Translations between Concept Hierarchies over Different Attribute Sets

Rainer Osswald

Department of Computer Science
Intelligent Information and Communication Systems
FernUniversität in Hagen
58084 Hagen, Germany
http://pi7.fernuni-hagen.de/osswald

Abstract. There are (at least) two different approaches to constructing concept lattices – one as the concept lattice of a given formal context, the other as a distributive lattice defined as the Lindenbaum algebra of a given set of (conjunctive or disjunctive) rules. It has been pointed out, first, that these two approaches are systematically related via Birkhoff's duality theorem between (finite) ordered sets and distributive lattices, and, second, that the concept lattices of Formal Concept Analysis are lattices only due to the restriction to conjunctive implications. The present paper shows how this duality can be naturally extended to describe the relation between concept hierarchies with differing sets of underlying attributes. To this end, an appropriate notion of attribute translation is introduced. Particular emphasis is given to the special case of extensions by attributes and statements and their effect on the corresponding concept hierarchies.

1 Introduction

It is well-known that within Formal Concept Analysis (FCA), the formal concepts defined by a formal context are uniquely determined by those attribute sets that are closed under all conjunctive implicational statements holding in the formal context. This systematic relationship between the statements holding in a formal context and the conceptual hierarchies consisting of closed attribute sets is not restricted to conjunctive implications and concept lattices but can be straightforwardly generalized to implications whose premise and conclusion are arbitrary *affirmative terms*, i.e., terms that may also contain disjunction, truth, and falsity [6, 14, 7, 2, 13]. The resulting hierarchies, which are not necessarily lattices any more, are also known as *information domains* [5]; they are directed-complete ordered sets and include all finite partial orders.

Other approaches towards constructing conceptual hierarchies based on a set of statements employ the distributive lattice of affirmative terms modulo the equivalence induced by the given statements [8–10], that is, the *Lindenbaum algebra* determined by the theory. Interestingly, both [10] and [9] deny any close connection of their approaches to FCA. As it has been pointed out in [12],

Radim Bělohlávek, Václav Snášel (Eds.): CLA 2005, pp. 92–106, ISBN 80–248–0863–3.

however, there is a close connection to the information domain and thus to the FCA approach: The concept lattice of FCA can be identified with the information domain of the conjunctive theory determined by the formal context, whereas the conceptual lattices studied in [8–10] are the Lindenbaum algebras of the respective theories.

In the present paper, the systemic correspondence between statements, information domains, and distributive lattices is extended to the case of translating between different base vocabulary. To this end, we introduce an appropriate notion of *translation* between statement sets over different vocabularies. Particular emphasis will be given to the special case of theory *extensions*. Proofs are omitted and can be found in [11].

2 Theories and Information Domains

2.1 Theories and Models

Suppose Σ is a set of (primitive) attributes that can be employed to classify the elements of a certain domain of discourse U. Let $\models$ be the corresponding *satisfaction* relation from U to Σ. For convenience, let us introduce two special attributes V and Λ which satisfy respectively everything and nothing in any universe of discourse. We allow to combine attributes by the standard Boolean connectives $\wedge$, $\vee$, and $\neg$. The term algebra of the Boolean attributes inductively defined that way will be denoted by $B[\Sigma]$. As usual, $\phi \rightarrow \psi$ stands for $\neg\phi \vee \psi$ and $\phi \leftrightarrow \psi$ for $\phi \rightarrow \psi \wedge \psi \rightarrow \phi$. The satisfaction relation $\models$ can be inductively extended to a relation from U to $B[\Sigma]$ in the obvious way: $x \in U$ satisfies $\phi \wedge \psi$ iff x satisfies ϕ and ψ; similarly, x satisfies $\neg\phi$ iff x does not satisfy ϕ; etc. A compound attribute $\phi \in B[\Sigma]$ is called *affirmative* (or *positive*) just in case $\neg$ does not occur in ϕ. The term algebra of affirmative terms over Σ will be denoted by $T[\Sigma]$.

The notions introduced in the preceding paragraph can be most easily reformulated within a standard logical setting by treating attributes as *monadic predicates* (cf. also [13]). Recall that within first-order predicate logic, monadic predicates are interpreted by subsets of a universe U. A satisfaction relation $\models$ from U to Σ is thus essentially the same as an *interpretation function* M from Σ to $\wp(U)$, with $M(p) = \{x \in U \mid x \models p\}$. Such an interpretation function M can be inductively extended to a function $\hat{M}$ from $B[\Sigma]$ to $\wp(U)$, with $\hat{M}(\phi) = \{x \in U \mid x \models \phi\}$. The set $\hat{M}(\phi)$ will be referred to as the *extent* of ϕ.

Making use of the (compound) attributes in statements that *hold* or are *true* with respect to an interpretation calls for *quantifying* over these attributes. The framework presented in this paper, which covers the approaches mentioned in the introduction, only employs *universal* quantification. That is, we restrict ourselves to *universal statements* of the form $\forall x(\phi x)$, or $\forall\phi$, with $\phi \in B[\Sigma]$. A *theory* over Σ is then defined as a set of universal statements of this form. First-order predicate logic gives us the following standard notions of truth and model: A statement $\forall\phi$ is *true* with respect to an interpretation if ϕ is satisfied by every element of the universe. An interpretation is a *model* of a theory Γ if every statement of Γ is true with respect to that interpretation.

Suppose Γ and Γ' are theories over Σ. Then Γ is said to *entail* Γ' if every model of Γ is also a model of Γ'. Notice that entailment in this "semantic" sense coincides with entailment by any sound and complete inference calculus for first-order predicate logic.[1] Keeping this in mind, we write $\Gamma \vdash \Gamma'$ if Γ entails Γ'. If two theories entail each other, they are said to be *equivalent*.

