_1 \cup \ldots Cb_n \cup Ca \cup Ic_2 \cup \ldots Ic_m) = Cb_1 \cup \ldots Cb_n \cup \underbrace{ICa} \cup Ic_2 \cup \ldots Ic_m}{I(Cb_1 \cup \ldots Cb_n \cup Ca \cup Ic_2 \cup \ldots Ic_m) = Cb_1 \cup \ldots Cb_n \cup \overbrace{Ca} \cup Ic_2 \cup \ldots Ic_m}$$

It is obvious from above that once we know how to add a I operator and how to transform it into a C, the general property mentioned earlier is R5.

4 Rough R4, R5 and Boolean S4, S5 Algebras

The rough R4 algebra is the quasi-Boolean version of the topological Boolean algebra. The topological Boolean algebras are algebraic models (see [10]) for the modal logic S4, where the interior I and closure C operations correspond to modal operators $\Box$ and $\Diamond$, respectively.

The topological Boolean algebras such that each open element is closed and each closed element is open form algebraic models for the modal logic S5. This justifies the following definition.

Definition 5 Boolean S4, S5 algebras. Any topological Boolean algebra $(A, 1, \cap, \cup, \neg, I)$ is called a Boolean S4 algebra.
An S4 algebra $(A, 1, \cap, \cup, \neg, I)$ such that for any $a \in A$, $CIa = Ia$, where $Ca = \neg I \neg a$, is called a Boolean S5 algebra.

It is obvious from the representation theorem for R5 algebras (see Section 2) that the principal Boolean property

$$a \cup \sim a = 1$$

does not hold neither in R5 nor in R4. It was proved in [18] that when we add the above property to the axioms of the quasi-Boolean algebra we obtain a Boolean algebra. The quasi complementation $\sim$ becomes in this case a classical set theoretical complementation.

This proves the following theorem.

Theorem 6. *A R4 (R5) algebra $(A, 1, \cup, \cap, \sim, I)$ with one of the following additional axioms (where 0 denotes ~ 1)*

$$a \cup \sim a = 1 \quad or \quad a \cap \sim a = 0$$

is called a S4 (S5) topological Boolean algebra.

We have added those two axioms to the set of axioms of R4 (R5, respectively) and let the prover run.

The prover has derived more than 300 properties for these topological S4, S5 Boolean algebras. Here are some we find interesting.

$$\begin{array}{ll} Ia \cup C \sim a = 1, & Ia \cap C \sim a = 0, \\ Ca \cup I \sim a = 1, & Ca \cap I \sim a = 0, \end{array}$$

$$\begin{array}{r} (Ca \cap Cb) \cup (Ca \cap c) \cup (I \sim b \cap \sim c) \cup I \sim a = 1, \\ (Ca \cap b) \cup Cc \cup (I \sim c \cap I \sim a) \cup (I \sim c \cap \sim b) = 1, \\ (C(a \cap I \sim b) \cap Cb) \cup (C(a \cap I \sim b) \cap \sim a) \cup (a \cap I \sim b) = C(a \cap I \sim b), \\ (Ia \cap C(a \cap C \sim a)) \cup (a \cap C \sim a) = a \cap C(a \cap C \sim a), \\ (I \sim a \cap Ib) \cup (\sim a \cap b \cap Ca) \cup (\sim a \cap b \cap C \sim b) = \sim a \cap b, \\ (Ca \cap \sim b) \cup (Ca \cap c) \cup (b \cap \sim c) \cup I \sim a = 1, \\ (Ia \cap b) \cup Cc \cup (I \sim c \cap C \sim a) \cup (I \sim c \cap \sim b) = 1. \end{array}$$

5 daTac — A Tool for Automated Deduction

The theorem prover daTac[3], for *Déduction Automatique dans des Théories Associatives et Commutatives*, has been developed at Loria, Nancy (France). This software is written in CAML Light (18000 lines), a language of the ML family.
daTac can be used for theorem proving or for straightforward deduction. It manipulates formulas of first order logic with equality expressed in a clausal form $A_1 \wedge \ldots \wedge A_n \Rightarrow B_1 \vee \ldots \vee B_m$. The deduction techniques implemented are detailed in [21]. We present here only a short overview of them.

5.1 Deduction Techniques

The prover is based on the first order logic with equality, hence the clauses use the equality predicate. But, we do not need to state all equality axioms. The symmetry, transitivity and functional reflexivity of the equality are simulated by a deduction rule called *Paramodulation* [22]. Its principle is to apply replacements in clauses. A paramodulation step in a positive literal is defined by

$$\frac{L_1 \Rightarrow l = r \vee R_1 \qquad L_2 \Rightarrow A[l'] \vee R_2}{(\ L_1 \wedge L_2 \Rightarrow A[r] \vee R_1 \vee R_2\ )\sigma}$$

where σ is a substitution (a mapping replacing variables by terms) unifying the term l and the subterm l' of A, i.e. $l\sigma$ is equal to $l'\sigma$. This last subterm is replaced by the right-hand side r of the equality $l = r$, and the substitution is applied on the whole deduced clause.
A similar paramodulation rule is defined for replacements in negative literals, i.e. literals on the left-hand side of the $\Rightarrow$ sign.

