

# Heyting Algebras with Boolean Operators for Rough Sets and Information Retrieval applications

Eric SanJuan <sup>a</sup>

<sup>a</sup> *LIA, Université d'Avignon et des pays du Vaucluse,  
339 chemin des Meinajariès, BP1228, 84911 AVIGNON Cedex 9 (FRANCE)  
E-mail : eric.sanjuan@univ-avignon.fr*

---

## Abstract

This paper presents an algebraic formalism for reasoning on **finite** increasing sequences over Boolean algebras in general and on generalizations of Rough Set concepts in particular. We argue that these generalizations are suitable for modeling relevance of documents in an Information Retrieval system.

*Key words:* Heyting algebras, Many Valued Logics, Rough Sets, Information Retrieval, Query Expansion, Implicative relations.

---

## 1 Introduction

In the 40's, Moisil found a construction of centered 3-valued Łukasiewicz algebras. In this construction, Moisil considered the set of all pairs  $(b_1, b_2)$  over a Boolean algebra with  $b_1 \leq b_2$  as the universe of a Heyting algebra with some additional unary operators. A strong generalization of this idea was investigated by Nadiu in (22), meanwhile similar constructions have been developed for generalizations of Post Algebras by Cat Ho and Rasiowa in (10) and by Serfati in (30) or for other algebras related to Łukasiewicz's ones by Iturrioz in (15).

It has been pointed out by Iturrioz in (14) how Moisil's construction could contribute to the understanding of the logic for Rough Sets introduced by Düntsch in (8) and the corresponding algebraic structure studied in (25). Rough Set theory has been developed since 1991 (26) beyond its primary goals of reasoning with different types of information incompleteness (24) and offers today a general framework for Data Mining (34) and Information Retrieval (12). Following Pawlak's idea, a Rough Set is a pair of approximations of a set whose

internal objects cannot be clearly discerned from external ones, due to lack of information. They are based on the concept of approximation space which is a frame  $(Ob, IND)$  such that  $Ob$  is a set of objects and  $IND$  an equivalence relation on  $Ob$ , called the indiscernibility relation.  $IND$  generates a monadic operator  $U$  and its dual  $L$  on the Boolean algebra  $(\wp(Ob), \cap, \cup, -, \emptyset, Ob)$ . The rough sets considered by Pawlak are the pairs  $(L(X), U(X))$  for any  $X \subseteq Ob$ . It has been shown in (8; 14) that the collection of all rough sets of an approximation space is a 3-valued Łukasiewicz algebra.

This paper presents an algebraic formalism introduced in (28) for reasoning on **finite** increasing sequences over Boolean algebras in general and on generalizations of Rough Set concept in particular. We associate with every finite poset  $\mathbf{T}$  a class of algebras called  $T$ -Rough algebras. The axiomatization of these algebras is simple and quite analogous to that of the  $L'_T$  propositional calculus introduced by A. Nour in (23).

Our work is closely related to *plain semi Post algebras* introduced in (10). Indeed, the fundamental examples of these algebras are constructed with partially ordered descending sequences over a Boolean algebra. Thus, plain semi Post algebras are Heyting algebras with unary operators and constants. Extensions of these algebraic systems, called *Perception Logics*, have been introduced by Rasiowa in (33) for modeling cooperating systems fully communicating. Perception logics can also be interpreted in knowledge based distributed systems.

Following Post's tradition, simple plain semi Post algebras are primal (every  $n$ -argument operation is definable). It follows from this strong property that, within this algebraic framework, we can only consider full collections of partially ordered sequences, whereas the collection of rough sets of an approximation space is not in general the full collection of increasing pairs. We therefore choose to follow the direction given by Łukasiewicz - Moisil algebras to consider many-valued facts.

To illustrate the usefulness of this algebraic formalism we introduce an innovating Information Retrieval (IR) model called  $k$ -Rough IR model where documents and queries are merged into a  $T$ -Rough algebra and a query expansion process is implemented using algebraic operators. In fact, the study of the IR system presented in this paper does not require  $T$ -Rough algebras where  $\mathbf{T}$  is not a chain. Meanwhile, we argue that these generalizations are suitable for further developments of our  $k$ -Rough IR system and other applications. They also give a better understanding of the large variety of algebraic frameworks that were introduced from the 40's to the 90's dealing with Heyting Algebras and MVL (17).

The rest of the paper is divided into two parts.

The first part of the paper presents an overview of Heyting Algebras with Boolean Operators (HABO) related to approximation reasoning and MVL. By Boolean operator we mean a unary function that maps the whole algebra onto a Boolean sub-algebra. A general and unified algebraic framework, called  $T$ -Rough algebras, is introduced in §2. In §3, homomorphisms and quotient algebras of  $T$ -Rough algebras are described. These results are applied in Section 4 to finite algebras. In Section 5, we show the connection between  $T$ -Rough algebras and the propositional calculus introduced and investigated in (23).

The second part deals with motivating applications of HABO. In §6 we introduce a generalization of Rough Set concept that allows the specification of  $k + 1$  degrees of approximation,  $k$  being a finite integer, and we show that these generalizations are Heyting algebras with operators. In §7 we show that this algebraic framework provides the intuition for defining several measures of implicative intensity. These measures are illustrated in §8

Finally, in section 9, we conclude and mention directions for future works.

## 2 Algebraic framework

We associate with every finite poset  $\mathbf{T}$  a class of algebras called  $T$ -Rough algebras that shall allow us to give a unified view of a large family of many-valued systems with a finite range of truth values, dealing with approximating reasoning.

### 2.1 $T$ -Rough algebras

Following (10, def. 1) and (14, def. 3.1), we associate a class of algebras with every poset  $\mathbf{T}$  such that  $|T| \geq 2$ , by the following definition. In this definition, the additional unary operators  $\pi_i$  are analogous to projection mappings of a product of Boolean algebras.

**Definition 1** *Given a finite poset  $\mathbf{T} = (T, \leq)$ , an abstract algebra  $\mathbf{H} = (H, \wedge, \vee, \rightarrow, (\pi_t)_{t \in T}, 0, 1)$  where  $\wedge, \vee, \rightarrow$  are binary operations,  $\pi_t$  for  $t \in T$  are unary operations, and  $0, 1$  are zero-argument operations, is said to be a  **$T$ -Rough algebra** provided the following conditions are satisfied:*

- (p0)  $(H, \wedge, \vee, \rightarrow, 0, 1)$  is a Heyting algebra;
- (p1)  $\pi_t(x \vee y) = \pi_t(x) \vee \pi_t(y)$ ;
- (p2)  $\pi_t(x \wedge y) = \pi_t(x) \wedge \pi_t(y)$ ;
- (p3)  $\pi_t \pi_u(x) = \pi_u(x)$ ;
- (p4)  $\pi_t(0) = 0$ ;

- (p5)  $\pi_t(x) \vee \neg\pi_t(x) = 1$  where  $\neg x = x \rightarrow 0$ ;  
 (p6)  $\pi_t(x \rightarrow y) = \bigwedge_{v \geq t} (\pi_v(x) \rightarrow \pi_v(y))$ ;  
 (p7)  $\bigwedge_{v \in T} \pi_v(x) \vee x = x$ .

for any  $x, y \in H, t, u, v \in T$ .

If  $\mathbf{T}$  is the chain:  $1 \leq \dots \leq k \leq k + 1$ , then  $\mathbf{H}$  is said to be a  **$k$ -Rough algebra**.

On the right-hand side of (p6) and on the left-hand side of (p7),  $\bigwedge$  denotes the greatest lower bound. Since  $T$  is finite, these bounds are reduced to finite conjunctions.

It follows from the above definition and from the fact that the class of all Heyting algebras is equationally definable (21), that given a poset  $T$ , the class of  $T$ -Rough algebras is also equationally definable. We shall denote by  $\mathcal{B}_{\mathbf{T}}$  this **equational class**. Like in (10; 14), the fact that the index set  $T$  of unary operations is ordered is not explicitly described by conditions (p0)–(p7).

The following definition recalls a fundamental example of Heyting Algebra with Boolean Operators (HABO) for algebraic models of MVL (18; 10; 15).

**Definition 2** Let  $\mathbf{B}$  be a Boolean algebra, let  $\mathbf{T} = (T, \leq)$  be a finite poset such that  $T = \{1, \dots, n\}$  is a finite set and  $\leq$  is a partial order on  $T$ . We denote by  $B_T$  the lattice of isotone applications from  $\mathbf{T}$  into  $\mathbf{B}$ :  $f \in B_T$  iff for any  $u, v \in T$ ,  $f(u) \leq f(v)$  whenever  $u \leq v$ . In the sequel we shall denote by  $f_t$  the image  $f(t)$  of  $t$  and we shall identify the application  $f$  with the sequence  $(f_t)_{t \in T}$ .

It is well known (22; 10; 15; 30) that  $(B_T, \wedge, \vee, \rightarrow, 0, 1)$  is a Heyting algebra where the implication  $\rightarrow$  is defined for any  $f, g \in \mathbf{B}_{\mathbf{T}}$  by:

$$(f \rightarrow g)_t = \bigwedge_{u \geq t} (\neg f_u \vee g_u)$$

Moreover, this Heyting algebra is linear (28) iff for any  $u, v, t \in T$ ,  $u \geq t$  and  $v \geq t$  imply  $u \leq v$  or  $v \leq u$ .

**Definition 3** For any  $t \in T$  we define a unary operator  $\pi_t$  on  $B_T$  by setting:  $(\pi_t(f))_u = f_t$ . We denote by  $\mathbf{B}_{\mathbf{T}}$  the Heyting algebra with operators:

$$\mathbf{B}_{\mathbf{T}} = (B_T, \wedge, \vee, \rightarrow, (\pi_t)_{t \in T}, 0, 1)$$

It is straightforward to check that any algebra  $\mathbf{B}_{\mathbf{T}}$  is in  $\mathcal{B}_{\mathbf{T}}$ .

## 2.2 Representation theorem

Let  $T$  be a poset and  $\mathbf{H} = (H, \wedge, \vee, \rightarrow, (\pi_t)_{t \in T}, 0, 1)$  a  $T$ -Rough algebra. We remind the reader that if  $\mathbf{H} = (H, \wedge, \vee, \rightarrow, 0, 1)$  is a Heyting algebra then the algebra  $\mathbf{B}(\mathbf{H}) = (B(H), \wedge, \vee, \neg, 0, 1)$ ,  $B(H)$  being the **set of complemented elements**, is a Boolean algebra. Moreover, the algebra  $(B(H), \wedge, \vee, \rightarrow, 0, 1)$  where  $\rightarrow$  is defined by  $x \rightarrow y = \neg x \vee y$  is a Heyting subalgebra of  $\mathbf{H}$ .

