Abstract Concept Lattices

Henry Soldano[1] and Véronique Ventos[2]

[1] L.I.P.N, UMR-CNRS 7030, Université Paris-Nord,
93430 Villetaneuse, France
[2] LRI, UMR-CNRS 8623, Université Paris-Sud,
91405 Orsay, France
`henry.soldano@lipn.univ-paris13.fr, ventos@lri.fr`

Abstract. We present a view of abstraction based on a structure preserving reduction of the Galois connection between a language $\mathcal{L}$ of terms and the powerset of a set of instances O. Such a relation is materialized as an *extension-intension lattice*, namely a concept lattice when L is the powerset of a set P of attributes. We define and characterize an *abstraction A* as some part of either the language or the powerset of O, defined in such a way that the extension-intension latticial structure is preserved. Such a structure is denoted for short as an *abstract lattice*. We discuss the extensional abstract lattices obtained by so reducing the powerset of O, together together with the corresponding abstract implications, and discuss *alpha lattices* as particular abstract lattices. Finally we give formal framework allowing to define a *generalized abstract lattice* whose language is made of terms mixing abstract and non abstract conjunctions of properties.

1 Introduction

There were in machine learning various attempts to formalize abstraction and characterize its desirable properties with respect to induction. An important statement was that abstraction should be order-preserving with respect to the partially ordered language in which hypotheses are searched for [21]. However, the question of what classes of abstractions are to be investigated for learning and reasoning is far from being exhausted. We present a view of abstraction based on a structure preserving reduction of the relation between a term t of a language $\mathcal{L}$, partially ordered following a general-to-specific partial order, and the *extension* of t on a set of objects (or instances) O, representing the subset of O whose elements satisfy the term.

In Formal Concept Analysis [15] and Galois lattice theory [4] $\mathcal{L}$ is a lattice, and the relation between $\mathcal{L}$ and the powerset $\mathcal{P}(O)$ is materialized as an *extension-intension lattice*. This lattice is the structure of the *definable* elements of $\mathcal{P}(O)$ [1], *i.e.* the subsets of O that are each the extension of some term of $\mathcal{L}$. In such a lattice, a definable set e represents the equivalence class of all the terms whose extension is e, and a node, also called a *concept*, is a pair (e, t) where t is the most specific term of the class, denoted as the *intension* of the concept. Formal Concept Analysis is primarily concerned with the relation between the powerset of a set of properties as a language and $\mathcal{P}(O)$. The extension-intension lattice is then denoted as a *concept lattice*. Various extensions have been recently proposed to ease the representation in more sophisticated languages

[6,18]. In particular *pattern structures* [16] have been recently introduced to represent complex data, associating such a *pattern structure* to each object. Logical Concept Analysis (LCA) [13] has been recently introduced as a general formalization in which $\mathcal{L}$ is a logical language and uses *object descriptions* in $\mathcal{L}$. Though we do not use here the LCA formulation and notations, for technical reasons, our construction of a Galois connection on a modal language is very similar to the construction presented in [12]. In a recent paper [20] particular mappings, denoted as *projections*, are used to reduce $\mathcal{L}$ or $\mathcal{P}(O)$ in such a way that the relation between the language and the extensional space is still materialized as a lattice. In other words, projections ensure that we have a coarser, yet structure preserving, view of concepts representatives of the universe we deal with. Independently [16] also uses projections on pattern structures.

In this paper we show that projections of a lattice are in a one to one correspondence with *abstractions* defined as parts of the lattice that are closed under least upper bound. We call the corresponding structures *abstract extension-intension lattices* and *abstract concept lattice*, and for short *abstract lattices*. We first briefly discuss intensional abstractions and investigate then more specifically extensional abstractions, i.e. parts of $\mathcal{P}(O)$ that are closed under set theoretic union.

More precisely, applying an extensional abstraction means that we will no longer consider instances of O as elements of the extensional space, but rather consider subsets of O given *a priori*. As an example consider $O = \{o_1, o_2, o_3, o_4\}$, and A be obtained by closing under union the part $\{\{o_1, o_2\}, \{o_1, o_3\}\}$. As a result, $\{o_1, o_2, o_3\}$ belongs to A but $\{o_2, o_3\}$ does not. Now, consider the smallest elements of A that contain a given instance o. We call these elements the *minimal abstractions* of o and consider them as *abstract instances*. In this example, $\{o_1, o_2\}$ is the unique minimal abstraction of o_2, $\{o_1, o_3\}$ is the unique minimal abstraction of o_3, $\{o_1, o_2\}$ and $\{o_1, o_3\}$ are the two minimal abstractions of o_1, and o_4 has no abstraction in A.

We relate then any term t to an *abstract extension* $ext_A(t)$ that turns out to be the union of all the abstract instances included in $ext(t)$. Going back to our example, suppose that $ext(t) = \{o_1, o_2, o_3, o_4\}$, then the abstract instances included in $ext(t)$ are $\{o_1, o_2\}$ and $\{o_1, o_3\}$ and therefore $ext_A(t) = \{o_1, o_2, o_3\}$. Clearly the abstract extension of a given term is always included in its original extension. In other words we reduce the extension of the term by excluding any instance that has no minimal abstraction included in the original extension. In the current example, o_4 is such an instance. The intuition here is a change in granularity: the new objects we deal with are the minimal abstractions of the original instances. Furthermore, as an extension-intension lattice represents a set of valid implications (see for instance [4]), our *abstract extension-intension lattice* also represents a set of *valid abstract implications*. Going back to our example, suppose $ext(p) = \{o_1, o_2, o_4\}$ and $ext(q) = \{o_1, o_2, o_3\}$, then the implication $p \rightarrow q$ is not valid on O. However, as we have $ext_A(p) = \{o_1, o_2\}$ and $ext_A(q) = \{o_1, o_2, o_3\}$, the abstract implication $p \rightarrow_A q$ is valid. As a matter of fact a general property of extensional abstractions is that they preserve validity of implications. Algorithms that extract valid implications [14,17] from a set of instances O, can be extended to extract *valid abstract implications*. As an example, we can interpret *alpha lattices* [28] as particular abstract lattices, and alpha implications are straightforwardly extracted by extending the method of N. Pasquier and collaborators [19] to alpha lattices.

Regarding extensional abstractions, the formal work closest to ours regards rough sets logics. Rough sets originally relies on an indiscernibility relation (see for instance [29]). More recently, generalization of rough sets using coverings, corresponding to our extensional abstractions, has been investigated from an algebraic perspective [5,8]. However these works do not investigate the extension-intension relationship.

