Information Algebra*

Jürg Kohlas and Cesar Schneuwly

University of Fribourg
Department of Informatics DIUF
Bd de Pérolles 90
1700 Fribourg, Switzerland
juerg.kohlas@unifr.ch, cesar.schneuwly@unifr.ch
http://diuf.unifr.ch/tcs/

1 Introduction

According to Shannon's classical information theory [19] information is measured by the reduction of uncertainty and the latter is measured by *entropy*. This theory is concerned with the transmission of symbols from a finite alphabet. The uncertainty concerns the question which symbol is sent and the information is given by a probabilistic model of the transmission channel and the symbol observed at the output of the channel. This leads to a statistical communication theory which is still the main subject of communication theory today.

There are some important elements in Shannon's approach that will be picked up and reconsidered here, although in another direction and with other goals than in Shannon's work. The first ingredient is that information relates to questions. In Shannon's case the question is fixed: what symbol is sent? In information processing in general several questions, whole systems of interrelated questions, will be considered. A piece of information may relate to a determined domain and must then be focussed on the question or questions of interest. Further, several pieces of information on related domains or questions may be available and must be aggregated to get the overall picture. These elements introduce an *algebraic flavor* into an extended information theory.

The theory proposed here can be sketched as follows: Questions can be represented by the possible answers they allow. There may be finer or coarser answers, which corresponds to a finer or coarser *granularity* of questions. This can be captured by a *partial order* between questions or the *domains* of possible answers. It will even be supposed that the system of questions or domains forms a *lattice*, such that two domains have a *supremum* or *join* representing the *combined question*, i.e. the possible answers to both questions. Two domains have also an *infimum* or *meet* representing the common part, the intersection, of both questions. Associated with this lattice of domains is a system of information consisting of pieces of information, each piece bearing on a determined domain from the lattice. Within this system the operations of *combination* of information, representing aggregation, and of *projection* to a given domain, representing

* Research supported by grant No. 200020–109510 of the Swiss National Foundation for Research.

G. Sommaruga (Ed.): Formal Theories of Information, LNCS 5363, pp. 95–127, 2009.

information extraction, are defined. This leads to a certain *two-sorted algebra* which is called an *information algebra* and which is the subject of this chapter.

First, in Section 2.1 the classical *relational algebra* associated with *relational databases* will be presented as a prototype of an information algebra. This serves as a motivation, since databases are surely depositories of information. In Section 2.2 the abstract axiomatic definition of information algebra is given. It is shown how information in this framework can be transported to arbitrary domains and thus relate to any question of the system considered. Further two equivalent variants of the algebraic structure are discussed: In Section 2.3 it is shown how focussing of information may in some cases be replaced by *variable elimination*. This positions information algebra in the context of *logics* and relates information extraction with *existential quantification*. The latter relation is elaborated in Section 3, especially in Subsections 3.2 and 3.3. The transport operation of information shows that the *same* piece of information may be represented equivalently relative to different domains. This leads to an equivalent *domain-free* version of the information algebra (Section 2.5). This variant may be better suited for some discussions than the original *labeled* version.

In Section 3 several instances or examples of information algebras are presented. They are mostly related to different systems of *logic* which provide besides databases a second basic form of representation of information. In particular the classical systems of *propositional* and *predicate* logic are presented as information algebras. This is clearly related to *algebraic theories* of logic as proposed for instance by [6, 7, 8]. The concept of *contexts* is proposed as a more general framework related to logic for obtaining information systems (Section 3.4). This concept is motivated by and related to *classifications* [1]. It is also connected to *concept analysis* [3]. Outside logic, a further example in Section 3.5 is linked to *fuzzy set theory* and *possibility theory*. These few examples should suffice to convince the reader about the justification and the interest of information algebras.