The reflexivity property of the equality predicate is simulated by a rule called *Reflexion*, defined by:

$$\frac{l = r \wedge L \Rightarrow R}{(\ L \Rightarrow R\ )\sigma}$$

where the substitution σ unifies the terms l and r.

The *Resolution* rule [20] permits to deal with the other predicate symbols than the equality:

$$\frac{A_1 \wedge L_1 \Rightarrow R_1 \qquad L_2 \Rightarrow A_2 \vee R_2}{(\ L_1 \wedge L_2 \Rightarrow R_1 \vee R_2\ )\sigma}$$

where σ unifies A_1 and A_2.

5.2 Strategies

In order to limit the number of possible deductions and to avoid useless deductions, the prover uses several strategies of deduction.

The first one is an *ordering strategy* [5]. It uses an ordering for comparing the terms and for orienting the equations. Hence, when a paramodulation step

[3] daTac Home Page (in the PROTHEO Group at the Loria):
`http://www.loria.fr/equipes/protheo/PROJECTS/DATAC/datac.html`

is applied from an equation $l = r$, it is checked that a term is never replaced by a bigger one, i.e. that the term r is not greater than the term l for this ordering. This technique is similar to the use of a rewriting rule $l \rightarrow r$.
The ordering is also used for selection of literals. For instance, it is possible to impose the condition that each deduction step has to use the maximal literal of a clause.

Another essential strategy is the *simplification*. We have defined some simplification rules whose purpose is to replace a clause by a simpler one, using term rewriting.
We have also defined some *deletion rules*. For instance, clauses which contain a positive equation $l = l$, or a same atom on the left-hand side and on the right-hand side of the implication sign $\Rightarrow$, are deleted. Another deletion technique is the *subsumption*, which can be schematized by: if a clause $L \Rightarrow R$ belongs to a set of clauses S, then any clause of the form $L\sigma \wedge L' \Rightarrow R\sigma \vee R'$ can be removed from S.

5.3 Deduction modulo E

The most important feature of daTac is the deduction modulo a set of equations E. The motivation for such deduction is the following.

The *commutativity property* of an operator f, $f(a,b) = f(b,a)$, cannot be oriented as a rewrite rule. This is a major problem when applying paramodulation steps.

Also, the *associativity property* of an operator f, $f(f(a,b),c) = f(a,f(b,c))$, has the disadvantage to provoke infinite sequences of paramodulation steps. For example, from an equation $f(d_1,d_2) = d_3$, we can derive

$$\begin{aligned} f(d_1, f(d_2, e_1)) &= f(d_3, e_1) \\ f(d_1, f(f(d_2, e_1), e_2)) &= f(f(d_3, e_1), e_2) \\ &\vdots \end{aligned}$$

Moreover, when these two properties (commutativity and associativity) are combined, it becomes very difficult to deduce useful clauses. For example, there are 1680 ways to write the term $f(a_1, f(a_2, f(a_3, f(a_4, a_5))))$, where f is associative and commutative. These 1680 terms are all semantically equivalent but none of them can be omitted, for the completeness of the deductions.

So, daTac is defined for being run modulo a set of equations E, composed of commutativity, and associativity and commutativity properties. These equations do not appear in the set of initial clauses. They are simulated by specific algorithms for equality checking, pattern matching and unification, in conjunction with some specially adapted paramodulation rules.

5.4 Advantages of daTac

daTac is an entirely automatic tool which, given a set of clauses, deduces new properties, consequences of these clauses. daTac is refutationally complete: if a set of clauses is incoherent, it will find a contradiction. The only reason that may it fail in its search for a contradiction is the limit of the memory size of the computer.

The implemented techniques involve ordering and simplification strategies combined with a deduction system based on paramodulation and resolution, as mentioned earlier, but other important strategies are also available, as the superposition and basic strategies [25].

At the end of an execution, the user can ask for a lot of extra information. Especially the prover can present a proof of a derived property, or of the contradiction found. Many statistics are also available, such as the number of deduction steps, the number of simplification steps, and the number of deletions.