To summarize, in this first part of the paper we characterize and discuss the properties of abstractions, and particularly of extensional abstractions, as a structure and order preserving reduction of extension-intension lattices.

We then remark that abstract implications are in fact implications between abstract terms, *i.e.* $t_1 \rightarrow t_2$ rewrites as $\Box t_1 \rightarrow \Box t_2$, where $\Box t$ means ”*Abstractly t*“ with respect to the abstraction A. Then, starting from the semantics, we discuss the corresponding *abstract modal logics*. This leads us to consider a language $\mathcal{L}\Box$ a term of which is made of a non abstract part together with an abstract part and to define a new, more expressive, extension-intension lattice.

2   Preliminaries

2.1   Galois Connection

We recall here the definitions of a Galois connection and a Galois lattice.

Definition 1. *Given two lattices $(E, \leq_E, \wedge, \vee)$ and $(F, \leq_F, \wedge, \vee)$, where $\leq_E$ and $\leq_F$ denote the order relations, and $\wedge, \vee$ the meet and the join operations, a Galois connection between E and F is a pair of mappings $(f : E \rightarrow F, g : F \rightarrow E)$ verifying the following properties:*

- *For any x and x' in E, we have that $x \leq_E x'$ implies $f(x) \geq_F f(x')$*
- *For any y and y' in F , we have that $y \leq_F y'$ implies $g(y) \geq_E g(y')$*
- *For any x in E and y in F we have that $g \circ f(x) \geq_E x$ and $f \circ g(y) \geq_F y$*

The Galois lattice defined by the Galois connection (f, g) is then the set $\{(x, y) \in E \times F \mid y = f(x) \text{ and } x = g(y)\}$ ordered by $\leq_E$.

2.2   Extension-Intension Lattices

FCA, in a broad sense, investigates the link between the terms of some language $\mathcal{L}$ and a universe O of elements of the universe denoted as *instances*. We also call *intensional* representations the terms of $\mathcal{L}$ and extensional representations the elements of the powerset $\mathcal{P}(O)$. In this presentation we suppose that we know whenever an instance o satisfies the term t, and we define accordingly its extension:

Definition 2. $ext_O(t) = \{o \in O \mid o \text{ satisfies } t\}.$

In what follows we consider O as a fixed and finite set of instances and we simply write $ext(t)$ the extension in O.

The language $\mathcal{L}$ is partially ordered by a general-to-specific relation. We write $t1 \preceq t2$ whenever t_1 is *less specific* than t_2, or equivalently t_1 is more general than t_2. The following proposition relate $\mathcal{L}$ and $\mathcal{P}(O)$ by a Galois connection.

Proposition 1. *Let $\mathcal{L}$ be a finite language, $\preceq$ a partial order on $\mathcal{L}$ denoted as speci-ficity, O be a finite set of instances and ext : $\mathcal{L} \to \mathcal{P}(O)$ be a mapping such that $t_1 \preceq t_2 \Rightarrow ext(t_1) \supseteq ext(t_2)$. We consider the following conditions:*

- *(condition 2.1) For any instance o, there is a* unique most specific term, *denoted as the object description $d(o)$, among all terms t such that $o \in ext(t)$*
- *(condition 2.2) $\mathcal{L}$ has a greatest element and is a lower semi lattice, i.e. each pair of terms t_1, t_2 of $\mathcal{L}$ has a unique greatest lower bound $t_1 \wedge_{\mathcal{L}} t_2$ in $\mathcal{L}$, also called the* least general generalisation *(for short lgg) of t_1 and t_2.*

Whenever conditions 2.1 and 2.2 are satisfied, $(\mathcal{L}, \preceq)$ is a lattice, i.e. two terms t_1 and t_2 also have a least upper bound denoted $t_1 \vee_{\mathcal{L}} t_2$, and the pair (int, ext) where

$$int(e) = \bigwedge_{o \in e_{\mathcal{L}}} d(o)$$

is a Galois connection.

In Formal Concept Analysis [15] and Galois Analysis [4], $\mathcal{L}$ is the powerset of a set P of attributes. Recent extensions to various languages have been performed [14,13]. In particular in [16], $\mathcal{L}$ is defined as a set of *pattern structures, i.e.* terms generated by first considering the set of object descriptions, and then closing it under the least general generalization $\wedge_{\mathcal{L}}$. Independently, E. Diday and R. Emilion start from the same assumptions [9]. Proposition 1 directly follows from, for instance, theorem 2 in [9].

We call *extension-intension lattice* the Galois lattice G coresponding to this Galois connection. G is ordered using the extensional order and each element G is a pair (e, t) such that $t = int(e)$ and $e = ext(t)$, *i.e.* t is the unique most specific term representing the equivalence class of terms whose extension is e, and is also denoted as the *intension* of e. Such most specific terms are also referred as *closed terms* or *closed motifs* in data mining [25]. In Machine Learning, the search space $\mathcal{L}$ may be in this way explored by minimally generalizing or specializing such closed terms [3].

Finally G is considered as the structure of O as perceived through $\mathcal{L}$ and can be also represented by the set T_O of all the implications $p \to q$ which are valid on O, i.e. such that $ext(p) \subseteq ext(q)$ is true. The equivalence class whose *intension* is $t = int(e)$, also contains various minimal (i.e. most general) terms t^m, also know as *generators*. The elements of T_O can be generated from the set of all the $t^m \to t$ implications also called the *min-max implications* basis of T_O [19].

2.3   Projections and Projected Extension-Intension Lattices

Now, consider the following problem; how to reduce $\mathcal{P}(O)$, $\mathcal{L}$, or both, in such a way that *i)* conditions 2.1 and 2.2 are still satisfied, *ii)* the resulting extension-intension lattice G' is isomorphic to part of G. In other words, how to reduce G by reducing the language or the powerset of the universe of instances in such a way that the structure is preserved, and its size reduced ? An answer to this question is given in [20] through the use of intensional and extensional projections whose images are also lattices:

Definition 3 (Projection). *p is a projection of a lattice $(M, \leq)$ iff for each pair (x, y) of elements of M, we have :*

If $x \leq y$ then $p(x) \leq p(y)$ (monotonicity)
$p(x) \leq x$ (minimality)
$p(x) \leq p(p(x))$ (semi-idempotence)

Let M be a lattice and p a projection of M, then $p(M)$ is also a lattice with join operator $\vee$ and meet operator $\wedge$ defined as, for any pair $m_1, m_2 \in p(M)$, $m_1 \wedge m_2 = p(m_1 \wedge_M m_2)$ and $m_1 \vee m_2 = m_1 \vee_M m_2$.