The last Section 4 establishes a first link of the theory of information algebras with Shannon's information measure, although it must be stressed that the algebraic theory so far is *not* a statistical theory. First we show how a natural *partial order* of information content arises from the algebra of information. It allows to compare information content both in an absolute way as well as with respect to a given question or domain. This order permits also to define particular algebras built form basic, finest information elements, called *atoms*. In those cases it will be possible to define a *quantitative information measure* using Hartley's measure (or entropy of uniform distributions) to quantify the *reduction of uncertainty* by an information element out of an information algebra. This measure is shown to respect the qualitative, partial order of information content. It is defined relative to any given domain, and there is also a *relative information measure* of a piece of information given another one. Several interesting properties of this measure are discussed. Again this is not a statistical theory of information, such that entropy displays not yet its full power. Motivated by relational algebra, dual information algebra and related measures in *Boolean information algebra* will also be considered in Section 4.2 and 4.4. Both measures have their proper interpretation and application.

Where do information algebras originate from? For Bayesian networks [14] proposed a so-called *local propagation algorithm*, which solved the dimensionality and efficiency problem of the naive solution problems. Based on this work [20] proposed a system of simple axioms which were sufficient for permitting local propagation and which were also sufficient for several formalisms of artificial intelligence. In [9] the algebraic theory of these, so-called *valuation algebras* was developed into some depth. And in particular, *information algebras* were proposed as valuation algebras which satisfy in addition the *idempotency property.* It is this property which allows the development of the information theory proposed here. So, whereas Shannon's theory is a theory of communication, resulting in efficient coding schemes, the theory of information algebra is a theory of *computation*, leading to efficient generic algorithms for important problems of *query processing.*

2 The Algebra of Information

2.1 A Prototype: Relational Algebra

Relational databases surely contain information. Therefore they may serve as a prototype example for the algebraic structure and theory we want to propose and discuss here. So let's summarize the basic elements of relational database theory.

Let $\mathcal{A}$ be a set of symbols, called *attributes.* For each $\alpha \in \mathcal{A}$ let D_α be a non-empty set, the set of possible values for attribute α. For example, if $\mathcal{A} = \{\texttt{name},\texttt{age},\texttt{income}\}$, then $D_{\texttt{name}}$ could be the set of strings, whereas $D_{\texttt{age}}$ and $D_{\texttt{income}}$ are both the set of nonnegative integers.

Let $x \subseteq \mathcal{A}$. A *x-tuple* is a function f with domain x and values $f(\alpha) \in D_\alpha$ for each $\alpha \in x$. The set of all x-tuples is denoted by E_x. For any x-tuple f and a subset $y \subseteq x$ the restriction $f[y]$ is defined to be the y-tuple g such that $g(\alpha) = f(\alpha)$ for all $\alpha \in y$.

A *relation* R over x is a set of x-tuples, i.e. a subset of E_x. The set of attributes x is called the *domain* of R and denoted by $d(R)$. For $y \subseteq d(R)$ the *projection* of R onto y is defined as follows:

$$\pi_y(R) = \{f[y] : f \in R\}.$$

The *join* of a relation R over x and a relation S over y is defined by

$$R \bowtie S = \{f : f \in E_{x \cup y}, f[x] \in R, f[y] \in S\}.$$

It is easy to see that the relations satisfy the following properties:

1. The join is an *associative* and *commutative* operation, and E_x is a *neutral* element for relations over x, i.e. $R \bowtie E_x = R$ if $d(R) = x$,
2. $d(R \bowtie S) = d(R) \cup d(S)$,

3. If $x \subseteq d(R)$, then $d(\pi_x(R)) = x$,
4. If $x \subseteq y \subseteq d(R)$, then $\pi_x(\pi_y(R)) = \pi_x(R)$,
5. If $d(R) = x$ and $d(S) = y$, then $\pi_x(R \bowtie S) = R \bowtie \pi_{x \cap y}(S)$,
6. If $x \subseteq y$, then $\pi_x(E_y) = E_x$,
7. If $x \subseteq d(R)$, then $R \bowtie \pi_x(R) = R$.

In fact, this algebraic system is part of *relational algebra* as defined in relational database theory [15]. Besides join and projection there are further operations like complement, union and difference. Relational algebra is used for query processing in relational databases. The operations of join and projection, and especially property (5) above, play a particularly important role [2]. There is even a special term for the formula $R \bowtie \pi_{x \cap y}(S)$ occurring in (5). It is called a *semijoin*.