Let  $\pi(H) = \{x \in H : (\forall t \in T) \pi_t(x) = x\}$  be the **set of fixed points** of operations  $\pi_t$ .

Let  $\mathbf{A}$  be an algebra and let  $A$  be its universe (underlying set). A subuniverse  $S$  of  $\mathbf{A}$  (3, Ch. II, def. 2.2) is a subset of  $A$  which is closed under the fundamental operations of  $\mathbf{A}$ . If  $S \neq \emptyset$ , then  $S$  is the universe of a subalgebra of  $\mathbf{A}$ .

**Lemma 4**  $\pi(H)$  is a subuniverse of  $\mathbf{B}(\mathbf{H})$ .

**PROOF.** It follows from (p0) and (p5) that for any  $x \in H$  and  $t \in T$ ,  $\pi_t(x)$  is a complemented element. Hence we have:  $\pi(H) \subseteq B(H)$ .

It follows from (p4) that  $0 \in \pi(H)$ . (p0) yields  $x \rightarrow x = 1$ , thus it follows from (p6) that for any  $t \in T$ ,  $\pi_t(1) = 1$  and consequently  $1 \in \pi(H)$ . Moreover, axioms (p1) and (p2) imply that  $\pi(H)$  is closed with respect to operations  $\wedge$  and  $\vee$ . Finally, to show that if  $x \in \pi(H)$  then  $\neg x \in \pi(H)$ , it is sufficient to prove the following equation:

$$\pi_u(\neg \pi_t(x)) = \neg \pi_t(x) \tag{1}$$

It follows from (p6) and (p4) that:  $\pi_v(\neg y) = \bigwedge_{w \geq v} \neg \pi_w(y)$ . Applying this equality we obtain:

$$\pi_u(\neg \pi_t(x)) = \bigwedge_{w \geq u} \neg \pi_w(\pi_t(x)) = \bigwedge_{w \geq u} \neg \pi_t(x) = \neg \pi_t(x)$$

□

We shall denote by  $\pi(\mathbf{H})$  the Boolean algebra defined on  $\pi(H)$ . It follows from previous Lemma 4 and (p1) that operators  $\pi_t$  are isotone mappings from  $\mathbf{H}$  into the Boolean algebra  $\pi(\mathbf{H})$ .

It is worth mentioning that previous lemma also yields  $\pi(H)$  and  $B(H)$  are subuniverses of  $\mathbf{H}$ .

The following lemma, which is a consequence of axioms (p6) and (p7) shows that for any  $t \in T$ ,  $\pi_t(x)$  can be treated as coordinates of  $x$  in  $\mathbf{H}$ .

**Lemma 5 (Determination principle)** *If  $\pi_t(x) = \pi_t(y)$  for any  $t \in T$ , then  $x = y$ .*

**PROOF.** The proof is analogous to that of determination principle in (15). Assume that  $\pi_t(x) = \pi_t(y)$  for all  $t \in T$ . It follows from axiom (p6) that  $\pi_t(x \rightarrow y) = \bigwedge_{w \geq t} (\pi_w(x) \rightarrow \pi_w(y)) = 1$ . Thus  $\bigwedge_{t \in T} \pi_t(x \rightarrow y) = 1$  and by (p7) we obtain  $x \rightarrow y = 1$ . This implies in a Heyting algebra that  $x \leq y$ . The proof of the other half is alike.  $\square$

Now we will show that for any  $x \in H$ , then the map  $\varphi_x : t \in T \mapsto \pi_t(x)$  is an isotone application.

**Lemma 6** *For any  $u, v \in T$ , if  $u \leq v$  then  $\pi_u(x) \leq \pi_v(x)$ .*

**PROOF.** Similar to (S12) in (15). Since  $\mathbf{H}$  is a Heyting algebra, by (p6) and Lemma 4 we obtain:

$$\pi_u(x) = \pi_u(1 \rightarrow x) = \bigwedge_{w \geq u} (\pi_w(1) \rightarrow \pi_w(x)) = \bigwedge_{w \geq u} \pi_w(x) \leq \pi_v(x)$$

$\square$

The next statement gives an embedding of any  $T$ -Rough algebra into  $\pi(\mathbf{H})_{\mathbf{T}}$ , the lattice of isotone applications from  $\mathbf{T}$  into the boolean algebra  $\pi(\mathbf{H})$ .

**Theorem 7**  *$\mathbf{H}$  is isomorphic to a subalgebra of  $(\pi(\mathbf{H}))_{\mathbf{T}}$ .*

**PROOF.** Let us consider the mapping  $h$  from  $H$  into  $(\pi(H))_{\mathbf{T}}$  defined by :  $h(x) = \varphi_x$ . It follows from Lemma 5 that  $h$  is one-to-one. Axioms (p1),(p2) and Lemma 4 yield that  $h$  is a homomorphism with respect to bounded lattice operations. From axiom (p6) we deduce that  $h$  is a Heyting homomorphism. Finally it follows from axiom (p3) that for any  $u \in T$ ,  $\varphi_{\pi_t(x)}(u) = \pi_u(\pi_t(x)) = \pi_t(x)$  and consequently that  $h(\pi_t(x)) = \pi_t(h(x))$ . Hence,  $h$  is a monomorphism of  $\mathbf{H}$  into  $(\pi(\mathbf{H}))_{\mathbf{T}}$ .  $\square$

### 2.3 Generator of the equational class

We shall denote by  $\underline{\mathbf{2}} = (\underline{\mathbf{2}}, \wedge, \vee, \neg, 0, 1)$  the **simple Boolean algebra**, i.e.  $\underline{\mathbf{2}} = \{0, 1\}$ . The algebra  $\underline{\mathbf{2}}_{\mathbf{T}}$  defined in Example 2 is isomorphic to the algebra:  $(\mathcal{F}(T), \cap, \cup, \text{imp}, g_1, \dots, g_{k+1}, \emptyset, T)$  where  $\mathcal{F}(T)$  is the set of filters  $F$  of  $\mathbf{T}$  (i.e.  $v \in F$  whenever  $u \in F$  and  $v \geq u$ ),  $\text{imp}$  is the binary operation defined by:  $\text{imp}(F, G) = \cup\{H \in \mathcal{F}(T) : F \cap H \subseteq G\}$  and for any  $t \in T$ ,  $g_t$  is the unary operator defined by:  $g_t(F) = T$  if  $t \in F$ ,  $g_t(F) = \emptyset$  otherwise. If  $\mathbf{T}$  is a chain, then  $\text{imp}(F, G) = T$  if  $F \subseteq G$  and  $\text{imp}(F, G) = G$  otherwise. Observe that any subset  $S$  of  $\underline{\mathbf{2}}_T$  closed with respect to  $\wedge, \vee, \rightarrow, 0$  and  $1$  is a sub-universe of  $\underline{\mathbf{2}}_{\mathbf{T}}$ . We are going to show that  $\underline{\mathbf{2}}_{\mathbf{T}}$  is a generator of the equational class  $\mathcal{B}_{\mathbf{T}}$ .

For any Boolean algebra  $\mathbf{B}$ , we denote by  $\mathcal{U}(B)$  the set of its **ultrafilters**. We associate with every  $\mathbf{H} \in \mathcal{B}_{\mathbf{T}}$  the algebra  $\widehat{\mathbf{H}}$  defined by:

$$\widehat{\mathbf{H}} = \left(\underline{\mathbf{2}}^{\mathcal{U}(\pi(\mathbf{H}))}\right)_{\mathbf{T}} \text{ if } \mathcal{U}(\pi(\mathbf{H})) \neq \emptyset \text{ and } \widehat{\mathbf{H}} = \underline{\mathbf{2}}_{\mathbf{T}} \text{ otherwise.}$$

**Lemma 8** *Any algebra  $\mathbf{H} \in \mathcal{B}_{\mathbf{T}}$  is isomorphic to a subalgebra of  $\widehat{\mathbf{H}}$ .*

**PROOF.** If  $\mathcal{U}(\pi(\mathbf{H})) = \emptyset$  then  $\pi(\mathbf{H}) = \{0, 1\}$  and the lemma follows from Theorem 7.

If  $\mathcal{U}(\pi(\mathbf{H})) \neq \emptyset$ , let  $\sigma$  be the mapping from  $H$  into  $\widehat{\mathbf{H}}$  defined by  $\sigma(x) = (\{U \in \mathcal{U}(\pi(\mathbf{H})) : \pi_t(x) \in U\})_{t \in T}$ . Then it follows from Theorem 7 and Stone's representation theorem for Boolean algebras that  $\sigma$  is a monomorphism.  $\square$

The following Lemma is inspired by Exponentiation Theorems in (22, §2).

**Lemma 9** *For any set  $U$  and any finite poset  $\mathbf{T}$  we have:  $(\underline{\mathbf{2}}^U)_{\mathbf{T}} \approx (\underline{\mathbf{2}}_{\mathbf{T}})^U$ .*

**PROOF.** Let  $x : t \in T \mapsto x_t \in \{0, 1\}^U$  be an element of  $(\underline{\mathbf{2}}^U)_{\mathbf{T}}$  and let  $\varphi$  a mapping from  $U$  into  $\underline{\mathbf{2}}_T$ . For any  $t \in T$ , we treat  $x_t$  as a subset of  $U$ , and for any  $u \in U$  we treat  $\varphi(u)$  as an element of  $\mathcal{F}(T)$ .

Consider the mapping  $\overrightarrow{\lambda} : x \mapsto \overrightarrow{x}$  defined by:  $\overrightarrow{x}(u) = \{t \in T : u \in x_t\}$  for any  $u \in U$ , and the mapping  $\overleftarrow{\lambda} : \varphi \mapsto \overleftarrow{\varphi}$  defined by:  $\overleftarrow{\varphi}_t = \{u \in U : t \in \varphi(u)\}$  for any  $t \in T$ .

Since  $u \in x_t \iff t \in \overrightarrow{x}(u)$ ,  $\overrightarrow{\lambda}$  is an order monomorphism. Likewise,  $\overleftarrow{\lambda}$  is also an order monomorphism. Thus  $\overrightarrow{\lambda}$  is an order isomorphism and consequently a Heyting isomorphism since any order isomorphism preserves infinite joins.

By definition of a product of algebras, for any  $t \in T$ ,  $\pi_t$  is defined on  $(\underline{\mathbf{2}}_T)^U$  by:  $(\pi_t(\varphi))(u) = \pi_t(\varphi(u))$ . Furthermore we have the following equivalences:  $\pi_t(\overrightarrow{x}(u)) = T$  iff  $t \in \overrightarrow{x}(u)$  iff  $u \in x_t$  iff for any  $v \in T$ ,  $u \in (\pi_t(x))_v$  iff  $\overrightarrow{\pi_t(x)}(u) = T$ . This yields that  $\overrightarrow{\lambda}(\pi_t(x)) = \pi_t(\overrightarrow{\lambda}(x))$  and consequently, that  $\overrightarrow{\lambda}$  is an isomorphism of  $T$ -Rough algebras.  $\square$

The following theorem shows that there is a *truth-table method* of verifying the truth or falsity of equations in a  $T$ -Rough algebra.