When considering respectively $M = \mathcal{L}$ and $M = \mathcal{P}(O)$ we obtain respectively intensional and extensional projections and both lead to *projected extension-intension lattices* [20,28].

Proposition 2. *Let (int, ext) be a Galois connection on $(\mathcal{P}(O), \preceq), (\mathcal{L}, \subseteq)$, G be the associated Galois lattice, and (e, t) be a node of G.*

- *Let p be a projection on $\mathcal{L}$, then $(p \circ int, ext)$ defines a Galois connection on $((p(\mathcal{L}), \preceq), (\mathcal{P}(O), \subseteq))$ and (e, t) is projected in the corresponding Galois lattice $p(G)$ on the node (e', t') such that $t' = p(t)$ and $e' = ext(t')$.*
- *Let p be a projection on $\mathcal{P}(O)$, then $(int, p \circ ext)$ defines a Galois connection on $((\mathcal{L}, \preceq), (p(\mathcal{P}(O)), \subseteq))$ and (e, t) is projected in the corresponding Galois lattice $p(G)$ on the node (e', t') such that $e' = p(e)$ and $t' = int(e')$.*

Note that in partial order theory, projections are known as *kernel operators* or *interior operators*. Their properties are well known [11] and are the basis of the next section.

3 Abstractions

3.1 From Projection to Abstractions

Reducing through projections $\mathcal{L}$, $\mathcal{P}(O)$ or both results in a reduced extension-intension representation whose latticial structure is preserved. However, while projections are technically useful, they do not always give a simple way to chose simplified representations. In what follows we present an equivalent view of projections.

Definition 4. *An abstraction of a lattice M is a subset [1] of M, closed under $\vee_M$.*

Building abstractions therefore simply means to chose any subset of M and close it by $\vee_M$. We note hereunder that abstractions are in a one to one correspondence with projections and so *abstract extension-intension lattices* are defined as projected extension-intension lattices.

Proposition 3. *Let A be an abstraction of a lattice M, then p_A defined as $p_A(x) = \bigvee_{c \in A, c \leq x} c$, is a projection of M. Let p be a projection then $A = p(M)$ is an abstraction of M, and p is the projection p_A associated to A.*

The equivalence between *interior systems* $p(M)$ and subsets A of M closed under union is a known property of interior operators [11]. As abstractions are closed under $\vee_M$, we have that $\vee_M = \vee_A$. From now on we simply write $\vee$ when no confusion is possible. An important point is that we only need the $\vee$-irreducible elements[2] of A:

---

[1] Note that $\perp_M$, the smallest element of M, belongs to all abstractions.
[2] Irreducible elements of A are elements that cannot be obtained as a result of applying $\vee_M$.

Proposition 4. *Let A be an abstraction of M, and A_I be the set of $\vee$-irreducible elements of A, the projection p_A is obtained by only considering elements of A_I:*

$$p_A(x) = \bigvee_{c \in A_I, c \leq x} c$$

Intensional abstractions An *intensional abstraction* is then simply obtained by selecting part of the language $\mathcal{L}$ and closing it by the join operator $\vee_\mathcal{L}$. For instance, consider the lattice $\mathcal{L}$ whose elements are intervals $[a, b]$ such that $a, b \in \{1, 2, 3, 4\}$. We have then $[1, 3] \vee_\mathcal{L} [3, 4] = [3]$. Consider first the abstraction A whose elements are the intervals containing 3. A is closed by $\vee_\mathcal{L}$ as intersecting two intervals containing 3 results in an interval also containing 3.

Consider now $\mathcal{L}' = \mathcal{L} - \{[1], [2], [3], [4]\}$. $\mathcal{L}'$ is obtained by simply deleting the most specific terms and is clearly also a lattice. However $[1, 3] \vee_{\mathcal{L}'} [3, 4]$ is now $[]$ and so clearly $\mathcal{L}'$ is not closed under $\vee_\mathcal{L}$, and therefore is not an abstraction. Note that there are in $\mathcal{L}'$ two most specific elements satisfied by $o = 3$, and as a consequence $d(o)$ is no longer defined.

Extensional abstractions. They abstract instances rather than abstracting the language of terms. Note that when considering extensional abstractions, $\vee$ is the set theoretic union $\cup$, and that the order relation is the set theoretic inclusion.

Example 1. As an example consider $O = \{1, 2, 3, 4\}$, and the abstraction A obtained by closing under union the part $\{\{1, 2\}, \{1, 3\}\}$ of $\mathcal{P}(O)$, so adding $\{1, 2, 3\}$ and $\emptyset$ to build A. The set of $\cup$-irreducible elements of A is $A_I = \{\{1, 2\}, \{1, 3\}\}$. For instance, $p_A(\{1, 2, 3\}) = \{1, 2\} \cup \{1, 3\} = \{1, 2, 3\}$ as both elements of A_I are included in $\{1, 2, 3\}$, and $p_A(\{2, 3\}) = \emptyset$ because neither $\{1, 2\}$ nor $\{1, 3\}$ is included in $\{2, 3\}$.

We remark in the next section that abstractions are partially ordered.

3.2   The Lattice of Abstractions

There is a partial order on projections of a lattice M [20] and therefore, on abstractions:

Definition 5. *Let M be a lattice and p_1 et p_2 two projections of M, we will state that $p_2 \leq p_1$, i.e. p_2 is less* concrete, *and so more* abstract *than p_1, iff there is some projection p defined on $p_1(M)$ such that for all c in M, $p_2(c) = p \circ p_1(c)$.*

This means that $A_2 = p_2(M)$ is more abstract than $A_1 = p_1(M)$ *iff* A_2 is an abstraction of A_1. This is also equivalent to saying that any $\vee$-irréductible element i_2 of A_2 may be written as a disjunction of $\vee$-irreducible elements of A_1, *i.e.* $i_2 = i_1^1 \vee \cdots \vee i_1^n$

Example 2 (Extensional abstractions). Let $M = \mathcal{P}(\{1, 2, 3, 4\})$, $A_{I1} = \{\{1, 2\}, \{1, 3\}\}$ and $A_{2I} = \{\{1, 2\}, \{1, 2, 3\}\}$, A_2 is more abstract than A_1 as $\{1, 2\} \in A_{I1}$ and $\{1, 2, 3\} = \{1, 2\} \cup \{1, 3\}$.