We propose in the next section to abstract an algebraic system from this example, which we claim covers important aspects of a general theory of information.

2.2 The Axioms

Relations R as defined in the previous section can be thought of as representing *pieces of information* indicating which tuples $f \in E_{d(R)}$ describe possible tuples of values of the attributes $\alpha \in d(R)$. A relation R with domain $d(R) = x$ answers the *question*, which of the elements of the cartesian space

$$D_x = \times_{\alpha \in d(R)} D_\alpha \tag{1}$$

represent the true values of the variables. A relation is however only a partial answer since it does not fix a unique, precise element as an answer. So, any piece of information R refers to a determined domain $d(R)$, which in turn represents a *question* related to the attributes in x, asking what are the possible elements of D_x. Further the *join* serves to combine or aggregate two pieces of information, represented by two relations R and S. The combined information, represented by the join $R \bowtie S$ refers to domain $d(R) \cup d(S)$, according to property (2) in the previous section. *Projection* serves to extract the information relative to a part $y \subseteq d(R)$ of the domain of an information R. It results in an information relative to domain y, see property (3) in the previous section.

Thus, in a general way, we assume a set D of elements which are called *domains* and which are thought to represent in an abstract sense questions. Domains may have different *granularity*, i.e. a domain $x \in D$ may be *coarser* than another domain $y \in D$, meaning that y represents a more precise question than x. This is modelled by a *partial order* in D. Thus, $x \leq y$ means that x is a coarser domain than y, or that domain y is finer than x. Moreover, given two domains x and y, there should be a coarsest domain, finer than both x and y, i.e. the *join* $x \vee y$ should exist within D. It represents the *combined question* composed of questions x and y. In the same way a finest domain coarser than both x and y should exist, i.e. the *meet* $x \wedge y$ should exist within D. This means that D is assumed to be a *lattice* [3].

In relational algebra the domains are represented by subsets x of the attribute set $\mathcal{A}$. The partial order is defined by set inclusion, $x \leq y$ if $x \subseteq y$. Join and meet of domains correspond to set union and intersection, i.e. $x \vee y = x \cup y$ and $x \wedge y = x \cap y$. This is a distributive lattice [3]. In many applications we will use subsets of attributes or variables as domains. We call this *multivariate domains.*

Alternatively, but equivalently, we could consider the domains D_x defined in equation (1). Then we have $D_x \leq D_y$ if $x \subseteq y$ and $D_x \vee D_y = D_{x \cup y}$ and $D_x \wedge D_y = D_{x \cap y}$. A cartesian product D_x induces a *partition* of the universe $D_{\mathcal{A}}$. In fact, another, more general and interesting class of domain lattices are given by lattices of partitions of a universe S [4]. We remark that such partition lattices are in general no more distributive.

Further we consider a set Φ of elements, called *pieces of information* whose generic elements we denote by $\phi, \psi, \ldots$ etc. Each information ϕ concerns a certain domain $d(\phi) \in D$, which is attached to ϕ as a *label* or *mark*. The *combination* of information is defined by a binary operation $\Phi \times \Phi \rightarrow \Phi$, which will be denoted by $(\phi, \psi) \mapsto \phi \otimes \psi$. If x is a domain out of D and $\phi \in \Phi$ an information such that $x \leq d(\phi)$, then $\phi^{\downarrow x}$ denotes the part of information ϕ which concerns domain x. This operation of *projection* (sometimes also called *marginalization*) is defined as a partial mapping $\Phi \times D \rightarrow \Phi$.

Formally, we have thus a two-sorted algebra (Φ, D) with the following operations:

1. *Meet, Join:* $D \times D \rightarrow D$, $(x, y) \mapsto x \wedge y, x \vee y$,
2. *Combination:* $\Phi \times \Phi \rightarrow \Phi$, $(\phi, \psi) \mapsto \phi \otimes \psi$,
3. *Projection:* $\Phi \times D \rightarrow \Phi$, $(\phi, x) \mapsto \phi^{\downarrow x}$, defined for $x \leq d(\phi)$.