We shall denote by  $V(\underline{\mathbf{2}}_T)$  the algebraic variety generated by the algebra  $\underline{\mathbf{2}}_T$ .

**Theorem 10**  $\mathcal{B}_T = V(\underline{\mathbf{2}}_T)$

**PROOF.** Since  $\underline{\mathbf{2}}_T \in \mathcal{B}_T$ , we obviously have  $\mathcal{B}_T \subseteq V(\underline{\mathbf{2}}_T)$ . Conversely, if  $\mathbf{H} \in \mathcal{B}_T$  then by Lemmas 8 and 9 we obtain that  $\mathbf{H}$  is isomorphic to a subalgebra of  $(\underline{\mathbf{2}}_T)^U$ . This yields that  $\mathbf{H} \in V(\underline{\mathbf{2}}_T)$ .  $\square$

Hence, an equation holds in any  $T$ -Rough algebra iff it holds in  $\underline{\mathbf{2}}_T$ . In particular, for  $T = \{1, 2\}$ , it is easy to check that the unary operation  $\sim$  defined by

$$\sim x = (x \vee \neg x) \wedge \neg \pi_1(x) \quad (2)$$

is a De Morgan negation. Hence, if  $T = \{1, 2\}$ , then  $\mathcal{B}_T$  is the class of three valued Łukasiewicz algebras.

Let us denote by  $\chi$  the unary operation defined by:

$$\chi(x) = \bigwedge_{t \in T} \pi_t(x) \quad (3)$$

$\chi$  is a monadic operator on a Heyting algebra. Such algebraic structures have been intensively studied by Bezhanishvili in (2). If  $T$  is the chain  $1 \leq \dots \leq k+1$ , then  $\chi = \pi_1$ .

**Corollary 11**  $\mathcal{B}_T$  is a discriminator variety.

**PROOF.** For any  $x \in \underline{\mathbf{2}}_T$ ,  $\chi(x) = 1$  if  $x = 1$  and  $\chi(x) = 0$  otherwise. This yields that the ternary term:  $\tau(x, y, z) = (x \wedge \chi(x \leftrightarrow y)) \vee (z \wedge \chi(x \leftrightarrow y))$  where  $x \leftrightarrow y = (x \rightarrow y) \wedge (y \rightarrow x)$ , is a discriminator term on  $\underline{\mathbf{2}}_T$ .  $\square$

It is known (3, Cor. 10.8, Ch. IV) that a finite algebra is primal iff it has a discriminator term, only one automorphism (the identity map) and only one subalgebra (itself). Let  $\mathbf{T}$  be a poset with at least two elements. Since  $\underline{\mathbf{2}}$  is a subalgebra of  $\underline{\mathbf{2}}_{\mathbf{T}}$  and  $\underline{\mathbf{2}} \neq \underline{\mathbf{2}}_{\mathbf{T}}$ , it follows that  $\mathcal{B}_{\mathbf{T}}$  is not a primal variety. This is the main difference with plain semi-Post algebras introduced in (10).

### 3 Deductive Systems

In this section we give a description of  $B_T$  lattices as Boolean products.

For any algebra  $\mathbf{A}$  we shall denote by  $\text{Cong}(\mathbf{A})$  the **set of congruences** on  $\mathbf{A}$ . Let  $\theta$  be a congruence, for any  $x \in A$ , we shall denote by  $\theta(x)$  the **equivalence class** of  $x$ . If  $\mathbf{A}$  is a Heyting algebra (in particular a Boolean algebra), the **collection of filters** of  $\mathbf{A}$  will be denoted by  $\mathcal{F}(\mathbf{A})$ , i.e.  $F \in \mathcal{F}(\mathbf{A})$  iff  $1 \in F$  and  $y \in F$  whenever  $x, x \rightarrow y \in F$ .

In the sequel,  $\mathbf{T}$  is a finite poset and  $\mathbf{H}$  is an arbitrary  $T$ -Rough algebra. We shall denote by  $H$  the universe of  $\mathbf{H}$ .

**Definition 12** *A deductive system of  $\mathbf{H}$  is a filter  $D \in \mathcal{F}(\mathbf{H})$  such that for any  $x \in D$ ,  $\chi(x) \in D$ . We shall denote by  $\mathcal{D}(\mathbf{H})$  the collection of all deductive systems of  $\mathbf{H}$ .*

#### 3.1 Congruence Lattice

Next lemma shows that there is a bijection between congruences and deductive systems of  $T$ -Rough algebras.

**Lemma 13** *For any  $\theta \in \text{Cong}(\mathbf{H})$ ,  $\theta(1) \in \mathcal{D}(\mathbf{H})$ . Conversely, if  $D \in \mathcal{D}(\mathbf{H})$  then the binary relation  $\theta_D$  defined on  $H$  by:  $(x, y) \in \theta_D \iff x \leftrightarrow y \in D$  is a congruence on  $\mathbf{H}$  such that  $\theta_D(1) = D$ .*

**PROOF.** Since  $\mathbf{H}$  is a Heyting algebra, it is well known (32, §13, Ch 1) that  $\theta(1)$  is a filter. Moreover, since  $\theta$  is a congruence on  $\mathbf{H}$ , if  $(x, 1) \in \theta$  then we have  $(\pi_t(x), \pi_t(1)) \in \theta$  for any  $t \in T$ . Since  $\mathbf{T}$  is finite, it follows by Lemma 4 that  $\chi(x) \in \theta(1)$ . This asserts that  $\theta(1)$  is a deductive system.

Conversely, it is also well known that  $\theta_D$  is a congruence with respect to Heyting operations:  $\wedge, \vee, \rightarrow, 0, 1$  and  $\theta_D(1) = D$ . Now assume that  $x \rightarrow y \in D$ . Then by definition of  $D$ ,  $\chi(x \rightarrow y) \in D$ . Therefore, for any  $t \in T$ ,  $\pi_t(x) \rightarrow \pi_t(y) \in D$  since by axiom (p6)  $\chi(x \rightarrow y) \leq \pi_t(x \rightarrow y) \leq \pi_t(x) \rightarrow \pi_t(y)$ . This

fact permits to state that if  $(x, y) \in \theta_D$  then  $(\pi_t(x), \pi_t(y)) \in \theta_D$  for any  $t \in T$ . Hence  $\theta_D$  is a congruence on  $\mathbf{H}$ .  $\square$

It follows from Lemma 13 that the mapping  $\theta \in \text{Cong}(\mathbf{H}) \mapsto \theta(\mathbf{1}) \in \mathcal{D}(\mathbf{H})$  is an order isomorphism and consequently that the lattice of congruences on  $\mathbf{H}$  is isomorphic to the lattice of deductive systems of  $\mathbf{H}$ .

We denote by  $\chi^{-1}$  the pullback mapping of  $\chi$ .

**Lemma 14** *For any  $D \in \mathcal{D}(\mathbf{H})$ ,  $D \cap \pi(H) \in \mathcal{F}(\pi(\mathbf{H}))$ . Conversely, if  $F \in \mathcal{F}(\pi(\mathbf{H}))$  then  $\chi^{-1}(F) = \{x \in H : \chi(x) \in F\}$  is a deductive system of  $\mathbf{H}$  such that  $\chi^{-1}(F) \cap \pi(H) = F$ .*

**PROOF.** Obviously  $D \cap \pi(H) \in \mathcal{F}(\pi(\mathbf{H}))$  since  $\pi(H)$  is a subuniverse of  $\mathbf{H}$ .

To prove that  $\chi^{-1}(F)$  is a deductive system, it is sufficient to show that it is a filter of  $\mathbf{H}$ . Let  $F \in \mathcal{F}(\pi(\mathbf{H}))$ . Since  $\chi$  is a monadic operator on  $H$ ,  $F \subseteq \chi^{-1}(F)$ . In particular  $1 \in \chi^{-1}(F)$ .

Assume that  $\{\chi(x), \chi(x \rightarrow y)\} \subseteq F$ , then  $\pi_t(x) \in F$  and by axiom (p6),  $\pi_t(x) \rightarrow \pi_t(y) \in F$  for any  $t \in T$ . Thus, since  $F$  is a filter of  $\pi(\mathbf{H})$ ,  $\pi_t(y) \in F$  for any  $t \in T$ . This yields that  $\chi(y) \in F$  and completes the proof.  $\square$

The main statement for representability theory connected with Boolean algebras deals with congruence lattices.

**Theorem 15**  $(\text{Cong}(\mathbf{H}), \subseteq) \approx (\mathcal{F}(\pi(\mathbf{H})), \subseteq)$ .

**PROOF.** It follows from Lemmas 13 and 14 that the mappings:

$$\lambda_\theta : \theta \in \text{Cong}(\mathbf{H}) \mapsto \theta(\mathbf{1}) \cap \pi(H) \in \mathcal{F}(\pi(\mathbf{H})) \quad (4)$$

$$\lambda_F : F \in \mathcal{F}(\pi(\mathbf{H})) \mapsto \theta_{\chi^{-1}(F)} \in \text{Cong}(\mathbf{H}) \quad (5)$$

are order monomorphisms with respect to inclusion.  $\square$

From Theorem 15, we obtain a Boolean product representation of algebras in  $\mathcal{B}_T$  and a description of simple algebras. For an explicit definition of Boolean products we refer the reader to (3, Ch IV, §8).

**Corollary 16** *For any ultrafilter  $U \in \mathcal{U}(\pi(\mathbf{H}))$ , let  $\mathbf{H}/\mathbf{U}$  be the quotient algebra  $\mathbf{H}/\theta_{\chi^{-1}(U)}$ . Then the following statements hold:*

- (1)  $\mathbf{H}/\mathbf{U}$  is a simple algebra whenever  $U \in \mathcal{U}(\pi(\mathbf{H}))$ ,
- (2) the simple algebras of  $\mathcal{B}_{\mathbf{T}}$  are the subalgebras of  $\underline{\mathbf{2}}_{\mathbf{T}}$ ,
- (3) if  $\mathbf{H}$  is not a simple algebra then the mapping  $\mathbf{H} \mapsto \mathbf{H}/\mathbf{U}$  is a representation as a Boolean product.

**PROOF.** Since the mapping  $\lambda_F$  defined by (5) in the proof of Theorem 15 is an order homomorphism, the set of maximal congruences of  $\mathbf{H}$  is:

$$\{\theta_{\chi^{-1}(U)} : U \in \mathcal{U}(\pi(\mathbf{H}))\}$$

This implies item 1 as well as the following equivalences, which prove item 2:  $\mathbf{H}$  is a simple algebra iff  $\mathcal{U}(\pi(\mathbf{H})) = \emptyset$  iff  $\pi(H) = \{0, 1\}$  iff by Theorem 7,  $\mathbf{H}$  is isomorphic to a subalgebra of  $\underline{\mathbf{2}}_{\mathbf{T}}$ .