A universal statement of the form $\forall(\phi \to \psi)$ (or $\forall(\phi \leftrightarrow \psi)$), with ϕ and ψ affirmative, is said to have *conditional* (or *biconditional*) *form*. In the following, we also write $\phi \preceq \psi$ and $\phi \equiv \psi$ for $\forall(\phi \to \psi)$ and $\forall(\phi \leftrightarrow \psi)$, respectively. A conditional form is called *normal*, if ϕ is purely conjunctive or V, and ψ is purely disjunctive or Λ. The normal form is called *reduced* if ϕ and ψ do not share any primitive attributes. A theory is said to have conditional (reduced normal) or biconditional form if each of its statements has this form. It is a standard exercise in elementary logic to check that every theory is equivalent to a theory in conditional (reduced normal) form as well as to a theory in biconditional form. Without restriction of generality, we can thus assume that a theory has conditional or biconditional normal form when appropriate.

2.2 Information Domains

Every theory Γ over Σ has a canonical "Henkin-style" model. Let the *canonical interpretation* of Σ in $\wp(\Sigma)$ take $p \in \Sigma$ to $\{X \subseteq \Sigma \,|\, p \in X\}$, i.e., $X \models p$ iff $p \in X$. The canonical model of Γ is then defined by eliminating all elements of $\wp(\Sigma)$ that are not compatible with the statements of Γ:

Definition 1 (Canonical Universe/Model). *The* canonical universe $C(\Gamma)$ *of a theory Γ is the set of all subsets of Σ which, under the canonical interpretation, satisfy ϕ for every statement $\forall\phi$ of Γ; that is, $C(\Gamma) = \{X \subseteq \Sigma \,|\, X \models \phi$ for every $(\forall\phi) \in \Gamma\}$. The* canonical model *of Γ has the universe $C(\Gamma)$ and takes each $p \in \Sigma$ to $\{X \in C(\Gamma) \,|\, p \in X\}$.*

It is not difficult to verify the following universal property the canonical model: a statement is entailed by Γ iff it is true with respect to the canonical model of Γ. The elements of $C(\Gamma)$ will be referred to as the *consistently Γ-closed subsets* of Σ. The canonical universe is naturally ordered by set inclusion. We say that any ordered set which is order-isomorphic to $C(\Gamma)$ "is" or represents the *information domain* of Γ, thereby adapting the terminology introduced in [5].[2]

Example 1. Let Γ be the theory consisting of the single statement $\mathit{human} \preceq \neg\mathit{feathered} \wedge \mathit{biped}$, which is equivalent to the theory $\{\mathit{human} \wedge \mathit{feathered} \preceq \Lambda, \mathit{human} \preceq \mathit{biped}\}$ in conditional normal form, and let Γ' consist of the statement $\mathit{human} \preceq \mathit{featherless} \wedge \mathit{biped}$. The information domains of Γ and Γ' are

[1] It is not difficult to devise a much simpler sound and complete inference calculus since we are working in a small fragment of first-order logic; see e.g. [11, Sect. 6.3].

[2] The canonical universe of Γ is the *free extent* of Γ in the sense of [6]. See [13] for a more detailed comparison of several terminologies.

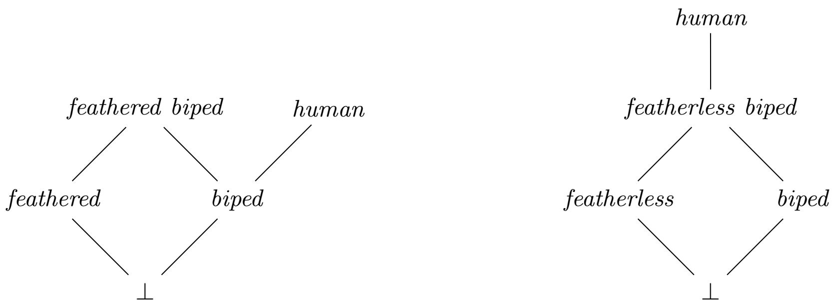

Fig. 1. Feathered and featherless bipeds

depicted on the left and the right of Figure 1, respectively, with subsets replaced by appropriate labels.

2.3 The Lindenbaum Algebra of Affirmative Terms

The Lindenbaum construction is the key step towards algebraizing logic. Its basic idea is to abstract away from syntactical differences between terms that are equivalent with respect to a given theory.

Definition 2 (Lindenbaum Algebra). *Let Γ be a theory over Σ. The* Lindenbaum algebra $L(\Gamma)$ (of affirmative terms) *of Γ is the quotient $T[\Sigma]/{\simeq_\Gamma}$, where $\phi \simeq_\Gamma \psi$ iff Γ entails $\phi \equiv \psi$.*

The Lindenbaum algebra of affirmative terms is a *distributive lattice with zero and unit*, i.e., an algebra of type $\langle 2, 2, 0, 0 \rangle$, with $[\phi] \wedge [\psi] = [\phi \wedge \psi]$, $[\phi] \vee [\psi] = [\phi \vee \psi]$, $0 = [\Lambda]$, and $1 = [V]$. It should be emphasized that the restriction to affirmative terms is essential in our definition of the Lindenbaum algebra because replacing $T[\Sigma]$ by $B[\Sigma]$ would give rise to a *Boolean lattice* (see also Section 4.2).

It follows by definition that two theories are equivalent if and only if they have the same Lindenbaum algebra. In [10], the elements of the Lindenbaum algebra $L(\Gamma)$ are referred to as the *semantic concepts* determined by Γ. Under this perspective, two affirmative terms "mean" the same (with respect to the given theory) just in case they represent the same semantic concept.

Example 2. The (non-isomorphic) Lindenbaum algebras of the theories Γ and Γ' of Example 1 are respectively depicted on the left and the right of Figure 2. The Lindenbaum algebra of Γ, for instance, is the quotient of the distributive lattice with zero and unit freely generated over the set $\{$*human*, *feathered*, *biped*$\}$ by the congruence relation generated by the pairs $\langle$*human* $\wedge$ *feathered*, $\Lambda\rangle$ and $\langle$*human*, *biped* $\wedge$ *human*$\rangle$.

In the rest of the paper, we speak of distributive lattices with zero and unit briefly as *algebras*.