We impose the following axioms on this two-sorted algebra:

1. *Lattice:* D is a lattice with respect to the operations of meet and join.
2. *Semigroup:* Φ is associative and commutative under combination.
3. *Labeling:* $d(\phi \otimes \psi) = d(\phi) \vee d(\psi)$.
4. *Neutrality:* For all $x \in D$ there is a neutral element e_x such that $d(e_x) = x$ and for all $\phi \in \Phi$ with $d(\phi) = x$, $\phi \otimes e_x = \phi$; and for all $y \in D$, $x \geq y$, we have $e_x^{\downarrow y} = e_y$.
5. *Nullity:* For all $x \in D$ there is a null element z_x such that $d(z_x) = x$ and for all $\phi \in \Phi$ with $d(\phi) = x$, $\phi \otimes z_x = z_x$; and for all $y \in D$, $y \geq x$, we have $z_x \otimes e_y = z_y$.
6. *Projection:* If $\phi \in \Phi$, $x \in D$, $x \leq d(\phi)$, then $d(\phi^{\downarrow x}) = x$.
7. *Transitivity:* If $x \leq y \leq d(\phi)$, then $(\phi^{\downarrow y})^{\downarrow x} = \phi^{\downarrow x}$.
8. *Combination:* If $d(\phi) = x$, $d(\psi) = y$, then $(\phi \otimes \psi)^{\downarrow x} = \phi \otimes \psi^{\downarrow x \wedge y}$.
9. *Idempotency:* If $x \leq d(\phi)$, then $\phi \otimes \phi^{\downarrow x} = \phi$.

A two-sorted algebra (Φ, D) satisfying these axioms is called an *information algebra* [9]. That D is a lattice means that the operations of meet and join are both associative and commutative, idempotent (i.e. $a \wedge a = a \vee a = a$) and absorbing (i.e. $a \vee (a \wedge b) = a$ and $a \wedge (a \vee b) = a$). Axiom (2) says that Φ is a *commutative semigroup* under combination. The sequence of how pieces of information are combined does not matter. The labeling axiom (3) states that the combination of pieces of information relative to domains x and y relates to the combined question $x \vee y$. Axiom (4) establishes the existence of a neutral element, which represents *vacuous information*. It is stable, in the sense that projection vacuous information yields vacuous information. Similarly, axiom (5) establishes the existence of null elements, representing *contradiction*. Axiom (6) means that if the part relative to domain x is extracted from an information, then the resulting information relates to domain x. Transitivity (axiom (7)) says that projection can be done in steps. The combination axiom (8) tells us, that, in order to extract the part relative to domain x from a combined information on x and y, we can as well first extract the part relative to $x \wedge y$ from the information on y and then combine the two pieces of information. Finally, idempotency means that combining a piece of information with a part of it, gives nothing new. These seem reasonable properties to assume for an algebra of information. For relational algebra, these axioms correspond to the properties derived in the previous section. Relational algebra is thus an information algebra.

The next three assertions are immediate consequences of the axioms:

Lemma 1. *1. If $d(\phi) = x$, then $\phi^{\downarrow x} = \phi$.*
2. $\phi \otimes \phi = \phi$.
3. $e_x \otimes e_y = e_{x \vee y}$.

Proof. (1) Let $x = d(\phi)$. Then, by the combination and stability axioms, we have $\phi^{\downarrow x} = (\phi \otimes e_x)^{\downarrow x} = \phi \otimes e_x^{\downarrow x} = \phi \otimes e_x = \phi$.

(2) Using (1) and idempotency, we obtain $\phi \otimes \phi = \phi \otimes \phi^{\downarrow x} = \phi$.