It follows that if  $\mathbf{H}$  is not a simple algebra then  $\Phi : \mathbf{H} \mapsto \mathbf{H}/\mathbf{U}$  is a representation as a subdirect product. Then it is easy to verify that  $\Phi$  is a Boolean representation where  $\mathcal{U}(\pi(\mathbf{H}))$  is endowed with the Boolean space topology. This shows item 3.  $\square$

It follows from previous corollary that for any integer  $k$ , there are  $2^k$  simple  $k$ -Rough algebras.

### 3.2 Weak deduction theorem

Following A. Monteiro (19, Th 3.9), we assert a weak deduction theorem that is the algebraic counterpart of A. Nour's theorem in (23, §3).

**Theorem 17** *Let  $\mathbf{H} \in \mathcal{B}_{\mathbf{T}}$ . For any  $D \in \mathcal{D}(\mathbf{H})$  and  $h \in H$ , the deduction system  $D_h$  generated by  $D \cup \{h\}$  is the set:  $\{x \in H : \chi(h) \rightarrow x \in D\}$ .*

**PROOF.** The theorem is a consequence of the deduction theorem for Boolean algebras and of Lemma 14. Indeed, from Lemma 14 we obtain the following equivalences for any  $D_1, D_2 \in \mathcal{D}(\mathbf{H})$ :

$$D_1 \subseteq D_2 \iff D_1 \cap \pi(H) \subseteq D_2 \cap \pi(H) \tag{6}$$

$$h \in D_1 \iff \chi(h) \in D_1 \cap \pi(H) \tag{7}$$

Then we have the following equalities:

$$D_h = \chi^{-1}(\{x \in \pi(H) : \chi(h) \rightarrow x \in D \cap \pi(H)\}) \quad (8)$$

$$= \{y \in H : \chi(h) \rightarrow \chi(y) \in D\} \quad (9)$$

Observe that by axioms (p1),(p3) and distributivity we have for any  $b \in \pi(H)$ ,  $\chi(b \vee y) = b \vee \chi(y)$ . Moreover, in a Heyting algebra, if  $b$  is a complemented element then  $b \rightarrow y = \neg b \vee y$ . Finally we obtain:

$$D_h = \{y \in H : \chi(\chi(h) \rightarrow y) \in D\} = \{y \in H : \chi(h) \rightarrow y \in D\}$$

□

## 4 Finite algebras

Given a finite poset  $\mathbf{T}$ , this section is devoted to finite  $T$ -Rough algebras. We apply results from previous §3 to obtain a description of finite algebras as direct products of simple algebras.

If  $\mathbf{P} = (P, \leq)$  is a poset, for any  $p \in P$  and  $Q \subseteq P$  we denote by  $\uparrow_Q p$  the set  $\{x \in Q : x \geq p\}$ . If  $\mathbf{B}$  is a Boolean algebra, we denote by  $At(\mathbf{B})$  the set of its atoms.

### 4.1 Direct products

Note that if  $\mathbf{H}$  is a finite algebra in  $\mathcal{B}_{\mathbf{T}}$ , then  $\mathcal{F}(\pi(\mathbf{H})) = \{\uparrow_{\pi(H)} b : b \in \pi(H)\}$  and  $\chi^{-1}(\uparrow_{\pi(H)} b) = \uparrow_H b$  for any  $b \in \pi(H)$ . It follows that  $\mathcal{D}(\mathbf{H}) = \{\uparrow_H b : b \in \pi(H)\}$  and for any  $b \in \pi(H)$ ,  $(x, y) \in \theta_{\uparrow_H b}$  iff  $x \leftrightarrow y \geq b$  where  $\theta_{\uparrow_H b}$  is defined in Lemma 13.

**Lemma 18** *If  $\mathbf{H}$  is finite and not simple then for any  $a \in At(\pi(H))$ , the algebra  $\mathbf{H} / \uparrow_{\pi(\mathbf{H})} \mathbf{a}$  is simple and:*

$$\mathbf{H} \approx \prod_{a \in At(\pi(H))} \mathbf{H} / \uparrow_{\pi(\mathbf{H})} \mathbf{a} \quad (10)$$

**PROOF.** The lemma is a consequence of the fact that the mapping  $\lambda_F$  defined by (5) is an order isomorphism.

If  $\mathbf{H}$  is not simple then  $At(\pi(H)) \neq \emptyset$ . Since for any  $a \in At(\pi(H))$  we have  $\uparrow_{\pi(H)} a \in \mathcal{U}(\pi(B))$ , it follows that  $\theta_{\uparrow_H a}$  is maximal and consequently  $\mathbf{H} / \uparrow_{\pi(\mathbf{H})} \mathbf{a}$

is simple. Moreover,  $\theta_{\uparrow_{\pi(H)}a}$  and  $\theta_{\uparrow_{\pi(H)}\neg a}$  are factor congruences on  $\mathbf{H}$  for any  $a \in At(\pi(H))$ . This yields:

$$\mathbf{H} \approx \mathbf{H} / \uparrow_{\pi(\mathbf{H})} \mathbf{a} \times \mathbf{H} / \uparrow_{\pi(\mathbf{H})} \neg \mathbf{a} \quad (11)$$

Let  $n = |At(\pi(H))|$ . On applying (11)  $n - 1$  times we obtain (10).  $\square$

Let  $\mathbf{H}$  be a finite algebra in  $\mathcal{B}_{\mathbf{T}}$ .

**Definition 19** For any  $b \in \pi(H)$ ,  $b \neq 0$ , we shall denote by  $\mathbf{H}_b$  the algebra  $(H, \wedge, \vee, \rightarrow_b, (\pi_t)_{t \in T}, 0, b)$  where  $H_b = \{x \in H : x \leq b\}$ , and  $\rightarrow_b$  is the binary operation defined by:  $x \rightarrow_b y = (x \rightarrow y) \wedge b$ .

This Definition 19 is sound since it follows from axioms (p1) or (p2) that the operators  $\pi_t$  are isotone and consequently by axiom (p3) that for any  $b \in \pi(H)$ ,  $x \leq b$  implies  $\pi_t(x) \leq b$ .

**Lemma 20** For any  $a \in At(\pi(H))$ ,  $\mathbf{H}_a \in \mathcal{B}_{\mathbf{T}}$  and we have:

$$\mathbf{H} / \uparrow_{\pi(\mathbf{H})} \mathbf{a} \approx \mathbf{H}_a$$

**PROOF.** It is well known that the mapping  $r_a : x \mapsto x \wedge a$  is a Heyting homomorphism from  $(H, \wedge, \vee, \rightarrow, 0, 1)$  onto  $(H, \wedge, \vee, \rightarrow_b, 0, b)$  (32, Ch IV, §8). Moreover, it follows from axioms (p1) and (p3) that  $r_a(\pi_t(x)) = \pi_t(x) \wedge a = \pi_t(r_a(x))$  and consequently that  $r_a$  is a homomorphism of  $T$ -Rough algebras. Finally, since  $r_a^{-1}(b) = \uparrow_H a$ , we have  $\mathbf{H}_a \approx r_a(\mathbf{H}) \approx \mathbf{H} / \uparrow_{\pi(\mathbf{H})} \mathbf{a}$ .  $\square$

## 4.2 Finite free algebras

Given a finite poset  $\mathbf{T}$ , we investigate finite free  $T$ -algebras.

Let  $\mathbf{H}$  be an algebra and  $S$  a subset of its universe  $H$ . In the sequel we denote by  $\mathbf{H}(S)$  the subalgebra of  $\mathbf{H}$  generated by  $S$ . Let  $\mathbf{T} = (T, \leq)$  be an arbitrary finite poset. Since the variety  $\mathcal{B}_{\mathbf{T}}$  is generated by the finite algebra  $\underline{\mathbf{2}}_{\mathbf{T}}$ , it follows that any free algebra in  $\mathcal{B}_{\mathbf{T}}$  generated by a finite set is finite. For any integer  $n$ , we denote by  $\mathbf{F}_{\mathbf{T}}(n)$  the free algebra of  $\mathcal{B}_{\mathbf{T}}$  with  $n$  generators.

**Theorem 21** Let  $n$  be an integer,  $n \leq 1$ , let  $G$  be a set such that  $|G| = n$  and let  $\underline{\mathbf{2}}_{\mathbf{T}}^G$  be the set of mappings from  $G$  into  $\underline{\mathbf{2}}_{\mathbf{T}}$ . Then:

$$\mathbf{F}_{\mathbf{T}}(n) \approx \prod_{f \in \underline{\mathbf{2}}_{\mathbf{T}}^G} \underline{\mathbf{2}}_{\mathbf{T}}(f(G))$$

where  $\underline{\mathbf{2}}_{\mathbf{T}}(\mathbf{f}(\mathbf{G}))$  is the subalgebra of  $\underline{\mathbf{2}}_{\mathbf{T}}$  generated by the image of  $G$  under  $f$ , i.e.  $\{f(g) : g \in G\}$ .