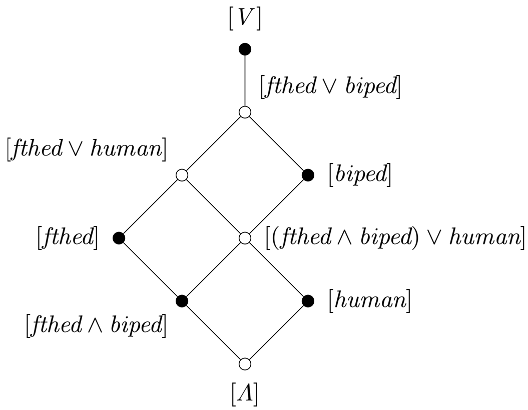

Fig. 2. Lindenbaum algebras of feathered and featherless bipeds

2.4 Birkhoff Duality

In the finite case, there is an intimate connection between the Lindenbaum algebra and the information domain of a theory in that they determine each other uniquely up to isomorphism. This correspondence is just an immediate consequence of a classical result of Birkhoff, according to which there is a categorical equivalence between the finite ordered sets and the finite distributive lattices with zero and unit (cf. [1, 4]).

A standard formulation of Birkhoff's duality is as follows: The (bounded) distributive lattice associated with a finite ordered set P is given by the lattice of upwards closed subsets of P. Conversely, the ordered set associated with a finite (bounded) distributive lattice D is given by the ordered set of $\vee$*-irreducible* elements of D. The only difference to the present situation is that we have to reverse the order of the $\vee$-irreducibles. For example, the $\vee$-irreducible elements in the lattice diagrams of Figure 2, which are marked by shaded circles, stand in an order-reversing one-to-one correspondence to the elements of the respective information domains; see Figure 1. (The $\vee$-irreducibles are characterized by the property of having precisely one element immediately below them.)

For a more explicit characterization of the Lindenbaum algebra of a theory in terms of its information domain we can employ the one-to-one correspondence between the upwards closed subsets of an ordered set and its *antichains*, where the antichain associated with an upwards closed set is the set of minimal elements of that set.[3]

Proposition 1. *In the finite case, there is a one-to-one correspondence between the elements of* $L(\Gamma)$ *and the antichains in* $C(\Gamma)$*, where an antichain* $\mathcal{S}$ *in* $C(\Gamma)$ *corresponds to the equivalence class of* $\bigvee\{\bigwedge X \mid X \in \mathcal{S}\}$ *in* $L(\Gamma)$*.*

[3] An *antichain* in an ordered set is a subset whose elements are pairwise incomparable with respect to the ordering relation.

3 Translations of Theories

3.1 Translations and Equivalences

Suppose two cognitive agents want to decide whether their theories about a certain universe of discourse are equivalent. If both agree that they use the same vocabulary in the same way, equivalence means that both theories entail each other, which is the case if every statement of one theory is entailed by the other and vice versa. This is precisely the definition of equivalent theories introduced in Section 2.1.

Consider now the case of two theories Γ and Γ' over different sets Σ and Σ' of primitives. Assume that both agents do not hinge on their chosen primitive terms but are willing to express them in terms of those of the other. Within the present framework this means to define a *translation function* μ from Σ to $T[\Sigma']$, which then can be naturally extended to a function $\hat{\mu}$ from $T[\Sigma]$ to $T[\Sigma']$. Since varying the base vocabulary makes it necessary to keep track of it, we consider theories from now on as pairs $\langle \Sigma, \Gamma \rangle$ where Γ is a theory over Σ in the sense introduced in Section 2.1. Keeping this in mind, we often write Γ instead of $\langle \Sigma, \Gamma \rangle$, if Σ is clear from the context or irrelevant.

Definition 3 (Theory Translation). *A* translation of theories *from $\langle \Sigma, \Gamma \rangle$ to $\langle \Sigma', \Gamma' \rangle$ is a function μ from Σ to $T[\Sigma']$ such that $\Gamma' \vdash \hat{\mu}(\Gamma)$. The translation μ is called* primitive *if $\mu(\Sigma) \subseteq \Sigma'$.*

The *composite* of two theory translations μ from Γ to Γ' and μ' from Γ' to Γ'' is the composite function $\widehat{\mu'} \circ \mu$ from Σ to $T[\Sigma'']$. It is straightforward to verify that $\widehat{\mu'} \circ \mu$ is indeed a theory translation from Γ to Γ''. In order to characterize the equivalence of theories, the notion of an invertible translation is too restrictive since, for instance, the empty theory over $\{a\}$ clearly should count as equivalent to the theory $\{b \equiv c\}$ over $\{b, c\}$ under any sensible definition of equivalence. We can overcome this problem by relaxing the notion of an invertible translation to that of a *quasi-invertible* one. Let $\imath_\Gamma$ be the *identity translation* on Γ that is given by the canonical inclusion of Σ into $T[\Sigma]$.

Definition 4 (Equivalence of Translations). *Let μ and ν be theory translations from $\langle \Sigma, \Gamma \rangle$ to $\langle \Sigma', \Gamma' \rangle$. Then μ is* equivalent *to ν, notation: $\mu \sim \nu$, iff $\Gamma' \vdash \mu(p) \equiv \nu(p)$ for every $p \in \Sigma$.*

Definition 5 (Quasi-Inverse). *Let μ be a theory translation from Γ to Γ' and let ν be a translation from Γ' to Γ. Then ν is* quasi-inverse *to μ iff $\nu \circ \mu \sim \imath_\Gamma$ and $\mu \circ \nu \sim \imath_{\Gamma'}$.*

We are now able to formulate an appropriate notion of equivalence between theories with different base vocabularies: Γ and Γ' are said to be *equivalent* if there is a quasi-invertible translation from Γ and Γ'; notation: $\Gamma \sim \Gamma'$. Quasi-invertible translations can be characterized by the following two conditions:

Definition 6 (Conservative Translation). *A translation μ from Γ to Γ' is* conservative *iff, for all statements α over Σ, $\Gamma' \vdash \hat{\mu}(\alpha)$ only if $\Gamma \vdash \alpha$.*