Proposition 5 (Lattice of abstractions). *Let M be a lattice, and $\mathcal{A}$ be the set of the abstractions of M. Let then $A_1 = p_1(M)$ and $A_2 = p_2(M)$ be elements of $\mathcal{A}$:*

- *The least upper bound $A_1 \vee_{\mathcal{A}} A_2$ is obtained by closing $A_1 \cup A_2$ under $\vee$.*
- *The greatest lower bound $A_1 \wedge_{\mathcal{A}} A_2$ is simply $A_1 \cap A_2$.*

This partial order, defined through projections, is known to order projected extension-intension lattices [20]. Therefore, abstract extension-intension lattices, where abstraction is performed on $\mathcal{L}$ or $\mathcal{P}(O)$, are also ordered following the lattice of abstractions.

4 Extensional Abstractions and Extensional Abstract Lattices

4.1 Extensional Abstract Lattices

An extensional abstraction is obtained by considering part of the powerset of O, and closing it by the union operator $\cup$. This means that we do not consider any more instances but rather subsets of instances, called *abstract instances*.

Definition 6. *Let A be an extensional abstraction and p_A the associated projection of $\mathcal{P}(O)$, then let t be a term of $\mathcal{L}$, $ext_A(t) = p_A \circ ext(t)$ is the* abstract extension *of t.*
   From Proposition 4 we also have

$$ext_A(t) = \bigcup_{u \in A_I, u \subseteq ext(t)} u$$

Definition 7. *Let A be an abstraction of $\mathcal{P}(O)$, the projected extension-intension lattice $p_A(G)$ is denoted as an* extensional abstract lattice, *and noted G_A.*

Hereunder we define *abstract instances* and *minimal abstractions* of a given instance.

Definition 8. *Let A be an extensional abstraction.*

- *An element of A_I is called an* abstract instance.
- *Let $A_I(o)$ be the subset of A_I whose elements contain the instance o. We denote as* minimal abstractions *of an instance o the minimal elements of $A_I(o)$, i.e. $A_m(o) = \{u \in A_I \mid o \in u \text{ and if } o \in u' \subset u, \text{ then } u' \notin A\}$.*

An interesting point is that A_I is the set of minimal abstractions of the instances:

Proposition 6.
$$A_I = \bigcup_{\{o \in O\}} A_m(o)$$

Example 3. In this example $O = \{1, 2, 3, 4\}$ and $A_I = \{\{1, 2\}, \{2, 3\}, \{3, 4\}\}$. Here $\mathcal{L} = \mathcal{P}(P)$ where $P = \{X, Y, Z\}$ is a set of properties. We give hereunder a table representing the context relating $\mathcal{P}(P)$ to $\mathcal{P}(O)$. We have added a column for each element of A_I. In Figure 1 we have represented the original concept lattice together with the extensional abstract lattice (here an *abstract concept lattice*) associated to A_I.

Instances	X	Y	Z	{1,2}	{2,3}	{3,4}
1	1	0	1	1	0	0
2	1	1	1	1	1	0
3	0	1	0	0	1	1
4	1	0	0	0	0	1
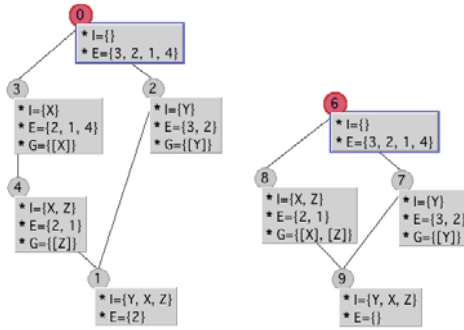
Fig. 1. The concept lattice (on the left) and the abstract concept lattice (on the right) corresponding to the context and abstraction given in Example 3. On each node I is the intension, E is the extension, and G is the set of most general terms whose extension is E. The nodes 3 and 4 of the original concept lattice are merged into the node 8 of the abstract concept lattice. As a result, the abstract implication $X \to Z$ is now valid as X and Z have the same abstract extension.

We may note that such an extensional abstract lattice is isomorphic to an extension-intension lattice relating $\mathcal{L}$ to the powerset of A_I. For that purpose, we reformulate A_I as a new instance set and we define $(ext' : \mathcal{L} \to A_I)$ as follows: let $u = \{o_1, \ldots o_k\}$ be an element of A_I, we have then that u, as a new instance, belongs to $ext'(t)$ whenever u, as a subset of O, is included in $ext(t)$. As a consequence $ext_A(t) = \bigcup_{u \in ext'(t)} u$ and therefore the images $ext'(\mathcal{L})$ and $ext_A(\mathcal{L})$ are in one to one correspondence.

As stated before, an extension-intension lattice corresponds to a set of valid implications. We define hereunder the abstract implications associated to an extensional abstract lattice.

Definition 9 (Abstract implication). *Let t_1 et t_2 be terms of $\mathcal{L}$. Whenever $ext_A(t_1) \subseteq ext_A(t_2)$ we say that the abstract implication $t_1 \to_A t_2$ is valid on A.*

Then G_A is represented by the whole set of abstract implications valid on A, or by any generating subset, as the *min-max basis of abstract implications* extending to the extensional abstract lattices the definition of the min-max basis of implications (see end of Section 2.2).

Now as $t_1 \to t_2$ means that $ext(t_1) \subseteq ext(t_2)$, we have that whenever some abstract instance u is included in $ext(t_1)$, we also have that u is included in $ext(t_2)$. This also means, by definition, that $ext_A(t_1)$ is included in $ext_A(t_2)$. As a consequence validity of implications is preserved by extensional abstraction:

Proposition 7. *Let A be an abstraction of $\mathcal{P}(O)$, and t_1 and t_2 two terms of $\mathcal{L}$. If $t_1 \to t_2$ is valid on O, then the abstract implication $t_1 \to_A t_2$ is valid on A.*

4.2   Alpha Lattices as Extensional Abstract Lattices

The partial order of abstractions (see Definition 5), means that A' is more abstract than A whenever any element of A' (or A'_I) may be written as the union of elements of A

(or A_I). An interesting case is *alpha abstraction*. Starting from a subset $\mathcal{C}$ whose union closure is an initial abstraction A, more concrete abstractions A^α are built. A^α is obtained by deriving from each initial category C the set of frequent enough parts of C, and closing under union the resulting set $\mathcal{C}^\alpha$. We have then that A^α is more abstract than $A^{\alpha'}$ iff $\alpha \geq \alpha'$. In a previous work the corresponding *alpha Galois lattices* were defined through projections and experimented in order to extract alpha association rules [28]. As an example, consider the implication "Animals that fly are oviparous", it is not valid on O because of the *bat*. When considering the categorization $\mathcal{C} = \{mammal, insect, bird\}$, the corresponding abstract implication is valid but never applies: none of the initial categories contains only flying animals. However by considering frequent enough parts of each category, the corresponding $(\alpha = 0.1)$- implication is valid, as the bat is eliminated from the premise (very few mammals fly), but still applies to flying birds and flying insects, as they represent large enough parts of their categories.