**PROOF.** Let  $B = \pi(F_T(n))$ . It follows from (10) and Lemma 20 that

$$\mathbf{F}_{\mathbf{T}}(\mathbf{n}) \approx \prod_{a \in \text{At}(B)} \mathbf{F}_{\mathbf{T}}(\mathbf{n}) / \uparrow_{\mathbf{B}} \mathbf{a} \quad (12)$$

For any  $a \in \text{At}(B)$ , let  $h_a$  be the natural homomorphism associated with  $\theta_{\uparrow_{\mathbf{B}} a}$ . It follows from Corollary 16, that  $\mathbf{F}_{\mathbf{T}}(\mathbf{n}) / \uparrow_{\mathbf{B}} \mathbf{a}$  can be identified with a subalgebra of  $\underline{\mathbf{2}}_{\mathbf{T}}$  and  $h_a$  with a mapping from  $\mathbf{F}_{\mathbf{T}}(\mathbf{n})$  into  $\underline{\mathbf{2}}_{\mathbf{T}}$ . This enables us to write:

$$\mathbf{F}_{\mathbf{T}}(\mathbf{n}) / \uparrow_{\mathbf{B}} \mathbf{a} \approx \underline{\mathbf{2}}_{\mathbf{T}}(\mathbf{h}_a(\mathbf{G}))$$

Since for any  $a_1, a_2 \in \text{At}(B)$ ,  $a_1 \neq a_2$ , there exists  $g \in G$  such that  $h_{a_1}(g) \neq h_{a_2}(g)$ , this yields that the mapping  $a \in \text{At}(B) \mapsto h_a \in \underline{\mathbf{2}}_{\mathbf{T}}^G$  is injective and consequently there exists a subset  $I \subseteq \underline{\mathbf{2}}_{\mathbf{T}}^G$  such that:

$$\mathbf{F}_{\mathbf{T}}(\mathbf{n}) \approx \prod_{f \in I} \underline{\mathbf{2}}_{\mathbf{T}}(\mathbf{f}(\mathbf{G}))$$

We now show that  $I = \underline{\mathbf{2}}_{\mathbf{T}}^G$ . Assume that  $f \in \underline{\mathbf{2}}_{\mathbf{T}}(f(G))$ . Since  $F_T(n)$  is free, there exists a homomorphism  $\hat{f} : F_T(n) \rightarrow \underline{\mathbf{2}}_{\mathbf{T}}$  such that for any  $g \in G$ ,  $\hat{f}(g) = f(g)$ . Since  $\hat{f}(F_T(n))$  is the universe of a simple algebra, the kernel of  $\hat{f}$  is a maximal congruence of  $F_T(n)$ . Since  $F_T(n)$  is finite, it follows from Theorem 15 that  $\bigwedge \hat{f}^{-1}(1) \in \text{At}(B)$ . Let us denote by  $a_f$  this atom. It follows from the determination principle (Lemma 5) that the mapping  $f \mapsto a_f$  is injective. Indeed, let  $f, h \in \underline{\mathbf{2}}_{\mathbf{T}}^G$  and  $g \in G$  such that  $f(g) \neq h(g)$ . Then there exists  $t \in T$  such that  $\pi_t(f(g)) \neq \pi_t(h(g))$  and consequently  $\hat{f}(\pi_t(g)) \neq \hat{h}(\pi_t(g))$ . Since for any  $x \in \underline{\mathbf{2}}_{\mathbf{T}}$ ,  $\pi_t(x) \in \{0, 1\}$ , this yields that  $\hat{f}^{-1}(1) \neq \hat{h}^{-1}(1)$  and consequently  $a_f \neq a_h$ . This means that  $|I| \geq |\underline{\mathbf{2}}_{\mathbf{T}}^G|$  and we conclude  $I = \underline{\mathbf{2}}_{\mathbf{T}}^G$ .

□

Let us remark that if  $\mathbf{T}$  is the chain  $1 \leq 2$ , then  $\mathbf{F}_{\mathbf{T}}(\mathbf{n}) = \underline{\mathbf{2}}_{\mathbf{T}}^q \times \underline{\mathbf{2}}^p$  where  $p = 2^n$ ,  $q = 3^n - p$ . Thus we obtain:

$$|F_T(n)| = 3^{3^n - 2^n} \times 2^{2^n}$$

This is the formula proved by A. Monteiro for free 3-valued Łukasiewicz algebras with  $n$  generators (20).

## 5 Propositional calculus

We now show the connection of  $T$ -Rough algebras with the propositional calculus introduced and investigated by A. Nour in (23).

Let  $\mathbf{T}$  be an arbitrary finite poset. We denote by  $\mathcal{F}$  the set of function symbols  $\{\wedge, \vee, \rightarrow, (\pi_t)_{t \in T}, 0, 1\}$  where  $\wedge, \vee, \rightarrow$  are binary,  $\pi_t$  are unary for any  $t \in T$  and  $0, 1$  are nullary. Let  $X$  be a set of distinct objects called variables and let  $\mathbf{T}(\mathbf{X}) = (T(X), \mathcal{F})$  be the term algebra of type  $\mathcal{F}$  over  $X$  (3, ch. II). Let  $\theta_X$  be the congruence on  $\mathbf{T}(X)$  defined by

$$\theta_X = \bigcap \{ \theta \in \text{Cong}(\mathbf{T}(X)) : \mathbf{T}(X)/\theta \in \mathcal{B}_{\mathbf{T}} \}$$

Then the  $\mathcal{B}_{\mathbf{T}}$ -free algebra over  $X$  is the algebra  $\mathbf{T}(X)/\theta_X$ .

We denote by  $\mathcal{A}_I$  the set of terms that are axioms of the positive propositional calculus of Hilbert and Bernays (32, Ch IX, §1), and by  $\mathcal{A}_T$  the finite set of axioms defined by the following schemas (23):

- (a1)  $\pi_t(x \vee y) \leftrightarrow \pi_t(x) \vee \pi_t(y)$  ;
- (p2)  $\pi_t(x \wedge y) \leftrightarrow \pi_t(x) \wedge \pi_t(y)$  ;
- (p3)  $\pi_t \pi_u(x) \leftrightarrow \pi_u(x)$  ;
- (p4)  $\neg \pi_t(0)$  ;
- (p5)  $\pi_t(x) \vee \neg \pi_t(x)$  ;
- (p6)  $\pi_t(x \rightarrow y) \leftrightarrow \bigwedge_{v \geq t} (\pi_v(x) \rightarrow \pi_v(y))$  ;
- (p7)  $\bigwedge_{v \in T} \pi_v(x) \rightarrow x$ .

for any  $t, u, v \in T$ .

Let  $\mathcal{T}h$  be the smallest subset of  $T(X)$  containing  $\mathcal{A}_I \cup \mathcal{A}_T$  such that:

- (1) if  $\{\tau_1, \tau_1 \rightarrow \tau_2\} \subseteq \mathcal{T}h$  then  $\tau_2 \in \mathcal{T}h$ ;
- (2) if  $\tau_1 \in \mathcal{T}h$  then  $\pi_t(\tau_1) \in \mathcal{T}h$  for any  $t \in T$ .

The results of A. Nour (23, §5) yield that for any  $\tau \in T(X)$ ,  $\tau \in \mathcal{T}h$  iff the algebra  $\mathbf{2}_{\mathbf{T}}$  satisfies the identity  $\tau = 1$ .

Following Lindenbaum, we define an equivalence relation  $\theta_L(X)$  on  $T(X)$  by:

$$(\tau_1, \tau_2) \in \theta_L \iff \tau_1 \leftrightarrow \tau_2 \in \mathcal{T}h$$

The next corollary states that the Lindenbaum algebra derived from A. Nour's propositional calculus is a free  $T$ -Rough algebras.

**Corollary 22**  $\theta_L = \theta_X$ .

**PROOF.** We have the following equivalences:  $(\tau_1, \tau_2) \in \theta_X$  iff  $\tau_1 \leftrightarrow \tau_2 \in \theta_X(1)$  iff the algebra  $\mathbf{2}_T$  satisfies the identity  $\tau_1 \leftrightarrow \tau_2 = 1$  by Theorem 10 iff  $\tau_1 \leftrightarrow \tau_2 \in Th$  by A. Nour's completeness theorem in (23, Cor 2).  $\square$

## 6 $k$ -Rough Sets

We start by introducing our notations and basic definitions on Rough Sets. Following Pawlak (26), we shall call :

**information system**, a triple  $(Ob, Att, Val)$  where  $Ob$  and  $Val$  are sets and  $Att$  is a finite collection of functions that map  $Ob$  into  $Val$ . The elements of  $Ob$  are called objects and are often associated with the records of a relational database. The elements of  $Att$  are called attributes. They correspond to the fields of the database. Therefore  $Val$  is the set of values of the attributes.

**approximation space**, a frame  $(Ob, IND_A)$ , where  $A \subseteq Att$  and  $IND_A$  is an equivalence relation on  $Ob$  such that  $(x, y) \in IND_A$  if the attributes in  $A$  cannot distinguish between  $x$  and  $y$ . (i.e. For all  $(a_1, a_2) \in Att^2$  and all  $x \in Ob$  we have  $a_1(x) = a_2(y)$ ). If  $A$  is a singleton we simply write  $IND_a$  instead of  $IND_{\{a\}}$ .

In the sequel we suppose that the set of values  $Val$  contains a distinguished element  $NULL$  such that  $a(x) = NULL$  whenever the field corresponding to  $a$  in the database is not defined for the record associated with the object  $x$ . Moreover, for any subset of attributes  $A \in Att$  we denote by  $IND_A(x)$  the equivalence class of  $x : \{y \in Ob : (x, y) \in IND_A\}$ .

We shall say that a collection  $\mathcal{C}$  of (possibly empty) subsets of  $\wp(Ob)$  is a partition of  $Ob$  if for any  $X, Y \in \mathcal{C}$ ,  $X \cap Y = \emptyset$  and  $\bigcup \mathcal{C} = Ob$ . For each subset  $A$  of attributes, we denote by  $\mathcal{C}_A$  the partition associated with  $IND_A$  :  $\mathcal{C} = \{IND_A(x) : x \in Ob\}$ . Conversely, let  $\mathcal{C}$  be a partition of  $Ob$ , then for any  $x \in Ob$ , we denote by  $\mathcal{C}(x)$  the class  $X \in \mathcal{C}$  such that  $x \in X$ . In other words, we identify  $\mathcal{C}$  with the natural map from  $Ob$  into  $\mathcal{C}$ . Therefore we have for every  $a \in Att$ ,  $IND_a = IND_{\mathcal{C}_a}$ .

In the rest of this section, we consider that  $\mathcal{C}$  is a partition of  $Ob$ . Let  $\mathbf{B}$  be the Boolean algebra  $(\wp(Ob), \cap, \cup, -, \emptyset, Ob)$  made of subsets of  $Ob$  and let  $B(\mathcal{C})$  be the collection of set-union of equivalence classes:

$$B(\mathcal{C}) = \left\{ \bigcup \mathcal{X} : \mathcal{X} \subseteq \mathcal{C} \right\}$$

It is worth mentioning that the algebra:  $\mathbf{B}(\mathcal{C}) = (B(\mathcal{C}), \cap, \cup, -, \emptyset, Ob)$  is a boolean sub-algebra of  $\wp(Ob)$ .

Let  $k$  be an integer. Following Ziarko in (35), we consider several usual ways of approximating a subset  $X \subseteq Ob$  called the  $j$ -lower approximation for  $1 \leq j \leq k$  and the upper approximation for  $j = k + 1$ . They are defined respectively as:

$$f_j^{\mathcal{C}}(X) = \bigcup \left\{ \mathcal{C}(x) : |\mathcal{C}(x) \cap X| \geq \frac{|\mathcal{C}(x)|}{j} \right\} \quad (13)$$

$$f_{k+1}^{\mathcal{C}}(X) = \bigcup \{ \mathcal{C}(x) : \mathcal{C}(x) \cap X \neq \emptyset \} \quad (14)$$

Note that for any  $1 \leq j \leq k + 1$  and  $X \in \mathcal{C}$ , we have  $f_j^{\mathcal{C}}(X) = X$ . Since we have  $f_1^{\mathcal{C}}(X) = \bigcup \{ Y \in \mathcal{C} : Y \subseteq X \}$ ,  $f_1^{\mathcal{C}}$  is the lower approximation operator  $L$  considered by Pawlak.