(3) By the labeling axiom, stability and idempotency we conclude that $e_x \otimes e_y = e_x \otimes e_y \otimes e_{x \vee y} = e_{x \vee y}^{\downarrow x} \otimes e_{x \vee y}^{\downarrow y} \otimes e_{x \vee y} = e_{x \vee y}$. □

A central problem in applications can be formulated as follows: Given a number of pieces of information $\phi_1, \ldots, \phi_n$ with domains $d(\phi_i) = x_i$ and a goal domain x. The part relating to domain x of the total combined information is to be computed. Formally stated, we want to compute

$$(\phi_1 \otimes \cdots \otimes \phi_n)^{\downarrow x}.$$

This is the *projection problem*. If this is computed as written here, then by the labeling axiom, an information on the possibly very large domain $x_1 \vee \cdots \vee x_n$ has to be computed and then projected. This may be computationally infeasible. Instead, based in particular on the combination axiom, methods can be devised where ideally never information on larger domains than x_1 to x_n must be computed. These are called *local computation methods* [9, 12]. They were first proposed by [14] for *probabilistic networks*. Later [21] noted that these local computation methods can be used, if the elements satisfy some abstract axioms. The axioms of an information algebra are modelled after the Shenoy-Shafer system. In particular the idempotency axiom is added, which is not essential for local computation. But we shall see below that this axiom is essential for the theory of information presented here.

2.3 Variable Elimination

If we consider information algebras with *multivariate domains*, then an interesting variant of information algebras can be formed. Let V be a finite or countable set of variables, denoted by $X, Y, \ldots$ etc. Consider an information algebra (Φ, D), where D is the lattice of subsets of V. Using projection we define a new operation called *variable elimination* for $X \in d(\phi)$:

$$\phi^{-X} = \phi^{\downarrow d(\phi)-\{X\}}.$$

The following properties hold for variable elimination:

1. If $X \in d(\phi)$, then $d(\phi^{-X}) = d(\phi) - \{X\}$.
2. If $X, Y \in d(\phi)$, then $(\phi^{-X})^{-Y} = (\phi^{-Y})^{-X}$.
3. If $X \in d(\psi)$, $X \notin d(\phi)$, then $(\phi \otimes \psi)^{-X} = \phi \otimes \psi^{-X}$.
4. If $X \in d(\phi)$, then $\phi \otimes \phi^{-X} = \phi$.
5. If $X \subseteq z \in D$, then $e_z^{-X} = e_{z-\{X\}}$.

(1) follows immediately from the projection axiom, if the definition of variable elimination is used. Similarly, (2) follows directly from the transitivity axiom, (4) is the idempotency axiom and (5) follows from the neutrality axiom. Only (3) is a little bit more involved. We have $(\phi \otimes \psi)^{-X} = (\phi \otimes \psi)^{\downarrow(x \cup y)-\{X\}}$ if $d(\phi) = x$ and $d(\psi) = y$. Note that $x \subseteq z = (x \cup y) - \{X\} \subseteq x \cup y$. We claim that

$$(\phi \otimes \psi)^{\downarrow z} = \phi \otimes \psi^{\downarrow y \cap z}. \tag{2}$$

Since $y \cap z = y \cap ((x \cup y) - \{X\}) = y - \{X\}$ because $X \in y$ and $X \notin x$, we have then $\phi \otimes \psi^{\downarrow y \cap z} = \phi \otimes \psi^{-X}$ which proves (3). In order to prove equation (2) we note that $z \cap (x \cup y) = z$. The labeling and combination axioms permit then to derive

$$\begin{aligned}(\phi \otimes \psi)^{\downarrow z} &= (\phi \otimes \psi)^{\downarrow z} \otimes e_z \\ &= (\phi \otimes \psi \otimes e_z)^{\downarrow z} \\ &= (\phi \otimes e_z) \otimes \psi^{\downarrow y \cap z} \\ &= (\phi \otimes \psi^{\downarrow y \cap z}) \otimes e_z.\end{aligned}$$

The first term in this combination has domain $x \cup (y \cap z) = z$. This shows then that equation (2) holds indeed.