5   The Extended Abstract Lattice $G_\square$

We will formalize the nature of abstract implications by interpreting them as classical implications relating modalized terms. So we rewrite $t_1 \rightarrow_A t_2$ as $\square_A t_1 \rightarrow \square_A t_2$. We first define a modal logics of abstraction built on a propositional modal language $\mathcal{L}_{mod}$. Then we restrict $\mathcal{L}_{mod}$ to a language $\mathcal{L}\square$ using only conjunction and non nested modal connectors. Finally we present an extended abstract lattice $G_\square$ where intensions have both modalized and non modalized parts.

5.1   Modal Logics of Abstraction

Hereunder we discuss the modal logics of abstractions. Note that in this section we denote as *worlds* the elements of O.

A modal logic, in its simplest form, is a propositional logic to which is added at least one unary modal connector $\square$, referred to as a *necessity* operator. *Classical modal logics* are the modal logics in which formulas are given truth values through *neighborhood semantics*, also known as *minimal models semantics* [7]. This wide class of modal logics includes in particular *normal modal logics* relying on Kripke possible world semantics. We will define hereunder *abstract modal logics* as particular classical modal logics. We hereunder very informally summarize neighborhood semantics in order to relate such semantics to our extensional abstractions.

The modal language $\mathcal{L}_{mod}$ is obtained by adding the modal connector $\square$ to a propositional language $\mathcal{L}$ built on a set of atomic propositions. So for instance, $\phi = a \wedge \square(b \wedge c)$ belongs to $\mathcal{L}_{mod}$. In order to give a truth value to a modal formula we first consider a set of worlds O, together with a *valuation* function ext, relating each atomic proposition p to the set of worlds in which p is true. For any formula ϕ without modal connectors, the computation of $ext(\phi)$ is then straightforward, for instance $ext(a \wedge b) = ext(a) \cap ext(b)$ and represents the worlds o in which ϕ is true, or in other words, the worlds that satisfies ϕ. In order to extend ext to modal formulas we need a neighboring function

$\mathcal{N}$ relating each world o to a set of subsets of O and we say that the world o is in $ext(\Box\phi)$ whenever $ext(\phi)$ belongs to $\mathcal{N}(o)$. It is known that neighboring functions $\mathcal{N}$ are in one to one correspondence with mappings $m : \mathcal{P}(O) \rightarrow \mathcal{P}(O)$ such that $m(e) = \{o \in O \mid e \in \mathcal{N}(o)\}$. To summarize we have now that

$$ext(\Box\phi) = m \circ ext(\phi)$$

Now recall that we defined abstract extension as $p \circ ext$ where p is a projection. We will so naturally define *abstract modal logics* as classical modal logics in which the mapping m is a projection. To characterize abstract modal logics we have just to translate the properties of projection as axioms and inference rules. Detailed results on modal logics of abstraction, including multimodal logics allowing to access to various levels of abstraction, are outside the scope of this paper [22]. Now, considering the abstract extension of a term as the extension of an abstract term leads to define extension-intension lattices built on languages whose terms contain both non abstract and abstract parts. This is the subject of the remaining of this section.

5.2   Mixing Abstract and Non Abstract Statements: The Language $\mathcal{L}\Box$

We consider here a language $\mathcal{L}\Box$ where classical properties and modalized properties appear simultaneously. Technically $\mathcal{L}\Box$ is a subset of $\mathcal{L}_{mod}$ whose terms contain $\wedge$ and $\Box$ as connectors, and in which the nesting of $\Box$ is not allowed. For instance, $(a \wedge b) \wedge \Box(b \wedge c) \wedge \Box(c \wedge d) \in \mathcal{L}\Box$, but $\Box((a \wedge b) \wedge \Box(b \wedge c)) \notin \mathcal{L}\Box$.

We first give an inductive definition of $\mathcal{L}\Box$.

Definition 10 (Inductive definition of $\mathcal{L}_P$ and $\mathcal{L}\Box$). *The inductive definition of $\mathcal{L}\Box$ according to P, a non-empty set of atoms, is the following (note that we define and use the language $\mathcal{L}_P$ corresponding to terms without any $\Box$ connective):*

- $\forall a \in P, a \in \mathcal{L}_P$
- *Constants $\top$ and $\bot \in \mathcal{L}_P$*
- *if F1 and F2 $\in \mathcal{L}_P$ then $(F1 \wedge F2) \in \mathcal{L}_P$*
- *if $F \in \mathcal{L}_P$ then $F \in \mathcal{L}\Box$*
- *if $F \in \mathcal{L}_P$ then $\Box F \in \mathcal{L}\Box$*
- *if $F1 \wedge F2 \in \mathcal{L}\Box$ then $(F1 \wedge F2) \in \mathcal{L}\Box$*

An intensional semantics for $\mathcal{L}\Box$

We define an intensional semantics for $\mathcal{L}\Box$ based on an algebraic approach similar to [27], [10] but for description logics. The definition is made in two steps. In the first step, an equational system which highlights the main properties of the $\mathcal{L}\Box$ connectives is given. During the second step, an homomorphism based on the equational system is defined. This homomorphism is used to map terms of $\mathcal{L}\Box$ to their *structural normal form* (snf) in the intensional semantics.