References

1. Banerjee, M.: A Cathegorical Approach to the Algebra and Logic of the indiscernible. Ph.D dissertation, Mathematics Department, University of Calcutta, India (1994).
2. Banerjee, M., Chakraborty, M. K.: Rough Algebra. Bull. Polish Acad. Sci. (Math.), 41(4):299–297 (1993).
3. Białynicki-Birula, A., Rasiowa, H.: On the representation of quasi-Boolean algebras. Bull. Ac. Pol. Sci. Cl. III, 5 (1957), pp. 259–261.
4. Bronikowski, Z.: Algebraic Structures of Rough Sets. In Ziarko, W., editor, Rough Sets, Fuzzy Sets and Knowledge Discovery, Workshops in Computing, pp. 242–247, Springer Verlag (1994).
5. Hsiang, J., Rusinowitch, M.: Proving Refutational Completeness of Theorem Proving Strategies: The Transfinite Semantic Tree Method. Journal of the ACM, 38(3):559–587 (1991).
6. Kelly, J.: General Topology. Van Nostrand (1955).
7. Kripke, S.: Semantics Analysis of Intuitionistic Logic. In Crossley, J.N., and Dummet, M.N.E., editors, Proc. of the Eight Logic Colloqium, Oxford 1963, pp. 92–130, North Holland Publishing C. (1965).
8. Kuratowski, C.: Topologie I, II. Warszawa (1958, 1961).
9. McKinsey, J. C. C.: A solution of the decision problem for the Lewis systems S.2 and S.4 with an application to topology. The Journal of Symbolic Logic 6, pp. 117–188 (1941).
10. McKinsey, J. C. C., Tarski, A.: The Algebra of Topology. Annales of Mathematics, 45:141–191 (1944).
11. McKinsey, J. C. C., Tarski, A.: On Closed Elements in Closure Algebras. Annales of Mathematics, 47:122–162 (1946).
12. Moisil, G. C.: Recherches sur l'algebre de la logique. Annales Sc. de l'Univerite de Jassy 22 (1935), pp. 1–117.
13. Nöbeling, G.: Grundlagen der analitishen Topologie. Berlin, Götingen, Heilderberg (1954).
14. Novotny, M., Pawlak, Z.: On Rough Equalities. Bull. Polish Acad. Sci. (Math.), 33(1-2):99–104 (1985).
15. Orłowska, E.: Semantics of vague concepts. In Dorn, G., Weingartner, P., editors, Foundations of Logic and Linguistics, Selected Contributions to the 7th International Congress of Logic, Methodology and Philosophy of Science, Saltzburg, Plenum Press, pp. 465–482 (1983).

16. Pagliani, P.: A Pure Logic-Algebraic Analysis of Rough Top and Rough Bottom Equalities. In Ziarko, W., editor, Rough Sets, Fuzzy Sets and Knowledge Discovery, Workshops in Computing, pp. 227–236, Springer Verlag (1994).
17. Pawlak, Z.: Rough Sets. Int. J. Comp. Inf. Sci., 11:341–356 (1982).
18. Rasiowa, H.: An Algebraic Approach to Non-Classical Logics. In Studies in Logic and the Foundations of Mathematics, vol. 78, North-Holland Publishing Company, Amsterdam, London - PWN, Warsaw (1974).
19. Rasiowa, H., Sikorski, R.: A proof of completeness theorem of Gódel. Fundamenta Mathematicae 37, pp. 193-200 (1950).
20. Robinson, J. A.: A Machine-oriented Logic Based on the Resolution Principle. Journal of the ACM, 12:23–41 (1965).
21. Rusinowitch, M., Vigneron, L.: Automated Deduction with Associative-Commutative Operators. Applicable Algebra in Engineering, Communication and Computing, 6(1):23–56 (1995).
22. Robinson, G. A., Wos, L. T.: Paramodulation and First-order Theorem Proving. In Meltzer, B., Mitchie, D., editors, Machine Intelligence, vol. 4, pp. 135–150, Edinburgh University Press (1969).
23. Stone, M. H.: Boolean Algebras and their Relation to Topology. In Proc. Nat. Ac. Sci., vol. 20, pp. 197–202 (1934).
24. Tarski, A.: Grundzüge des Syntemenkalküls. Ester Teil, Fundamenta Mathematicae 25, pp. 503–526 (1935).
25. Vigneron, L.: Associative-Commutative Deduction with Constraints. In Bundy, A., editor, Proceedings 12th International Conference on Automated Deduction, Nancy (France), vol. 814 of Lecture Notes in Artificial Intelligence, pp. 530–544, Springer-Verlag (1994).
26. Wasilewska, A.: Topological Rough Algebras. In Lin, T. Y., editor, Rough Sets and Database Mining, Kluwer (to appear).
27. Wasilewska, A., Banerjee, M.: Rough Sets and Topological Quasi-Boolean Algebras. In ACM CSC'95 23rd Annual Workshop on Rough Sets and Database Mining, Conference Proceedings, San Jose State University, San Jose, California, pp. 61–67 (1995).
28. Wasilewska, A., Banerjee, M.: Representation Theorem for Topological Rough Algebras. Bull. Polish Acad. Sci. (Math.) (to appear).

This article was processed using the LaTeX macro package with LMAMULT style