We may take properties (1) to (5) above for variable elimination as new axioms instead of axioms (4), (6), (7), (8) and (9) together with the remaining axioms (1), (2), (3) and (5). This gives a variant of an information algebra. In this system, property (2) above allows to define unambiguously the elimination of several variables $X_1, \ldots, X_n \in d(\phi)$ by

$$\phi^{-\{X_1,\ldots,X_n\}} = (\cdots((\phi^{-X_1})^{-X_2})\cdots)^{-X_n}.$$

According to property (2) the actual elimination sequences does not matter.

Variable elimination is only defined for *finite* sets of variables. Therefore, in general, it is less powerful than projection. If D in an information algebra (Φ, D)

is the lattice of *finite* subsets of a set of variables, then, for $x \in D$, projection may be defined in terms of variable elimination as follows:

$$\phi^{\downarrow x} = \phi^{-(d(\phi)-x)}.$$

It can easily be verified that, with this definition, the axioms of an information algebra are satisfied, if variable elimination satisfies properties (1) to (4) above. Thus, for multivariate systems with pieces of information always relating to finite sets of variables, the algebras with projection and variable elimination are *equivalent*.

2.4 Transport of Information

So far, information can only be projected to subdomains of its domain. However, transport of information from one domain to another one can be defined more generally. Let (Φ, D) be an information algebra. Then, for $y \geq d(\phi)$, we define a new operation

$$\phi^{\uparrow y} = \phi \otimes e_y,$$

called the *vacuous extension* of ϕ to domain y. This term is justified, since, for $d(\phi) = x$,

$$(\phi^{\uparrow y})^{\downarrow x} = (\phi \otimes e_y)^{\downarrow x} = \phi \otimes e_y^{\downarrow x} = \phi \otimes e_x = \phi$$

by the combination and stability axioms. So, vacuous extension indeed does not add or change otherwise information. Now, more generally, for $d(\phi) = x$ and $y \in D$ arbitrary, we define the operation

$$\phi^{\rightarrow y} = (\phi^{\uparrow x \vee y})^{\downarrow y}.$$

This is called the *transport operation*; it permits to transport a piece of information from its original domain to any other domain. Note that projection and vacuous extension are just special cases of this transport operation, namely for $y \leq d(\phi)$ or $y \geq d(\phi)$ respectively. Note further that $\phi^{\uparrow x \vee y} = \phi \otimes e_{x \vee y} = \phi \otimes e_y \otimes e_{x \vee y} = \phi \otimes e_y$, hence

$$\phi^{\rightarrow y} = (\phi \otimes e_y)^{\downarrow y} = \phi^{\downarrow x \wedge y} \otimes e_y = (\phi^{\downarrow x \wedge y})^{\uparrow y}.$$

In the following lemma we collect some properties of the transport operation.

Lemma 2. *1.* $(\phi^{\rightarrow y})^{\rightarrow z} = (\phi^{\rightarrow y \wedge z})^{\rightarrow z}$.
2. If $d(\phi) = x$, *then* $(\phi \otimes \psi)^{\rightarrow x} = \phi \otimes \psi^{\rightarrow x}$.
3. If $d(\phi) = x$, *then* $\phi^{\rightarrow x} = \phi$.
4. If $d(\phi) = x$, *then* $\phi \otimes \phi^{\rightarrow y} = \phi^{\uparrow x \vee y}$.

Proof. (1) If $y \leq z$, then we claim that $\phi^{\rightarrow y} = (\phi^{\rightarrow z})^{\rightarrow y}$. In fact, assume $d(\phi) = x$, then

$$\begin{aligned}(\phi^{\rightarrow z})^{\rightarrow y} &= ((\phi^{\uparrow x \vee z})^{\downarrow z})^{\downarrow y} = (\phi^{\uparrow x \vee z})^{\downarrow y} \\ &= (((\phi^{\uparrow x \vee y})^{\uparrow x \vee z})^{\downarrow x \vee y})^{\downarrow y} = (\phi^{\uparrow x \vee y})^{\downarrow y} = \phi^{\rightarrow y}.\end{aligned}$$