Definition 7 (Essentially Surjective Translation). *A translation μ from Γ to Γ' is* essentially surjective *iff for every $p \in \Sigma'$ there is a $\phi \in T[\Sigma]$ such that $\Gamma' \vdash p \equiv \hat{\mu}(\phi)$.*

Proposition 2. *A theory translation μ has a quasi-inverse if and only if μ is conservative and essentially surjective.*

Example 3. Let Γ be the theory over $\Sigma = \{a_0, a_1, \dots a_k\} \cup \{b_0, b_1, \dots b_k\}$, with k finite, that consists of the statements

$$a_n \equiv a_{n+1} \vee b_{n+1} \qquad \text{and} \qquad a_n \wedge b_n \equiv \Lambda \qquad (0 \leqslant n < k).$$

The information domain $C(\Gamma)$ of Γ consists of the sets $\varnothing$, $\{a_0, a_1, \dots, a_k\}$, and $\{a_0, a_1, \dots, a_{n-1}, b_n\}$ for all $n \leqslant k$. Since the elements of $C(\Gamma) \setminus \{\varnothing\}$ are pairwise incomparable with respect to set inclusion, it follows that $C(\Gamma)$ is *flat* and hence order-isomorphic to the information domain of the theory $\Gamma' = \{c_m \wedge c_n \equiv \Lambda \,|\, m \neq n\}$ over $\Sigma' = \{c_0, c_1, \dots, c_{k+1}\}$. Let μ be the function from Σ to $T[\Sigma']$ with

$$\mu(a_n) = c_{n+1} \vee \dots \vee c_{k+1} \qquad \text{and} \qquad \mu(b_n) = c_n \qquad (0 \leqslant n \leqslant k).$$

Clearly Γ' entails $\hat{\mu}(\Gamma)$, since $(c_{n+1} \vee \dots \vee c_{k+1}) \wedge c_n \equiv (c_{n+1} \wedge c_n) \vee \dots \vee (c_{k+1} \wedge c_n) \equiv \Lambda$ for all $n \leqslant k$. So μ is a translation from Γ to Γ'. Now consider the function ν from Σ' to $T[\Sigma]$ such that $\nu(c_{k+1}) = a_k$ and $\nu(c_n) = b_n$ for every $n \leqslant k$. We want to show that ν is a translation from Γ' to Γ. To this end, observe that Γ entails $b_m \preceq a_n$ and hence $b_m \wedge b_n \preceq \Lambda$ for all $m > n$. In addition, Γ entails $a_m \preceq a_n$ if $m \geqslant n$, from which it follows that $a_m \wedge b_n \equiv a_m \wedge a_n \wedge b_n \equiv \Lambda$ whenever $m \geqslant n$. All in all, this proves that Γ entails $\hat{\nu}(\Gamma')$. It remains to check that ν is quasi-inverse to μ. The only nontrivial claim is that Γ entails $\nu(\mu(a_n)) \equiv a_n$, i.e., that $\Gamma \vdash a_n \equiv b_{n+1} \vee \dots \vee b_k \vee a_k$, which follows easily by induction. Figure 3 illustrates the situation for $k = 1$ in terms of Lindenbaum algebras and information domains; in addition, the figure depicts the extents of the primitives. The shaded circles in the diagrams of $L(\Gamma)$ and $L(\Gamma')$ correspond to join-irreducible elements, which stand in an (order-reversing) one-to-one correspondence to the elements of $C(\Gamma)$ and $C(\Gamma')$. Clearly, the equivalence μ from Γ to Γ' induces isomorphisms between $L(\Gamma)$ and $L(\Gamma')$ and between $C(\Gamma)$ and $C(\Gamma')$.

Every theory translation μ from Γ to Γ' canonically gives rise to an algebra homomorphism $L(\mu)$ from $L(\Gamma)$ to $L(\Gamma')$ as well as to an order-preserving function $C(\mu)$ from $C(\Gamma')$ to $C(\Gamma)$ which preserves suprema of directed sets (cf. e.g. [11, Sect. 8.2]).

Proposition 3. *Suppose μ and ν are theory translations from Γ to Γ'. Then $L(\mu) = L(\nu)$ iff $\mu \sim \nu$. In particular, $L(\Gamma) \simeq L(\Gamma')$ iff $\Gamma \sim \Gamma'$.*

Proposition 4. *A translation μ of theories is conservative iff $C(\mu)$ is onto; μ is essentially surjective iff $C(\mu)$ is an order embedding.*

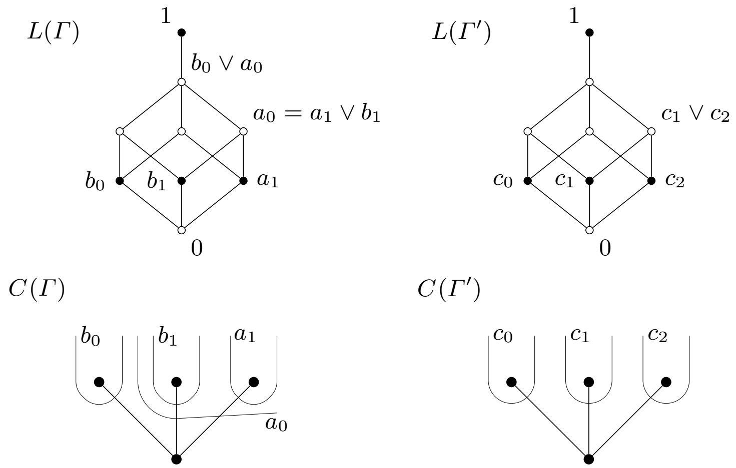

Fig. 3. Lindenbaum algebras and informations domains of Example 3 for $k = 1$

3.2 Example: Minimal Representations

Let P be a finite ordered set. Then P can be represented by a subset system over P, where $x \in P$ corresponds to $\downarrow x = \{y \in P \mid y \leqslant x\}$. Moreover, every subset system over a finite set Σ is the information domain of a theory over Σ. Of course, P may have a representation by a subset system over a set Σ with lower cardinality than P. For instance, if P is bounded and distributive, the set of join-irreducible elements of P can be chosen for Σ. Another type of example is given by the fact that a full binary tree P of height $k+1$, which has $2^{k+1} - 1$ nodes and 2^k leaves, can be embedded in $\wp(\Sigma)$ with $|\Sigma| = 2k$.