Definition 11 . (The equational system $EQ_\square$)
$\forall$ F1, F2, F3 $\in \mathcal{L}\square$, $\forall$ F, F' $\in \mathcal{L}_P$:

1. $(F1 \wedge F2) \wedge F3 = F1 \wedge (F2 \wedge F3)$
2. $F1 \wedge F2 = F2 \wedge F1$
3. $F1 \wedge F1 = F1$
4. $\top \wedge F1 = F1$
5. $\bot \wedge F1 = \bot$
6. $\square F = \square F \wedge F$
7. $\square(F \wedge F') = \square(F \wedge F') \wedge \square F \wedge \square F'$

The equational system $EQ_\square$ fixes the main properties of the connectives and can be used to define an equivalence relation between terms of $\mathcal{L}\square$. Equality modulo axioms of $EQ_\square$ is denoted by $\equiv_{EQ_\square}$. We use it to formalize the subsumption relation in $\mathcal{L}\square$.

Definition 12 (Subsumption in $\mathcal{L}\square$). *Let F1, and F2 be two terms of $\mathcal{L}\square$, F1 $\preceq$ F2 (i.e. F1 subsumes or is less specific than F2) iff F1 $\wedge$ F2 $\equiv_{EQ_\square}$ F2.*

$EQ_\square$ induces a class of algebras. From this class, a structural algebra can be proposed, which provides $\mathcal{L}\square$ with an intensional semantics called $\mathcal{CL}\square$. The elements of $\mathcal{CL}\square$ are structures whose definition is given below. These elements are *structural normal forms* of terms of $\mathcal{L}\square$ allowing us to obtain an unique class representative for each equivalence classes of terms. $\mathcal{CL}\square$ can be viewed as a normalized subset of $\mathcal{L}\square$ where the lgg is unique.

The intensional semantics $\mathcal{CL}\square$
An element of $\mathcal{CL}\square$ corresponding to a term T of $\mathcal{L}\square$ denoted $snf(T)$ is a pair defined as follows: $< Eclassical, E\square >$. *Eclassical* is a set of atoms belonging to P , $E\square$ is a set of subsets of P. Intuitively, *Eclassical* contains every explicit and implicit classical properties of T (i.e. properties not at reach to a $\square$ connective). The data structure used can then be a simple set of atoms. $E\square$ contains $\square$ properties of T. Since for instance $\square(a \wedge b)$ and $\square(b \wedge d)$ cannot be compared, we must keep both of them. The data structure used is a set of sets.

   This definition presents the data structure of elements of $\mathcal{CL}\square$ but not how to associate to a term of $\mathcal{L}\square$ its corresponding element in $\mathcal{CL}\square$. To make this computation and then to define $\mathcal{CL}\square$, an homomorphism from the set of terms of $\mathcal{L}\square$ and the set of elements of $\mathcal{CL}\square$ need to be defined.

   The homomorphism from $\mathcal{L}\square$ into $\mathcal{CL}\square$ is sketched below. It allows us to associate to each term T of $\mathcal{L}\square$ its *structural normal form* denoted $snf(T)$. This homoporphism takes into account axioms of $EQ_\square$ and the normalization strategy choosen. Indeed, to obtain a normal form many normalization strategies may be applied (*e.g.* deletion of redundant information). We chose to add implicit information in the classical part, this strategy is a kind of partial saturation which is a trick largely used to make easier subsumption and lgg computation. On the other hand, for complexity reasons we only keep maximal subsets in the $\square$ part. The $\square$ part is a Sperner family i.e., an antichain in the inclusion lattice over the power set of $\mathcal{L}_P$ (for more details see [2]).

Homomorphism snf from $\mathcal{L}\Box$ into $\mathcal{CL}\Box$:

Term of $\mathcal{L}\Box$	Element in $\mathcal{CL}\Box$
$\top$	$t = <\emptyset,\{\emptyset\}>$
$\bot$	$bo = <P,\{P\}>$
$a \in P$	$<\{a\},\{\emptyset\}>$
$F1 \wedge F2$	$snf(F1) \vee_s snf(F2)$
$\Box F$	$<E_F,\{E_F\}>$

where $snf(F) = <E_F,\emptyset>$ (the $\Box$ part is empty since there is no $\Box$ nesting).

$\vee_s$ is the join operator in $\mathcal{CL}\Box$. It uses the classical union set operator and $\vee_a$ which represents the union operator in antichains.
Let $K1$ and $K2$ be two antichains:

$$K1 \vee_a K2 = \{x \in K1 \cup K2 \mid \nexists y \in K1 \cup K2 \text{ s.t. } x \subset y\}$$

Let $snf(Fi)$ be $<E_{Fi}, Ki>$ for $i = 1, 2$:

$$snf(F1) \vee_s snf(F2) = <E_{F1} \cup E_{F2}, K1 \vee_a K2>$$

Example 4. $snf((a \wedge b) \wedge \Box(b \wedge c) \wedge \Box b \wedge \Box(c \wedge d)) = <\{a,b,c,d\},\{\{b,c\},\{c,d\}\}>$
c and d are added in the classical part according to axiom 6 and 7, $\{b\}$ is removed from the $\Box$ part since $\{b\} \subseteq \{b,c\}$.

Definition 13 (Structural Subsumption in $\mathcal{CL}\Box$). *Let S1 and S2 be two elements of* $\mathcal{CL}\Box$,

$$S1 \preceq_s S2 \text{ (i.e. S1 subsumes S2) iff } S1 \vee_s S2 = S2$$

Proposition 8. *Subsumption in $\mathcal{L}\Box$ is equivalent to structural Subsumption in $\mathcal{CL}\Box$: Let F1, and F2 be two terms of $\mathcal{L}\Box$,*

$$F1 \preceq F2 \text{ iff } snf(F1) \preceq_s snf(F2)$$

5.3   The Galois Lattice G$_\Box$

We consider the two following posets : $(\mathcal{P}(O),\subseteq)$ and $(\mathcal{CL}\Box, \preceq_s)$. In order to obtain a Galois connection between the two posets, we need to define the least general generalisation in $\mathcal{CL}\Box$. The *lgg* uses the following definition of $\wedge_a$ the intersection operator in antichain: any antichain A corresponds to a lower set $L_A = \{x \in A \mid \exists y \in A \text{ s.t. } x \subseteq y\}$. Let A and B be two antichains and L_A and L_B their lower sets:

$$A \wedge_a B = \{x \in L_A \cap L_B \mid \nexists y \in L_A \cap L_B \text{ s.t .} x \subset y\}$$

Definition 14 (Least General Generalisation in $\mathcal{CL}\Box$). *Let S1 = $<$E1,K1$>$ and S2 = $<$E2, K2$>$ be two elements in $\mathcal{CL}\Box$:*

$$lgg(S1,S2) = <E1 \cap E2,K1 \wedge_a K2>$$

Let EnsS be a finite set of elements of $\mathcal{CL}\Box$ = $\{s1,s2,\ldots,sn\}$:

$$lgg(EnsS) = lgg(s1,lgg(s2,\ldots lgg(sn\text{-}1,sn)))$$

Each instance o is described by a term T of $\mathcal{L}\square$. In order to obtain a unique most specific description in $\mathcal{CL}\square$ satisfied by o, we simply use $snf(T)$ denoted $d(o)$ in the following.