In order to prove (1) we apply this result and obtain

$$(\phi^{\to y \wedge z})^{\to z} = ((\phi^{\to y})^{\to y \wedge z})^{\to z} = ((\phi^{\to y})^{\downarrow y \wedge z})^{\uparrow z} = (\phi^{\to y})^{\to z}.$$

(2) This follows from the combination axiom

$$\begin{aligned}(\phi \otimes \psi)^{\to x} &= (\phi \otimes \psi)^{\downarrow x} = \phi \otimes \psi^{\downarrow x \wedge y} = \phi \otimes e_x \otimes \psi^{\downarrow x \wedge y} \\ &= \phi \otimes (\psi^{\downarrow x \wedge y})^{\uparrow x} = \phi \otimes \psi^{\to x}.\end{aligned}$$

(3) This follows since $\phi^{\to x} = \phi^{\downarrow x} = \phi$.
(4) Here we have, using the idempotency axiom,

$$\phi \otimes \phi^{\to y} = \phi \otimes (\phi^{\downarrow x \wedge y})^{\uparrow y} = \phi \otimes \phi^{\downarrow x \wedge y} \otimes e_y = \phi \otimes e_y = \phi^{\uparrow x \vee y}.$$ □

These properties of transport are similar to the transitivity, combination, projection and idempotency axioms of the information algebra. In fact, they could replace them.

2.5 Domain-Free Information Algebras

Assume that, in an information algebra (Φ, D), for two elements $\phi, \psi \in \Phi$ with domains $d(\phi) = x$ and $d(\psi) = y$ it holds that

$$\phi^{\to y} = \psi, \quad \psi^{\to x} = \phi. \tag{3}$$

Then ϕ and ψ represent in some sense the *same information*, in particular $\phi^{\downarrow x \wedge y} = \psi^{\downarrow x \wedge y}$ and $\phi^{\uparrow x \vee y} = \psi^{\uparrow x \vee y}$. We write $\phi \equiv \psi$ if (3) holds. This is clearly an *equivalence relation*. Moreover it is a *congruence* in the information algebra (Φ, D) in the following sense [9]: First $\phi_1 \equiv \phi_2$ and $\psi_1 \equiv \psi_2$ imply

$$\phi_1 \otimes \psi_1 \equiv \phi_2 \otimes \psi_2,$$

and secondly, also for any $z \in D$,

$$\phi_1^{\to z} \equiv \phi_2^{\to z}.$$

In fact in the last relation equality holds.

Let then $\Phi/\equiv$ denote the *equivalence classes* $[\phi]$ of this congruence in Φ. Then, in this quotient algebra the following two operations are well defined:

1. *Combination:* $[\phi] \otimes [\psi] = [\phi \otimes \psi]$.
2. *Focussing:* $[\phi]^{\Rightarrow x} = [\phi^{\to x}]$.

In the two-sorted algebra $(\Phi/\equiv, D)$ with the two operations just defined, the following properties hold:

Theorem 1. *Let $\Psi = \Phi/\equiv$ and denote generic elements of Ψ by $\psi, \eta, \ldots$ etc. Then*

1. *Semigroup: Ψ is associative and commutative under combination.*
2. *Support: If $\psi \in \Psi$, then there is a $x \in D$ such that $\psi = \psi^{\Rightarrow x}$.*

3. *Neutrality: There is a neutral element* e *such that* $\psi \otimes e = \psi$ *for all* $\psi \in \Psi$ *and* $e^{\Rightarrow x} = e$.
4. *Nullity: There is a null element* z *such that* $\psi \otimes z = z$ *for all* $\psi \in \Psi$ *and* $z^{\Rightarrow x} = z$.
5. *Transitivity: If* $\psi \in \Psi$ *and* $x, y \in D$, *then* $(\psi^{\Rightarrow x})^{\Rightarrow y} = \psi^{\Rightarrow x \wedge y}$.
6. *Combination: If* $\psi, \eta \in \Psi$ *and* $x \in D$, *then* $(\psi^{\Rightarrow x} \otimes \eta)^{\Rightarrow x} = \psi^{\Rightarrow x} \otimes \eta^{\Rightarrow x}$.
7. *Idempotency: If* $\psi \in \Psi$ *and* $x \in D$, *then* $\psi \otimes \psi^{\Rightarrow x} = \psi$.