The *representation problem* arising here has the following general form: Given a finite ordered set P, find the least number n such that P can be represented as a subset system over a set of cardinality n. The solution of this problem is of practical importance because representations of ordered sets as subset systems can be easily implemented on a computer via *bit-vector* encoding. Unfortunately, the representation problem is NP-complete (see e.g. [3]).

In terms of theories and translations the representation problem can be rephrased as follows: Given a theory Γ over a finite set Σ, find a set Σ' with minimal cardinality such that Γ is equivalent to a theory Γ' over Σ'. Notice that if $|\Sigma'| = n$ then the width of $C(\Gamma')$ is at most $\binom{n}{\lfloor \frac{n}{2} \rfloor}$, by Sperner's Lemma (cf. e.g. [15]), which gives us a lower bound for the cardinality of Σ'.[4]

[4] The *width* of an ordered set P is the cardinality of a maximal antichain of P.

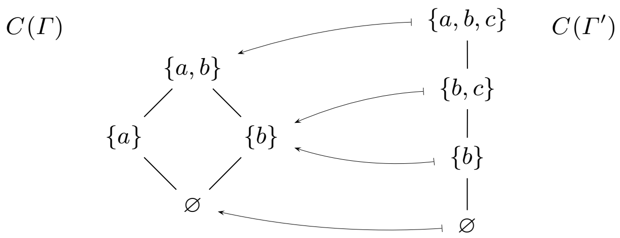

Fig. 4. Extension ε where $C(\varepsilon)$ is neither one-to-one nor onto

4 Theory Extensions

A translation of theories from $\langle\Sigma,\Gamma\rangle$ to $\langle\Sigma',\Gamma'\rangle$ is called an *extension* if its underlying function is an inclusion of Σ into Σ', and $\Gamma \subseteq \Gamma'$. We then say that $\langle\Sigma',\Gamma'\rangle$ is an *extension* of $\langle\Sigma,\Gamma\rangle$.

Proposition 5. *If ε is an extension of theories from $\langle\Sigma,\Gamma\rangle$ to $\langle\Sigma',\Gamma'\rangle$ then $C(\varepsilon)(Y) = Y \cap \Sigma$. In particular, $C(\varepsilon)$ is an order embedding if $\Sigma = \Sigma'$.*

Example 4. Let Γ be the empty theory over $\{a,b\}$ and let ε be its extension to Γ' by a single primitive c and the statements $a \preceq c$ and $c \preceq b$. The induced function $C(\varepsilon)$ from $C(\Gamma')$ to $C(\Gamma)$ is depicted by Figure 4. Notice that $C(\varepsilon)$ is neither one-to-one nor onto.

4.1 Conservative Extensions and Rule Extensions

Suppose Γ and Γ' are theories over Σ such that $\Gamma \subseteq \Gamma'$. Then Γ' is called a *rule extension* of Γ. The corresponding extension from Γ to Γ' is the identity function on Σ. A rule extension is trivially essentially surjective and the induced function of canonical universes is an inclusion, by Proposition 5; in particular, $C(\Gamma') \subseteq C(\Gamma)$. The definition of the canonical universe of a theory $\langle\Sigma,\Gamma\rangle$ given in Section 2.2 can be seen as an example of this fact: $C(\Gamma)$ is included in the powerset $\wp(\Sigma)$ of Σ, which is the canonical universe of the empty theory over Σ.

An extension $\langle\Sigma',\Gamma'\rangle$ of $\langle\Sigma,\Gamma\rangle$ is called *conservative* if the corresponding extension translation is conservative, that is, if $\Gamma' \vdash \alpha$ just in case $\Gamma \vdash \alpha$, for every statement α over Σ. In other words, conservative extensions do not entail additional statements over the base vocabulary. It follows by Propositions 5 and 4 that $C(\Gamma) = \{Y \cap \Sigma \mid Y \in C(\Gamma')\}$ in this case.

Proposition 6. *Let μ be a theory translation from Γ to Γ'.* (i) *If μ is conservative, Γ' is equivalent to a conservative extension of Γ.* (ii) *If μ is essentially surjective, Γ' is equivalent to a rule extension of Γ.*

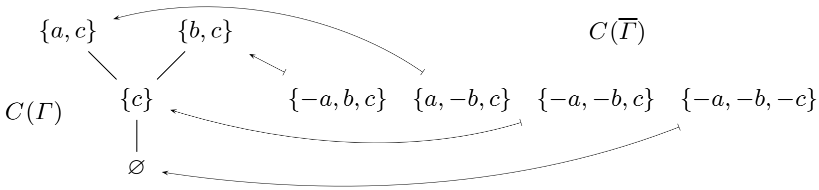

Fig. 5. Function of canonical universes induced by Booleanization

4.2 Booleanization

Let Γ be a theory over Σ. The *Booleanization* $\overline{\Gamma}$ of Γ is defined as follows: Σ is extended by a disjoint copy $\{-p \mid p \in \Sigma\}$ of Σ, and Γ is extended by all statements $p \wedge -p \preceq \Lambda$ and $V \preceq p \vee -p$, with $p \in \Sigma$, i.e., by all instantiations of the laws of contradiction and excluded middle for primitives.

Observe that no element of $C(\overline{\Gamma})$ is a proper subset of another element of $C(\overline{\Gamma})$, because each element of $C(\overline{\Gamma})$ contains either p or $-p$, but not both, for all $p \in \Sigma$.

Proposition 7. *The information domain of the Booleanization of a theory is an antichain.*

Let ε be the extension from Γ to $\overline{\Gamma}$. By Proposition 5, $C(\varepsilon)$ takes $Y \in C(\overline{\Gamma})$ to $Y \cap \Sigma$. Observe that $C(\varepsilon)$ is onto and one-to-one; in particular, Booleanization is conservative, by Proposition 4.