Proposition 9. *Let ($\mathcal{CL}\square$, $\preceq_s$) and ($\mathcal{P}(O)$,$\subseteq$) be two posets. The pair (ext,int) defined by:*

$$\forall\, S \in \mathcal{CL}\square,\ ext(S) = \{o \in O \mid S \preceq_s d(o)\}$$
$$\forall\, E \in \mathcal{P}(O),\ int(E) = lgg(d(E))\ with\ d(E) =\{\, d(o) \mid \forall\, o \in E\}$$

defines a Galois connection between ($\mathcal{CL}\square$, $\preceq_s$) and ($\mathcal{P}(O)$,$\subseteq$)

Proof : as stated in proposition 4 p11 in [15], (ext,int) is a Galois connection iff:

$$S \preceq_s int(E) \Leftrightarrow E \subseteq ext(S)$$

1) $S \preceq_s int(E) \Rightarrow S \preceq_s lgg(d(E))$
$\forall o \in E, lgg(d(E)) \preceq_s d(o)$ since $lgg(d(E)) = lgg(d(o), lgg(d(E - \{o\}))$
$\Rightarrow S \preceq_s d(o)$ since $S \preceq_s lgg(d(E)) \Rightarrow o \in ext(S)$ by definition of $ext \Rightarrow E \subseteq ext(S)$
2) $E \subseteq ext(S) \Rightarrow (\forall o \in E \Rightarrow o \in ext(S)) \Rightarrow \forall o \in E, S \preceq_s d(o) \Rightarrow S \preceq_s int(E)$

Proposition 10. *The extension-intension lattice $G_\square$ defined by the Galois connection of proposition 9 is denoted as an extended abstract lattice.*

When considering a given extensional abstraction A, we have now a "pure" extensional abstract lattice G_A, as defined in section 4 together with the new intension/extension lattice $G_\square$ defined on the connection between $\mathcal{CL}\square$ and $\mathcal{P}(O)$ and relying, for its modalized part, on $ext(\square t) = p_A \circ ext(t)$. We are now interested in the exact relations between these lattices and the original concept lattice G relating the language $\mathcal{P}(P)$ to $\mathcal{P}(O)$. We first remark that $\mathcal{P}(P)$ is isomorphic to an abstraction of $\mathcal{CL}\square$ as the classical part of $\mathcal{CL}\square$ is closed under the lgg operator. Therefore G is obtained as an intensional abstraction of $G_\square$. A second remark is that the extensional abstract lattice $p_A(G_\square)$ is isomorphic to G_A: consider some intension t in G, its abstract extension $e = ext_A(t)$ rewrites as $ext(\square t)$, and the representation of $\square t$ is the most specific element in $\mathcal{CL}\square$ whose extension contains e. As a consequence, G_A and G are both less abstract than $G_\square$. We draw Figure 2 these three lattices in a very simple example with two atomic properties X and Y, 4 instances and the following abstraction $A_I = \{\{1, 2\}, \{2, 3\}, \{3, 4\}\}$. Note that in the table hereunder, the columns $\square t$ represent the abstract extensions $ext_A(t)$. For instance the column $\square XY$ represents the empty set since $ext(XY) = \{2\}$ and that no abstract instance of A_I is included in $ext(XY)$.

Instance	X	Y	□ X	□ Y	□ XY	{1,2}	{2,3}	{3,4}
1	1	0	1	0	0	1	0	0
2	1	1	1	1	0	1	1	0
3	0	1	0	1	0	0	1	1
4	1	0	0	0	0	0	0	1
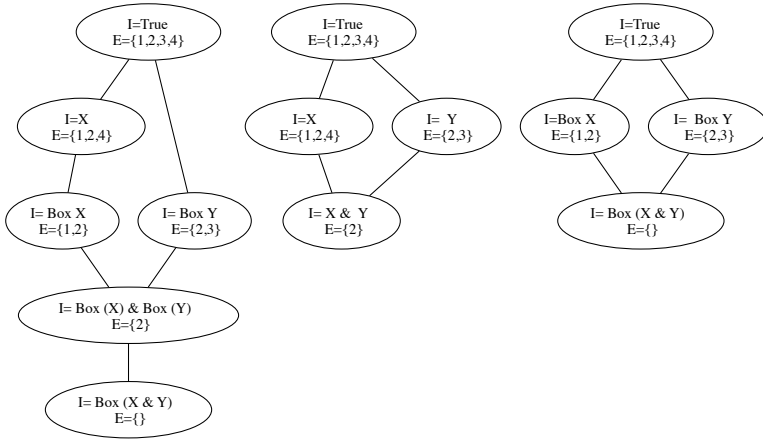
Fig. 2. The three intension/extension lattices corresponding to the example of section 5.2. The node label is a minimal representation of the intension I (here *Box* and *&* stand for $\square$ and $\wedge$) together with its extension E in $\{1, 2, 3, 4\}$. The leftmost lattice is $G_\square$, the centermost lattice is the original concept lattice G, and the rightmost lattice is the projected lattice $p_A(G_\square)$ isomorphic to the abstract lattice G_A.

6   Related Work and Conclusion

In this paper we have proposed abstractions as reductions that preserve the extension-intension lattice structure.They are simply defined as parts of either the intensional language or the extensional space that are closed under the *join* operator. A second contribution is the investigation of extensional abstractions. We have shown that defining an extensional abstraction $A \subseteq \mathcal{P}(O)$ consists in *a priori* defining as units particular subsets of instances, denoted as abstract instances, so applying a change in extensional granularity. A noticeable effect of so preserving the Galois lattice structure is that validity of implications is preserved through extensional abstractions. Finally, we interpret abstract implications as classical implications between two modalized terms. This leads to define *abstract modal logics* and to define an *extended abstract concept lattice* relating $\mathcal{P}(O)$ to a language whose elements have abstract and non abstract parts. We can then search for concepts as "*Oviparous and abstractly apt-to-fly*" where "*o* satifies *abstractly apt-to fly*" means that all the instances of some minimal abstraction of *o* share this property. Regarding the implementation of abstract concept lattices, we can benefit from implementations of alpha lattices, and in particular their incremental construction [23] adapted from [26], and a software, based on `Galicia` [24], is available[3]. However, *extended abstract concept lattice* construction still has to be investigated, both from an algorithmic and practical point of view. Note that as *extended abstract lattices* belongs to the class of logical concept lattices [13] yet there exists a way to implement them.