According to Pawlak (26), a **Rough Set** of an approximation space  $(Ob, IND_{\mathcal{C}})$  is a pair  $(f_1^{\mathcal{C}}(X), f_{k+1}^{\mathcal{C}}(X))$  for some  $X \subseteq Ob$ .

**Example 23** Let  $Ob$  be the set  $\{a_1, b_1, b_2, c_1, c_2, c_3\}$  and let  $\mathcal{C} = \{C_a, C_b, C_c\}$  a partition of  $Ob$  in three classes where  $C_a = \{a_1\}$ ,  $C_b = \{b_1, b_2\}$  and  $C_c = \{c_1, c_2, c_3\}$ . Then the collection of rough sets of the approximation space  $(Ob, IND_{\mathcal{C}})$  is the collection of pairs  $(X, Y) \in \wp(Ob)^2$  such that  $X \subseteq Y$  and  $C_a \subseteq X$  whenever  $C_a \subseteq Y$ .

We are now going to define  $k$ -Rough Sets which are special cases of  $\alpha$ -Rough Sets introduced in (35) to generalize Pawlak's Rough Set concept.

**Definition 24** Let  $X$  be a subset of  $Ob$ , for any integer  $k > 0$  the  **$k$ -Rough Set** generated by  $X$  is the sequence:  $\vec{X} = (f_i^{\mathcal{C}}(X))_{1 \leq i \leq k+1}$ .

We denote by  $RS(\mathcal{C})$  the collection of all  $k$ -Rough Sets:  $RS(\mathcal{C}) = \{ \vec{X} : X \subseteq Ob \}$ .

Since for an equivalence class  $X$  the sequence  $(f_j(X))$  is constant, this yields a set-embedding of  $B(\mathcal{C})$  into  $k$ -rough sets. In the sequel we shall identify every element  $X$  of  $B(\mathcal{C})$  with the constant sequence  $\vec{X} = (X, \dots, X)$ .

Relationships between  $k$ -Rough representations can be modeled using HABO operations defined on the lattice  $RS(\mathcal{C})$ .

**Definition 25** We denote by  $\vec{0}$  the  $k + 1$ -sequence  $(\emptyset, \dots, \emptyset)$  and by  $\vec{1}$  the  $k + 1$ -sequence  $(Ob, \dots, Ob)$ . Let  $X, Y$  be two subsets of  $Ob$ ,

- (1) we define  $k + 1$  unary operations  $\pi_i$  on  $RS(\mathcal{C})$  by setting:  $\pi_i(\vec{X}) = \overrightarrow{f_i^{\mathcal{C}}(X)}$  for any  $1 \leq i \leq k + 1$ .
- (2) moreover we define three binary operations  $\wedge, \vee, \rightarrow$  by the following equalities:

$$\vec{X} \wedge \vec{Y} = \left( f_i^{\mathcal{C}}(X) \cap f_i^{\mathcal{C}}(Y) \right)_{1 \leq i \leq k+1} \quad (15)$$

$$\vec{X} \vee \vec{Y} = \left( f_i^{\mathcal{C}}(X) \cup f_i^{\mathcal{C}}(Y) \right)_{1 \leq i \leq k+1} \quad (16)$$

$$\vec{X} \rightarrow \vec{Y} = \left( \bigcap_{i \leq j \leq k+1} \left( -f_j^{\mathcal{C}}(X) \cup f_j^{\mathcal{C}}(Y) \right) \right)_{1 \leq i \leq k+1} \quad (17)$$

Let  $X, Y \subseteq Ob$ , we shall write  $\vec{X} \leq \vec{Y}$  whenever for any approximation operator  $f_i$  we have :  $f_i^{\mathcal{C}}(X) \subseteq f_i^{\mathcal{C}}(Y)$

Observe that for any  $C \in \mathcal{C}$ ,  $\vec{C} \leq \vec{X}$  iff  $C \subseteq f_i^{\mathcal{C}}(X)$  for any  $1 \leq i \leq k+1$ .

**Lemma 26** *RS(C) is closed under the operations  $\wedge, \vee, \rightarrow$  and  $\pi_i$  for every  $1 \leq i \leq k+1$ .*

**PROOF.** It is clear that  $RS(\mathcal{C})$  is closed under  $\pi_i$  since  $\pi_i(X) = \overrightarrow{f_i^{\mathcal{C}}(X)}$ . We shall prove that for any  $X, Y \subseteq Ob$  there exists  $Z_{op} \subseteq Ob$  such that  $\overrightarrow{Z_{op}} = \vec{X} \text{ op } \vec{Y}$  where  $\text{op} \in \{\wedge, \vee, \rightarrow\}$ .

For every  $C \in \mathcal{C}$  let:

$$C_{\wedge} = \begin{cases} X \cap C & \text{if } |X \cap C| \leq |Y \cap C| \\ Y \cap C & \text{otherwise} \end{cases} \quad (18)$$

$$C_{\vee} = \begin{cases} X \cap C & \text{if } |X \cap C| \geq |Y \cap C| \\ Y \cap C & \text{otherwise} \end{cases} \quad (19)$$

$$C_{\rightarrow} = \begin{cases} C & \text{if } \vec{C} \leq \vec{X} \rightarrow \vec{Y} \\ Y \cap C & \text{otherwise} \end{cases} \quad (20)$$

Then we can take :

$$Z_{op} = \bigcup \{C_{op} : C \in \mathcal{C}\} \quad (21)$$

for every  $\text{op} \in \{\wedge, \vee, \rightarrow\}$ .

This is obvious except for  $\text{op} = \rightarrow$ . In this case observe that :

$$C \subseteq \bigcap_{i \leq j \leq k+1} \left( -f_j^{\mathcal{C}}(X) \cup f_j^{\mathcal{C}}(Y) \right) \quad (22)$$

holds iff  $C \subseteq f_j^c(Y)$  whenever  $C \subseteq f_j^c(X)$  and  $j \geq i$ . Since  $C \not\subseteq f_i^c(X)$  yields  $C \not\subseteq f_n^c(X)$  for every  $1 \leq n \leq i$ , it follows that for all  $C \in \mathcal{C}$ , (22) holds for every  $1 \leq i \leq k+1$  or whenever  $C \subseteq f_i^c(Y)$ . This shows that if for some  $C \in \mathcal{C}$ ,  $\overrightarrow{C} \not\subseteq \overrightarrow{X} \rightarrow \overrightarrow{Y}$  then  $\overrightarrow{C} \wedge (\overrightarrow{X} \rightarrow \overrightarrow{Y}) = \overrightarrow{C} \wedge \overrightarrow{Y} = \overrightarrow{C} \cap \overrightarrow{Y} = \overrightarrow{C}$ , and proves (21) for  $op = \rightarrow$ .  $\square$

We denote by  $\mathbf{H}_C$  the system:  $(RS(\mathcal{C}), \wedge, \vee, \rightarrow, \pi_1, \dots, \pi_{k+1}, \overrightarrow{0}, \overrightarrow{1})$ .

The following corollary is straightforward.

**Corollary 27**  $\mathbf{H}_C$  is a  $k$ -Rough algebra.

## 7 Application to Information Retrieval Systems

Following (1), an Information Retrieval (IR) model is a quadruple  $(\mathbf{D}, \mathbf{Q}, \mathcal{F}, R)$  where:

- (1)  $\mathbf{D}$  is a set composed of logical views for the documents in the collection.
- (2)  $\mathbf{Q}$  is a set composed of queries.
- (3)  $\mathcal{F}$  is a framework for modeling documents representations, queries, and their relationships.
- (4)  $R$  is a binary ranking function which associates a real number with a query  $X_q \in \mathbf{Q}$  and a document's logical view  $X_d \in \mathbf{D}$ .

Let us suppose that documents are indexed by a set  $Ob$  of key-words and that there exists a classification  $\mathcal{C}$  of these key-words into topics.

We consider the new IR model induced by:

- (1) document logical views  $\mathbf{D}$  and queries  $\mathbf{Q}$  are subsets of  $Ob$ ;
- (2) the framework  $\mathcal{F}$  is the algebra  $\mathbf{H}_C$ .

The rest of the section is devoted to the definition in this algebraic framework of the ranking function  $R$ .

Let  $k$  be an integer and let  $K$  be the set of smaller integers  $\{0, 1, \dots, k\}$ . The following definition introduces the concept of fuzzy membership function ( $k$ -membership function) corresponding to  $k$ -Rough sets following (26).

**Definition 28** For any  $X \subseteq Ob$ , we define a  $k$ -membership function  $\lambda_X :$

$\mathcal{C} \longrightarrow K$  associated with  $\vec{X}$  in the following way for any  $C \in \mathcal{C}$ :

$$\lambda_X(C) = \begin{cases} 0 & \text{if } C \cap X = \emptyset \\ 1 / \min\{j \in K : C \subseteq f_j(X)\} & \text{otherwise} \end{cases}$$

We shall denote by  $\lambda_{\mathcal{C}}$  the set  $\{\lambda_X : X \subseteq Ob\}$ .

Obviously there is a one-to-one correspondence between  $RS(\mathcal{C})$  and  $\lambda_{\mathcal{C}}$  since by definition for any  $X \subseteq Ob$ ,  $C \in \mathcal{C}$  and  $j \in K$  we have:

$$\lambda_X(C) \geq 1/j \iff C \subseteq f_j(X)$$

$k$ -membership functions can be extended to the operations between  $k$ -Rough sets in the usual way for every  $X, Y \subseteq Ob$  and  $C \in \mathcal{C}$ :

$$(\lambda_X \wedge \lambda_Y)(C) = \min\{\lambda_X(C), \lambda_Y(C)\} \quad (23)$$

$$(\lambda_X \vee \lambda_Y)(C) = \max\{\lambda_X(C), \lambda_Y(C)\} \quad (24)$$

$$(\lambda_X \rightarrow \lambda_Y)(C) = \begin{cases} 1 & \text{if } \lambda_X(C) \leq \lambda_Y(C) \\ \lambda_Y(C) & \text{otherwise} \end{cases} \quad (25)$$

$$\pi_i(\lambda_X)(C) = \begin{cases} 1 & \text{if } C \subseteq f_i^c(X) \\ 0 & \text{otherwise} \end{cases} \quad (26)$$

We now suppose that  $Ob$  is a finite set. We define a measure on  $RS(\mathcal{C})$ , by analogy with finite measures on Boolean algebras and usual indices of dissimilarity for categorical variables (5).