Example 5. Consider the theory $\Gamma = \{a \wedge b \preceq \Lambda, a \vee b \preceq c\}$ over $\Sigma = \{a, b, c\}$. Then $C(\Gamma)$ consists of $\varnothing$, $\{c\}$, $\{a, c\}$, and $\{b, c\}$, whereas $C(\overline{\Gamma})$ consists of $\{-a, -b, -c\}$, $\{-a, -b, c\}$, $\{a, -b, c\}$, and $\{-a, b, c\}$. Figure 5 depicts the induced function $C(\varepsilon)$ from $C(\overline{\Gamma})$ to $C(\Gamma)$.

The Lindenbaum algebra of $\overline{\Gamma}$ is easily seen to be *complemented*, i.e., $L(\overline{\Gamma})$ is a *Boolean lattice*. Since $C(\varepsilon)$ is onto, $L(\varepsilon)$ is an embedding of $L(\Gamma)$ into $L(\overline{\Gamma})$. It is not difficult to prove the following universal property of $L(\varepsilon)$: For every homomorphism h from $L(\Gamma)$ to a complemented algebra A, there exists a unique homomorphism h' from $L(\overline{\Gamma})$ to A such that $h = h' \circ L(\varepsilon)$. This universal characterization of $L(\overline{\Gamma})$ provides an additional justification of the term ‘Booleanization’. Moreover, the Boolean lattice $L(\overline{\Gamma})$ is isomorphic to the *Lindenbaum algebra* $B[\Sigma]/{\simeq_\Gamma}$ *of Boolean terms* of Γ.

4.3 Rule Completion

Besides Booleanization there is a another possible way to extend a theory $\langle \Sigma, \Gamma \rangle$ to a theory whose information domain is an antichain. The idea is to find an extension of Γ whose information domain is the set of maximal elements of the information domain of Γ. (Since every information domain is directed-complete, it has maximal elements by Zorn's Lemma.) Now observe that if such an extension of Γ exists at all, it can be realized by a rule extension, according to Propositions 4 and 6. We speak of a *rule completion* of Γ in this case. A possible rule completion of the theory Γ of Example 5 is given by $\Gamma' = \Gamma \cup \{ V \preceq a \vee b\}$. Then $C(\Gamma')$ consists of $\{a, c\}$ and $\{b, c\}$.

Rule completion, however, is not always possible. A simple counterexample is provided by the full binary exclusion theory $\Gamma = \{p \wedge q \preceq \Lambda \mid p \neq q\}$ over an *infinite* set Σ. Then $C(\Gamma) = \{\varnothing\} \cup \{\{p\} \mid p \in \Sigma\}$. It can be shown (see [13, p. 27]) that there is no theory Γ' over Σ such that $C(\Gamma') = \{\{p\} \mid p \in \Sigma\}$. Hence, there is no rule completion of Γ. Intuitively, what is needed here is the statement $V \preceq \bigvee \Sigma$, that is, an *infinite disjunction*.

4.4 Direct Sums

Consider the task of combining two theories $\langle \Sigma, \Gamma\rangle$ and $\langle \Sigma', \Gamma'\rangle$. Let us assume that Σ and Σ' and hence Γ and Γ' are disjoint. The most obvious way to combine the two theories into one is to take their *disjoint union*

$$\langle \Sigma, \Gamma\rangle \uplus \langle \Sigma', \Gamma'\rangle \;=\; \langle \Sigma \cup \Sigma', \Gamma \cup \Gamma'\rangle,$$

which is also known as their *direct sum.* Correspondingly, one can define the direct sum of an arbitrary family $\langle\langle \Sigma_i, \Gamma_i\rangle\rangle_{i\in I}$ of theories, given that the Σ_i's are pairwise disjoint:

$$\biguplus_{i\in I}\langle \Sigma_i, \Gamma_i\rangle \;=\; \langle \textstyle\bigcup_{i\in I} \Sigma_i, \bigcup_{i\in I} \Gamma_i\rangle.$$

It is not difficult to describe the canonical universe of $\biguplus_i \Gamma_i$ in terms of those of the Γ_i's: Let ε_i be the extension from $\langle \Sigma_i, \Gamma_i\rangle$ to the direct sum $\langle \bigcup_i \Sigma_i, \bigcup_i \Gamma_i\rangle$. According to Proposition 5, the function $C(\varepsilon_i)$ from $C(\biguplus_i \Gamma_i)$ to $C(\Gamma_i)$ takes X to $X \cap \Sigma_i$. Moreover, if $X_i \in C(\Gamma_i)$ for every i, then $\bigcup_i X_i \in C(\biguplus_i \Gamma_i)$; hence

$$C(\textstyle\biguplus_i \Gamma_i) \;=\; \{\bigcup_i X_i \mid X_i \in C(\Gamma_i)\}.$$

It follows that the information domain of $\biguplus_i \Gamma_i$ is order-isomorphic to the Cartesian product $\prod_i C(\Gamma_i)$ ordered by *coordinatewise inclusion*:

$$C(\textstyle\biguplus_i \Gamma_i) \;\simeq\; \prod_i C(\Gamma_i).$$

Example 6. Let Γ be a theory over Σ and let Σ' be a set disjoint to Σ. We call $\langle \Sigma \cup \Sigma', \Gamma\rangle$ the *extension of* $\langle \Sigma, \Gamma\rangle$ *by primitives* Σ'. Since $\langle \Sigma \cup \Sigma', \Gamma\rangle$ is identical to $\langle \Sigma, \Gamma\rangle \uplus \langle \Sigma', \varnothing\rangle$, it follows that $C(\langle \Sigma \cup \Sigma', \Gamma\rangle) \simeq C(\langle \Sigma, \Gamma\rangle) \times \wp(\Sigma')$. In particular, $|C(\langle \Sigma \cup \Sigma', \Gamma\rangle)| = |C(\langle \Sigma, \Gamma\rangle)| \cdot 2^{|\Sigma'|}$.