---

[3] `http://www-lipn.univ-paris13.fr/~ champesme/alphabetagalicia`

References

1. Ben-David, S., Ben-Eliyahu-Zohary, R.: A modal logic for subjective default reasoning. Artif. Intell. 116(1-2), 217–236 (2000)
2. Bollobás, B.: Combinatorics: set systems, hypergraphs, families of vectors, and combinatorial probability. Cambridge University Press, New York (1986)
3. Brézellec, P., Soldano, H.: Tabata: a learning algorithm performing a bidirectional search in a reduced search space using a tabu stratégie. In: Eur. Conf. in Art. Int., ECAI 1998, Brighton, England, pp. 420–424. Wiley and Sons, Chichester (1998)
4. Caspard, N., Monjardet, B.: The lattices of closure systems, closure operators, and implicational systems on a finite set: a survey. Discrete Appl. Math. 127(2), 241–269 (2003)
5. Cattaneo, G., Ciucci, D.: Lattices with interior and closure operators and abstract approximation spaces. T. Rough Sets 10, 67–116 (2009)
6. Chaudron, L., Maille, N.: Generalized formal concept analysis. In: Ganter, B., Mineau, G.W. (eds.) ICCS 2000. LNCS, vol. 1867, pp. 357–370. Springer, Heidelberg (2000)
7. Chellas, B.F.: Modal Logic: an introduction. Cambridge University Press, Cambridge (1980)
8. Davvaz, B., Mahdavipour, M.: Rough approximations in a general approximation space and their fundamental properties. International Journal of General Systems 37(3), 373–386 (2008)
9. Diday, E., Emilion, R.: Maximal and stochastic galois lattices. Discrete Appl. Math. 127(2), 271–284 (2003)
10. Dionne, R., Mays, E., Oles, F.J.: The equivalence of model-theoretic and structural subsumption in description logics. In: Int. Joint Conf. on Art. Int. (IJCAI), pp. 710–717 (1993)
11. Erné, M., Koslowski, J., Melton, A., Strecker, G.E.: A Primer on Galois Connections. Annals of the New York Academy of Sciences 704(Papers on General Topology and Applications), 103–125 (1993)
12. Ferré, S.: Negation, opposition, and possibility in logical concept analysis. In: Missaoui, R., Schmidt, J. (eds.) ICFCA 2006. LNCS (LNAI), vol. 3874, pp. 130–145. Springer, Heidelberg (2006)
13. Ferré, S., Ridoux, O.: An introduction to logical information systems. Information Processing and Management 40(3), 383–419 (2004)
14. Ganascia, J.-G.: TDIS: an algebraic formalization. In: Int. Joint Conf. on Art. Int (IJCAI), vol. 2, pp. 1008–1013 (1993)
15. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer, Heidelberg (1999)
16. Ganter, B., Kuznetsov, S.O.: Pattern structures and their projections. In: Delugach, H.S., Stumme, G. (eds.) ICCS 2001. LNCS (LNAI), vol. 2120, pp. 129–142. Springer, Heidelberg (2001)
17. Guigues, J.L., Duquenne, V.: Famille non redondante d'implications informatives résultant d'un tableau de données binaires. Mathématiques et Sciences Humaines 95, 5–18 (1986)
18. Liquière, M., Sallantin, J.: Structural machine learning with Galois Lattice and Graphs. In: ECML 1998, pp. 305–313. Morgan Kaufmann Publishers Inc., San Francisco (1998)
19. Pasquier, N., Taouil, R., Bastide, Y., Stumme, G., Lakhal, L.: Generating a condensed representation for association rules. Journal of Intelligent Information Systems (JIIS) 24(1), 29–60 (2005)

20. Pernelle, N., Rousset, M.-C., Soldano, H., Ventos, V.: Zoom: a nested Galois lattices-based system for conceptual clustering. J. of Experimental and Theoretical Artificial Intelligence 2/3(14), 157–187 (2002)
21. Saitta, L., Zucker, J.-D.: A model of abstraction in visual perception. Applied Artificial Intelligence 15(8), 761–776 (2001)
22. Soldano, H.: A modal view on abstract learning and reasoning. Technical report, L.I.P.N, UMR-CNRS 7030, Université Paris-Nord (2010)
23. Soldano, H., Ventos, V., Champesme, M., Forge, D.: Incremental construction of alpha lattices and association rules. In: Setchi, R., Jordanov, I., Howlett, R.J., Jain, L.C. (eds.) KES 2010. LNCS, vol. 6277, pp. 351–360. Springer, Heidelberg (2010)
24. Valtchev, P., Grosser, D., Roume, C., Hacene, M.R.: Galicia: An open platform for lattices. In: In Using Conceptual Structures: Contributions to the 11th Intl. Conference on Conceptual Structures (ICCS 2003), pp. 241–254. Shaker Verlag (2003)
25. Valtchev, P., Missaoui, R., Godin, R.: Formal concept analysis for knowledge discovery and data mining: The new challenges. In: Eklund, P.W. (ed.) ICFCA 2004. LNCS (LNAI), vol. 2961, pp. 352–371. Springer, Heidelberg (2004)
26. Valtchev, P., Missaoui, R., Godin, R., Meridji, M.: Generating frequent itemsets incrementally: two novel approaches based on Galois lattice theory. J. of Experimental and Theoretical Artificial Intelligence 2/3(14), 115–142 (2002)
27. Ventos, V.: A deductive study of the c-classicde description logic. In: Description Logics. AAAI Technical Report, vol. WS-96-05, pp. 192–196. AAAI Press, Menlo Park (1996)
28. Ventos, V., Soldano, H.: Alpha Galois Lattices: An Overview. In: Ganter, B., Godin, R. (eds.) ICFCA 2005. LNCS (LNAI), vol. 3403, pp. 299–314. Springer, Heidelberg (2005)
29. Yao, Y.Y., Lin, T.Y.: Generalization of rough sets using modal logics. Intelligent Automation and Soft Computing 2, 103–120 (1996)