**Definition 29** *A finite measure on  $RS(\mathcal{C})$  is an isotone non negative real valued unary function  $\mu$  such that for any  $X, Y \subseteq Ob$  we have  $\mu(\vec{X}) = 0$  whenever  $\vec{X} = \vec{0}$ ,  $\mu(\vec{X} \vee \vec{Y}) = \mu(\vec{X}) + \mu(\vec{Y})$  whenever  $\vec{X} \wedge \vec{Y} = \vec{0}$ . Moreover, we shall say that  $\mu$  is normalized if  $\mu(\vec{1}) = 1$  and it is positive if  $\vec{1}$  is the only element at which  $\mu$  takes the value 1.*

**Definition 30** *For any set  $U$ , we say that a non negative real valued binary function  $d$  is a dissimilarity on  $U$  if for any  $V_1, V_2, V_3 \subseteq U$ ,  $d(V_1, V_1) = 0$  and  $d(V_1, V_2) = d(V_2, V_1)$ .*

We remind the reader that if  $m$  is a finite normalized measure on a Boolean algebra  $\mathbf{B}$ , then the binary function defined by  $x, y \mapsto 1 - m(x \leftrightarrow y)$  is a

metric on  $\mathbf{B}$ .

We shall now introduce a positive measure  $\mu_{\mathcal{C}}$  and a dissimilarity  $\delta_{\mathcal{C}}$  on  $RS(\mathcal{C})$ .

We denote by  $\vec{X} \leftrightarrow \vec{Y}$  the  $k$ -rough set:  $(\vec{X} \rightarrow \vec{Y}) \wedge (\vec{Y} \rightarrow \vec{X})$ .

**Definition 31** We denote by  $\mu_{\mathcal{C}}$  the application from  $RS(\mathcal{C})$  into the set of real numbers defined for any  $X \subseteq Ob$  by:

$$\mu_{\mathcal{C}}(\vec{X}) = \frac{\max\{ |Y| : \vec{Y} = \vec{X}, Y \subseteq Ob \}}{|Ob|}$$

The proof of the following lemma is straightforward.

**Lemma 32**  $\mu_{\mathcal{C}}$  is a normalized positive finite measure on  $RS(\mathcal{C})$  and the real binary function  $\delta_{\mathcal{C}}$  defined for every  $X, Y \subseteq Ob$  by  $\delta_{\mathcal{C}}(\vec{X}, \vec{Y}) = 1 - \mu_{\mathcal{C}}(\vec{X} \leftrightarrow \vec{Y})$  is a dissimilarity.

We shall now show how to compute the number  $\mu_{\mathcal{C}}(\vec{X})$  for any  $X \subseteq Ob$ .

**Lemma 33** Let  $S$  be a subset of  $Ob$  and  $j$  an integer such that  $j < |S|$ . We denote by  $\lfloor S \rfloor_j$  the integer  $n$  such that:

- $n = |S|$  if  $j = 0$ ,
- $n = \frac{|S|}{j} - 1$  if  $j \neq 0$  and  $|S|$  is a multiple of  $j$ ,
- $\frac{|S|}{j} - 1 \leq n \leq \frac{|S|}{j}$  otherwise

Then we have for any  $X \subseteq Ob$ :

$$\mu_{\mathcal{C}}(\vec{X}) = \frac{\sum \{ \lfloor C \rfloor_{1/\lambda_X(C)-1} : C \in \mathcal{C}, C \subseteq f_{k+1}^{\mathcal{C}}(X) \}}{|Ob|}$$

Hence, we can define the ranking function  $R$  by setting:

$$R(X_q, X_d) = |Ob| \cdot \mu_{\mathcal{C}}(\vec{X}_q \rightarrow \vec{X}_d) + \mu_{\mathcal{C}}(\vec{X}_d \rightarrow \vec{X}_q)$$

for any  $X_q \in \mathbf{Q}$  and  $X_d \in \mathbf{D}$ .

The value for  $\mu_{\mathcal{C}}(\vec{X}_q \rightarrow \vec{X}_d)$  (i) is maximal if the relevant topics with regard to query  $q$  are also relevant for document  $d$ . Thus the function  $R$  ranks these documents at the top of the list. Conversely, the value of  $\mu_{\mathcal{C}}(\vec{X}_d \rightarrow \vec{X}_q)$  (ii) is maximal if all the topics that characterize document  $d$  are also relevant for

query  $q$ . Thus two documents,  $d_1, d_2$  will have the same value for (i), the  $R$  function will use the value of (ii) to differentiate them.

This shows that HABO allow to formalize implicative relationships in the sense of (9), between documents and queries as logical implications to which a measure of uncertainty is associated such as in Rijsbergen (27) and Sebastiani (29), but not in terms of conditional probability. Given two documents whose terms overlap the same topics in a more or less similar manner, they will be associated irrespective of whether or not they share common terms. Like in (11) or in fuzzy IR (1, §2.6), our approach is basically an indirect way of ranking documents as relevant documents, a promising compromise between Boolean and Fuzzy IR.

## 8 Examples

We illustrate the basic elements of the algebraic structure of  $k$ -rough sets by means of a small hierarchical classification system on computing methodologies specified in (12) with two levels (a unique class  $L_1(t)$  and supra-class  $L_2(t)$  is given to each term  $t \in Ob$ ). Table 1 gives the corresponding information system  $HS = (Ob, Att, Val)$  where  $Ob$  is a set of 15 terms and  $Att = \{L_1, L_2\}$ .

$HS$  hierarchical means that we have the functional dependency :  $L_1 \rightarrow L_2$ , thus  $IND_{Att} = IND_{\{L_1, L_2\}} = IND_{\{L_1\}}$  ( $L_1$  is a key of the database). However, this is not compulsory and what follows remains correct for any Information System by taking  $\mathcal{C} = \mathcal{C}_A$  where  $A$  can be the whole set  $Att$  of attributes, or any key of the database.

Then we have :  $\mathcal{C}_{L_1} = \{A, B, C, D, E\}$  and  $\mathcal{C}_{L_2} = \{A \cup B, C \cup D \cup E\}$  where  $A = \{a_1, a_2\}$ ,  $B = \{b_1, b_2, b_3\}$ ,  $C = \{c_1, \dots, c_6\}$ ,  $D = \{d_1, d_2\}$  and  $E = \{e_1, e_2\}$  are five disjoint clusters of terms in table 1.

Table 2 shows three examples of Library entries in (12). Braced lists are associated terms, and parentheses enclose summaries of the books.

It follows that in this classification system, each document  $d_i$  has associated with it a set of index terms  $t_1, \dots, t_n$  that capture the essence of it. Every document then becomes synonymous with its set of terms. The terms, in turn, are divided into clusters  $A, \dots, E$  according to the topics to which they refer.

In keeping with the format of the  $k$ -Rough model described so far, we merge the documents into the dependence spaces induced by the partitions  $\mathcal{C}_{L_1}$  and  $\mathcal{C}_{L_2}$  of the set  $Ob$  of indexed terms. Then every set of terms associated with a document defines a 2-Rough Set in each dependence space as described in

Table 1  
Example of hierarchical classification system

index term	class	supra-class
$a_1$ : Algebraic algorithms	Algorithms	Algebraic Manipulations
$a_2$ : Analysis of algorithms	Algorithms	Algebraic Manipulations
$b_1$ : Evaluation strategies	Languages and Systems	Algebraic Manipulations
$b_2$ : Nonprocedural languages	Languages and Systems	Algebraic Manipulations
$b_3$ : Special purpose hardware	Languages and Systems	Algebraic Manipulations
$c_1$ : Cartography	Applications	Artificial Intelligence
$c_2$ : Games	Applications	Artificial Intelligence
$c_3$ : Industrial Automation	Applications	Artificial Intelligence
$c_4$ : Law	Applications	Artificial Intelligence
$c_5$ : Medecine and Science	Applications	Artificial Intelligence
$c_6$ : Office automation	Applications	Artificial Intelligence
$d_1$ : Analogies	Learning	Artificial Intelligence
$d_2$ : Concept learning	Learning	Artificial Intelligence
$e_1$ : Manipulators	Robotics	Artificial Intelligence
$e_2$ : Sensors	Robotics	Artificial Intelligence

Table 3.

These three documents share few terms in common, meanwhile  $d_1$  and  $d_6$  have the same 2-Rough approximation in  $(Ob, IND_{L_2})$ . We are now going to compute the implications between these three documents in  $(Ob, IND_{L_1})$ . This can easily be done using membership functions in Table 4.

Table 5 illustrates how the Heyting implication between two documents  $d_1$  and  $d_2$  can be used in an query expansion process. Indeed, for every  $1 \leq i \leq k$ , the following formula based on the Heyting implication and on the Boolean operators gives the topics (classes) that are  $1/i$ -relevant to document  $d_1$  and that are developed in document  $d_2$  :

$$\pi_i(\vec{d}_1) \wedge \pi_1(\vec{d}_1 \rightarrow \vec{d}_2)$$

It follows that the measure  $\mu(\vec{d}_1 \rightarrow \vec{d}_2)$  should point out a document  $d_2$  which could be interesting for any reader of  $d_1$ . The values of this measure for the three documents in the previous example are given in Table 6. By way of

Table 2

Example of Library entries

- $d_1$  “Artificial Justice: Expert systems and Legal Consultation”. Fergus, W. W.  
 {Law, Evaluation strategies, Special purpose hardware}  
 (An evaluation of expert systems applied to the field of legal consultation).
- $d_3$  “Computer system review”. Aron and Cornick.  
 {Algebraic algorithms, Analysis of algorithms, Nonprocedural languages,  
 Cartography, Games, Office automation, Concept learning, Manipulators, Sensors}  
 (An encyclopedic review of computer systems technology).
- $d_6$  “Programming Heuristics and Heuristic Programming”. Enver, N.  
 {Nonprocedural languages , Special purpose hardware, Analogies, Concept learning}  
 (An exploration of the programming of intelligent systems with heuristic learning).

Table 3

2-Rough Sets generated by documents

$X \subseteq Ob$	$\vec{X}$ in $(Ob, IND_{L_1})$	$\mu_{C_{L_1}}(\vec{X})$	$\vec{X}$ in $(Ob, IND_{L_2})$	$\mu_{C_{L_2}}(\vec{X})$
$d_1$	$(\emptyset, B, B \cup C)$	$\frac{4}{15}$	$(\emptyset, \emptyset, Ob)$	$\frac{6}{15}$
$d_3$	$(A \cup E, Ob - B, Ob)$	$\frac{11}{15}$	$(\emptyset, Ob, Ob)$	$\frac{13}{15}$
$d_4$	$(D, B \cup D, B \cup D)$	$\frac{4}{15}$	$(\emptyset, \emptyset, Ob)$	$\frac{6}{15}$

Table 4

Membership functions associated with documents in  $(Ob, IND_{L_1})$ 

$X$	$\lambda_X(A)$	$\lambda_X(B)$	$\lambda_X(C)$	$\lambda_X(D)$	$\lambda_X(E)$
$d_1$	0	1/2	1/3	0	0
$d_3$	1	1/3	1/2	1/2	1
$d_6$	0	1/2	0	1	0

example, this table shows that a reader of a technical book on programming intelligent systems ( $d_6$ ) should be interested by the application of expert systems to legal consultation ( $d_1$ ), more than by the general encyclopedic review ( $d_3$ ). Note that usual statistical symmetrical dissimilarity measures (5, p. 85) would have ranked document  $d_3$  before  $d_1$ .