Example 7. Let Γ be a theory over Σ. Call a primitive $p \in \Sigma$ *free with respect to* Γ if p does not occur in any statement of Γ. Let $\sigma(\Gamma)$ be the set of primitives occurring in at least one of the statements of Γ, that is, $\Sigma' = \Sigma \setminus \sigma(\Gamma)$ is the set of free primitives. Then $\langle \Sigma, \Gamma\rangle$ is the direct sum of $\langle \sigma(\Gamma), \Gamma\rangle$ and $\langle \Sigma', \varnothing\rangle$ (and $\langle \Sigma, \Gamma\rangle$ is the extension of $\langle \sigma(\Gamma), \Gamma\rangle$ by primitives Σ'). Hence $C(\langle \Sigma, \Gamma\rangle) \simeq C(\langle \sigma(\Gamma), \Gamma\rangle) \times \wp(\Sigma')$.

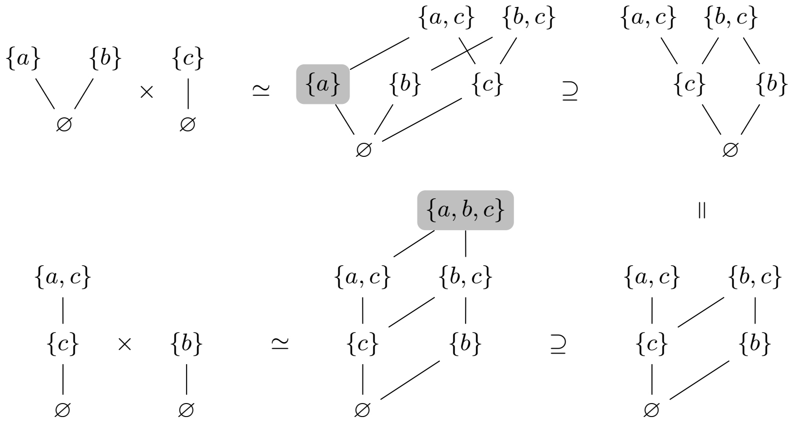

Fig. 6. Information domain of extension as subset of product

4.5 Extensions Decomposed

An extension of theories from $\langle \Sigma, \Gamma \rangle$ to $\langle \Sigma', \Gamma' \rangle$ can be decomposed into an extension of primitives from $\langle \Sigma, \Gamma \rangle$ to $\langle \Sigma', \Gamma \rangle$ followed by a rule extension from $\langle \Sigma', \Gamma \rangle$ to $\langle \Sigma', \Gamma' \rangle$. Recall from Example 6 that

$$C(\langle \Sigma', \Gamma \rangle) = \{X \cup Y \mid X \in C(\langle \Sigma, \Gamma \rangle),\, Y \subseteq \Sigma \setminus \Sigma'\} \simeq C(\langle \Sigma, \Gamma \rangle) \times \wp(\Sigma \setminus \Sigma').$$

Moreover, $C(\langle \Sigma', \Gamma' \rangle)$ consists of all elements of $C(\langle \Sigma', \Gamma \rangle)$ that are consistently closed with respect to $\Gamma' \setminus \Gamma$. So we can construct $C(\langle \Sigma', \Gamma' \rangle)$ from $C(\langle \Sigma, \Gamma \rangle)$ by taking first the product of $C(\langle \Sigma, \Gamma \rangle)$ and $\wp(\Sigma \setminus \Sigma')$ and then deleting those elements that are not consistently closed with respect to $\Gamma' \setminus \Gamma$.

Example 8. Let Γ' be the theory $\{a \wedge b \preceq \Lambda,\, a \preceq c\}$ over $\Sigma' = \{a, b, c\}$. Viewed as an extension of the theory $\{a \wedge b \preceq \Lambda\}$ over $\{a, b\}$, the construction of $C(\Gamma')$ by product and deletion is as depicted by the upper row of Figure 6. The lower row of the figure shows the construction if Γ' is viewed as an extension of the theory $\{a \preceq c\}$ over $\{a, c\}$. The shaded elements are subject to deletion because they are not consistently closed with respect to $\Gamma' \setminus \Gamma$.

4.6 A Simple Algorithmic Construction Scheme

Let Γ be a theory over a finite set Σ. Choose a strictly increasing sequence $\Sigma_0, \Sigma_1, \ldots, \Sigma_n$ of sets, with $\Sigma_0 = \varnothing$ and $\Sigma_n = \Sigma$, and let $\Gamma_i = \Gamma|_{\Sigma_i}$ be the set of all statements of Γ with primitives in Σ_i. Then $C(\Gamma)$ can be constructed from $C(\Gamma_0)$ by applying the two-step method described in Section 4.5 iteratively to the extension from Γ_{i-1} to Γ_i, for every i. As we have seen, $C(\Gamma_i)$ can be constructed from $C(\Gamma_{i-1})$ by taking all sets of the form $X \cup Y$, with $X \in C(\Gamma_{i-1})$ and $Y \subseteq \Sigma_i \setminus \Sigma_{i-1}$, such that $X \cup Y$ is consistently closed with respect to $\Gamma_i \setminus \Gamma_{i-1}$.

```
function  C (Σ: set;  Γ: theory):  system of sets;
begin
    if  not  cc? (∅, Γ|∅)  then
        C := ∅
    else begin
        C  := {∅};
        Σ' := ∅;
        while  Σ ≠ ∅  and  C ≠ ∅  do begin
            F  := any nonempty subset of Σ;
            Σ' := Σ' ∪ F;
            Γ' := Γ|Σ';
            C' := ∅;
            foreach  X ∈ C,  Y ⊆ F  do begin
                X' := X ∪ Y;
                if  cc? (X', Γ')  then  C' := C' ∪ {X'}
            end;
            Σ := Σ \ F;
            Γ := Γ \ Γ';
            C := C'
        end
    end
end;

function  cc? (X: set;  Γ: theory):  boolean;
{ true if X is consistently Γ-closed, false otherwise }
begin
    foreach  (φ ⪯ ψ) ∈ Γ  do
        if  X ⊨ φ  and  X ⊭ ψ  then  return (false);
    return (true)
end;
```

Fig. 7. A generic algorithmic scheme for information domain construction

An algorithmic formulation of this iteration scheme is shown in Figure 7, where the variables F and Σ' take the place of $\Sigma_i \setminus \Sigma_{i-1}$ and Σ_i, respectively. (Notice that the algorithm yields $C = \varnothing$ for inconsistent theories like $\{V \preceq a,\, a \preceq \Lambda\}$.) Calculating the information domain of the i-th extension requires to check $|C(\Gamma_{i-1})| \cdot 2^{k_i}$ sets against $|\Gamma_i \setminus \Gamma_{i-1}|$ statements, with $k_i = |\Sigma_i \setminus \Sigma_{i-1}|$. Consequently, $C(\Gamma_i)$ should be of low cardinality (more in the order of $|\Sigma_i|$ than of $2^{|\Sigma_i|}$) and k_i should be near to one. In other words, it is important to choose the partition of $\Gamma \setminus \Gamma_0$ into the sets $\Gamma_i \setminus \Gamma_{i-1}$ $(1 \leqslant i \leqslant n)$ in such a way that keeps $|C(\Gamma_i)|$ small during the construction process. How to do this in a systemic way is a topic for future research.

Notice that a nonredundant theory is not necessarily the best choice. For though less rules reduce the number of tests a single candidate set of primitives has to undergo (in the subroutine *cc*? of Figure 7), the total number of sets to be tested during an extension step can increase because additional (redundant) statements would have pruned $C(\Gamma_i)$ at an earlier stage of the construction.

Suppose Γ has reduced normal form. Then every statement of Γ can be represented by a *sequent*, i.e., a pair $\langle P, Q\rangle$ of disjoint finite subsets of Σ, where ϕ is the conjunction of the elements of P and ψ is the disjunction of the elements of Q (with V and Λ arising respectively by conjunction and disjunction of nothing). In terms of the sequent representation of Γ, the **if**-statement of the subroutine cc? becomes an elementary condition on sets:

$$\textbf{if}\quad P \subseteq X \quad \textbf{and} \quad X \cap Q = \varnothing \quad \textbf{then} \quad \textbf{return}\ (\text{false});$$

5 Conclusion

We have presented a framework for translating between classifications, taken as theories, over different base vocabulary and studied their effect on the associated information domains, which can be regarded as conceptual hierarchies. In particular, we have studied extensions of theories, with Booleanization as a special case. Moreover, it has been shown how theory extensions can be straightforwardly employed for constructing conceptual hierarchies in a step-by-step fashion.

Though the presentation has abstained from using the language of category theory, some readers surely will have noticed that categorical concepts are at least implicit in our treatment of theory translations. Readers interested in a more explicit categorical setting are invited to consult [11].

References

1. Garrett Birkhoff. *Lattice Theory.* AMS Colloquium Publications, Vol. 25. AMS, Providence, RI, 3rd edition, 1967.
2. Radim Bělohlávek and Vladimír Sklenář. Formal concept analysis constrained by attribute-dependency formulas. In Bernhard Ganter and Robert Godin, editors, *Proceedings of ICFCA 2005*, Lecture Notes in Artificial Intelligence 3403, pages 176–191, Berlin, 2005. Springer.
3. Yves Caseau, Michel Habib, Lhouari Nourine, and Olivier Raynoud. Encoding of multiple inheritance hierarchies and partial orders. *Computational Intelligence*, 15(1):50–62, 1999.
4. Brian A. Davey and Hilary A. Priestley. *Introduction to Lattices and Order.* Cambridge University Press, Cambridge, 2nd edition, 2002.
5. Manfred Droste and Rüdiger Göbel. Non-deterministic information systems and their domains. *Theoretical Computer Science*, 75:289–309, 1990.
6. Bernhard Ganter and Rudolf Wille. Contextual attribute logic. In William Tepfenhart and Walling Cyre, editors, *Proceedings of ICCS 1999*, Lecture Notes in Artificial Intelligence 1640, pages 377–388, Berlin, 1999. Springer.
7. Hele-Mai Haav. A semi-automatic method to ontology design by using FCA. In Vaclav Snášel and Radim Bělohlávek, editors, *Proceedings of CLA 2004*. VŠB – Technical University of Ostrava, Department of Computer Science, 2004.

8. Jørgen Fischer Nilsson. A logico-algebraic framework for ontologies. In P. Anker Jensen and P. Skadhauge, editors, *Ontology-based Interpretation of Noun Phrases, Proceedings of the 1st International OntoQuery Workshop*, pages 43–56, University of Southern Denmark – Kolding, 2001.
9. Nikolaj Oldager. Taxonomies with lattice algebras. In K. I. Simov and A. Kiryakov, editors, *Proceedings of the 1st International Workshop on Ontologies and Lexical Knowledge Bases (OntoLex)*, pages 106–117, 2000.
10. Frank J. Oles. An application of lattice theory to knowledge representation. *Theoretical Computer Science*, 249:163–196, 2000.
11. Rainer Osswald. A logic of classification – with applications to linguistic theory. Technical Report 303 – 9/2003, FernUniversität in Hagen, Department of Computer Science, 2003.
12. Rainer Osswald. Two views on building conceptual hierarchies. In Gabriele Kern-Isberner, Wilhelm Rödder, and Friedhelm Kulmann, editors, *Proceedings of the 2nd Workshop on Conditionals, Information, and Inference*, 2004. Ulm, September 21.
13. Rainer Osswald. Concept hierarchies from a logical point of view. In *Supplementary Proceedings of the 3rd International Conference on Formal Concept Analysis (ICFCA-2005)*, pages 15–30, Lens, France, 2005.
14. Rainer Osswald and Wiebke Petersen. A logical approach to data driven classification. In *KI-2003: Advances in Artificial Intelligence*, Lecture Notes in Artificial Intelligence 2821, pages 267–281, Berlin, 2003. Springer.
15. William T. Trotter. Partially ordered sets. In Ronald Graham, Martin Grötschel, and László Lovász, editors, *Handbook of Combinatorics*, pages 433–480. North-Holland, Amsterdam, 1995.