Yet, another strategy for query expansion is to consider similar documents. Table 6 gives the values of  $|Ob| \times \mu_C(\vec{d}_i \leftrightarrow \vec{d}_j)$ , for every  $i, j \in \{1, 3, 6\}$ . They induce by definition the values of the dissimilarity  $\delta_C$  between two documents

Table 5  
Implications between documents

$\vec{X} \rightarrow \vec{Y}$	$d_1$	$d_3$	$d_6$
$d_1$	$\vec{1}$	$(Ob - B, Ob - B, Ob)$	$(A \cup D \cup E, Ob - C, Ob - C)$
$d_3$	$(B, B, B \cup C)$	$\vec{1}$	$\overline{B \cup D}$
$d_6$	$\overline{Ob - D}$	$(A \cup C \cup E, Ob - B, Ob)$	$\vec{1}$

$d_i$  and  $d_j$  which is small whenever the two documents cover the same topics with the same intensity. The table shows that document  $d_1$  is closest to  $d_6$  than to  $d_3$ . This is because, like in Kulczynski and Jaccard statistics (5, p. 85), the computation of  $\vec{d}_i \leftrightarrow \vec{d}_j$  takes into account the number of disagreements between two categorical variables.

Table 6  
Implication measures and dissimilarities between documents

$ Ob  \times \mu_C(\vec{Y} \rightarrow \vec{X})$				$ Ob  \times \mu_C(\vec{X} \leftrightarrow \vec{Y})$			
$X \setminus Y$	$d_1$	$d_3$	$d_6$	$X \setminus Y$	$d_1$	$d_3$	$d_6$
$d_1$	15	13	8	$d_1$	15	3	6
$d_3$	5	15	5	$d_3$	3	15	3
$d_6$	13	12	15	$d_6$	6	3	15

## 9 Related and future works

$T$ -Rough algebras developed in this paper are special cases of finitely generated varieties of Heyting algebras with operators that have been intensively investigated in universal algebra theory. Suitable and optimal representations have been developed for Monadic Heyting Algebras (2), logics based on Łukasiewicz-Moisil algebras (16) in Kripke-style and relative-Stones Heyting algebras generated by finite chains (6) which have all many similarities with the  $k$ -Rough algebras introduced in this paper. In a more general way, a Priestley style duality theory has been developed for varieties of Heyting algebras (7) and duality for bounded lattices with operators has been studied and applied to the semantics of non classical logics in (31).

However, in this paper, we based the study of finite many valued algebraic logics on the simple model of Heyting algebras of finite increasing partially ordered sequences of subsets, following Moisil and Rasiowa tradition. We have

generalized to  $T$ -Rough algebras the fundamental representation theorem already established in (10) for plain semi post algebras and for Symmetrical Heyting Algebras with a finite order type of Operators in (15). We deduce from this theorem that, given a poset  $\mathbf{T}$ , the variety of  $T$ -Rough algebras is generated by a finite algebra.

To illustrate the potential applications of this algebraic frame, we have introduced an IR model that extends the clean and simple boolean model with the functionality of partial matching. The basic idea, like in fuzzy IR, is to expand the set of index terms in a query with related terms in the same topic. Hence the model deals with the representation of classes whose boundaries are not well defined like in fuzzy set theory. Meanwhile, the simplicity of our system allows us to consider large collections of documents. Moreover the sound algebraic formalism based on Heyting algebras with Boolean operators can incorporate new facts (close sentences) and rules (entailments) as they become available and suggest a way to progress in the direction of a more sophisticated question-answering system.

## References

- [1] R. BAEZA-YATES, B. RIBEIRO - NETO, *Modern Information Retrieval*, ACM Press, Addison-Wesley, 1999, 510 p.
- [2] G. BEZHANISHVILI, *Varieties of Monadic Heyting Algebras*, *Studia Logica*, Part I **61** (1998) pp. 367–402, Part II **62** (1999) pp. 21–48, Part III **64** (2000) pp. 215–256.
- [3] S. BURRIS, H.P. SANKAPPANAVAR, *A Course in Universal Algebra*, Graduate Texts in Mathematics, Springer-Verlag, 1981.
- [4] C. CARPINETO, G. ROMANO, *Order-Theoretical Ranking*, *Journal of the American Society for Information Science*, **51 - 7** (2000) pp. 587–601.
- [5] B.V. CUTSEM (ed.), *Classification and Dissimilarity Analysis*, *Lectures Notes in Statistics* **93**, 1994, 238 p.
- [6] B. A. DAVEY, *Dualities for equational classes of Brouwerian and Heyting algebras*, *Trans. Amer. Math. Soc.* **221** (1976), pp. 119-146.
- [7] B. A. DAVEY, H. A. PIESTLEY *Optimal Natural Dualities for varieties of Heyting algebras*, *Studia Logica* **56** (1996), pp. 67-96.
- [8] I. DÜNTSCH, *A logic for rough sets*, *Theoretical Computer Science*, **179 1-2**(1997) pp. 427-436.
- [9] R. GRAS, *L'implications statistique - Nouvelle méthode exploratoire de données*, Le pensée sauvage éditions, 1996.
- [10] Ng. CAT HO, H. RASIOWA, *Plain semi-Post algebras as a poset-based generalization of Post algebras and their representability*, *Studia Logica* **48** (1989), pp. 509–530.
- [11] S. DEERWESTER, S. T. DUMAIS, G. W. FURNAS, T. K. LANDAUER,

- R. HARSHMANN, *Indexing by Latent Semantic Analysis*, Journal of the American Society for Information Science **41 - 6** (1990) pp. 391–407.
- [12] J. GREEN, N. HORNE, E. ORLOWSKA, P. SIEMENS, *A Rough Set Model Information Retrieval*, Fundamenta Informaticae **28** (1996) pp. 273–296.
- [13] L. ITURRIOZ, *Many-valued logics for computer science applications, COST Action 15, Final Report*, Office for Official Publications of the European Communities, EUR 19204, 2000.
- [14] L. ITURRIOZ, *Rough Sets and three-valued structures*, Chapter 33 in Logic at Work, Essays Dedicated to the Memory of Helena Rasiowa, E. Orłowska (ed.), Physica-Verlag, Heilderberg (1999) pp. 596–603.
- [15] L. ITURRIOZ, *Symmetrical Heyting algebras with a finite order type of operators*, Studia Logica **55** (1995), pp. 89–98.
- [16] L. ITURRIOZ, E. ORLOWSKA, *A Kripke-style and relationnal semantics for logics based on Lukasiewicz algebras*, Conference in honour of J. Lukasiewicz, Dublin, 1996.
- [17] L. ITURRIOZ, E. ORLOWSKA, E. TURUNEN (eds), *Atlas of many-valued logic structures*, Tampere University of Technology, Mathematics Report **75**, 2000.
- [18] Gr. C. MOISIL, *Essais sur les logiques non chrysippiennes*, Editions de l'académie de la république socialiste de Roumanie, Bucarest 1972, 820 p.
- [19] A. MONTEIRO, *Sur les algèbres de Heyting symétriques*, Portugaliae Mathematica **39** (1980), pp. 1–237.
- [20] A. MONTEIRO, *Séminaire sur les algèbres de Lukasiewicz trivalentes*, July 1966, Universidad Nacional del Sur, Bahía Blanca, Argentina.
- [21] A. MONTEIRO, *Axiomes indépendants pour les algèbres de Brouwer*, Revista de la Unión Matemática Argentina y de la Asociación Física Argentina, vol. **XVII** de Homenaje a Beppo Levi, Buenos Aires (1953) pp. 149–160.
- [22] Gh. S. NADIU, *Sur la logique de Heyting*, in Logique Automatique Informatique, Académie de la R. S. de Roumanie, Bucarest 1971, pp. 41–70.
- [23] A. NOUR, *Sémantique algébrique d'un système logique basé sur un ensemble ordonné fini*, Mathematical Logic Quartely, **45 (4)** (1999), pp. 457–466.
- [24] E. ORLOWSKA, *What You Always Wanted to Know about Rough Sets*, Chapter 1 in Incomplete Information : Rough Set Analysis, Studies in Fuzziness and Soft Computing, E. Orłowska (ed.), Physica-Verlag, Heidelberg, 1998, pp.1–20.
- [25] P. PAGLIANI, *Rough-Sets Theory and logic algebraic structures in Modeling incomplete information - fundamentals and application*, Ewa Orłowska (ed.), springer, 1997.
- [26] Z. PAWLAK, *Rough Sets, Theoretical Aspects of Reasoning about Data*, Kluwer, Dordrecht, 1991.
- [27] C. J. RIJSBERGEN, *A non classical logic for information retrieval*, The computer journal, **29 (6)** (1986), pp. 481–485.

- [28] E. SANJUAN, *Algèbres de Heyting avec opérateurs Booléens et applications aux systèmes d'information*, PhD thesis, Claude Bernard university, Lyon, France, 2000, 204 p.
- [29] F. SEBASTIANI, *A Probabilistic Terminological Logic for Modelling Information Retrieval*, in Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, W.B. Croft and C.J. Rijsbergen sc. eds., Springer-Verlag, Dublin, Ireland, (1994) pp. 122–130.
- [30] M. SERFATI, *The lattice theory of r-ordered partitions* Discrete Mathematics **194 1-3** (1999) pp.205–227.
- [31] V. SOFRONIE-STOKKERMANS, *Duality and canonical extensions of bounded distributive lattices with operators, and applications to the semantics of non classical logics*, Studia Logica **64** (2000) pp. 93–132, 151 – 172.
- [32] H. RASIOWA, R. SIKORSKI, *The Mathematics of Metamathematics*, Polish Scientific Publ., 1963, 3rd ed., 1970.
- [33] H. RASIOWA, *On approximation logics: A survey*, Jahrbuch 1990, Kurt Gödel Gessellschaft, Vienna 1990, pp. 63–87.
- [34] W. ZIARKO, *Review of basics of rough sets in the context of data mining*, In Proceedings of the Fourth International Workshop on Rough Sets, Fuzzy Sets and Machine Discovery, S.Tsumoto (ed.), Tokyo Nov. 6-8 1996, pp. 447 – 457.
- [35] W. ZIARKO, *Variable precision rough set model*, Journal of computer and System Science **46** (1993) pp. 39 – 59.