Proof. (1) Associativity and commutativity of combination in Ψ is inherited from Φ.

(2) By Lemma 2 (3) we have $[\phi] = [\phi^{\rightarrow x}] = [\phi]^{\Rightarrow x}$ if $d(\phi) = x$.

(3) The equivalence class $[e_y]$ is the neutral element and $[e_y]^{\Rightarrow x} = [e_y^{\rightarrow x}] = [e_x]$ proves that the neutral element is stable under focussing.

(4) The equivalence class $[z_y]$ is the null element, and $[z_y]^{\Rightarrow x} = [z_y^{\rightarrow x}] = [z_x]$ proves the stability of the null element under focussing.

(5) By Lemma 2 (1) we have $(\phi^{\rightarrow x})^{\rightarrow y} = (\phi^{\rightarrow x \wedge y})^{\rightarrow y}$. Since $(\phi^{\rightarrow x \wedge y})^{\rightarrow y} \equiv \phi^{\rightarrow x \wedge y}$ we obtain $([\phi]^{\Rightarrow x})^{\Rightarrow y} = [(\phi^{\rightarrow x})^{\rightarrow y}] = [\phi^{\rightarrow x \wedge y}] = [\phi]^{\Rightarrow x \wedge y}$.

(6) Since $d(\phi^{\rightarrow x}) = x$, we obtain, using Lemma 2 (2)

$$\begin{aligned}([\phi]^{\Rightarrow x} \otimes [\psi])^{\Rightarrow x} &= [(\phi^{\rightarrow x} \otimes \psi)^{\rightarrow x}] = [\phi^{\rightarrow x} \otimes \psi^{\rightarrow x}] \\ &= [\phi]^{\Rightarrow x} \otimes [\psi]^{\Rightarrow x}.\end{aligned}$$

(7) This follows from Lemma 2 (4). In fact, if $d(\phi) = y$, then $[\phi] \otimes [\phi]^{\Rightarrow x} = [\phi \otimes \phi^{\rightarrow x}] = [\phi^{\uparrow x \vee y}] = [\phi]$. □

A two-sorted algebra (Ψ, D) with the operations of combination and focussing, satisfying the properties of Theorem 1, is called a *domain-free information algebra*. Theorem 1 says that any information algebra induces a domain-free information algebra. In order to distinguish the original algebra form the domain-free one, we call it a *labeled information algebra*.

In a domain-free information algebra (Ψ, D) a domain $x \in D$ is called a support of $\psi \in \Psi$, if $\psi = \psi^{\Rightarrow x}$. This means that no information is lost, when ϕ is focussed on domain x or, in other words, the whole information in ϕ is *carried* by domain x. According to the support property (2) in Theorem 1 any element of Ψ has a support. Here are a few properties of supports:

Lemma 3. *1.* x *is a support of* $\psi^{\Rightarrow x}$.
2. *If* x *and* y *are supports of* ψ, *then* $x \wedge y$ *is a support of* ψ.
3. *If* x *is a support of* ψ *and* $x \leq y$, *then* y *is a support of* ψ.
4. *If* x *is a support of* ψ, y *a support of* η, *then* $x \vee y$ *is a support of* $\psi \otimes \eta$.

Proof. (1) By transitivity (Theorem 1 (5)) we have $(\psi^{\Rightarrow x})^{\Rightarrow x} = \psi^{\Rightarrow x \wedge x} = \psi^{\Rightarrow x}$.

(2) Again, by (5) of Theorem 1, we obtain $\psi^{\Rightarrow x \wedge y} = (\psi^{\Rightarrow x})^{\Rightarrow y} = \psi^{\Rightarrow y} = \psi$.

(3) If $x \leq y$, then $x = x \wedge y$. So, once more by Theorem 1 (5), we conclude that $\psi^{\Rightarrow y} = (\psi^{\Rightarrow x})^{\Rightarrow y} = \psi^{\Rightarrow x \wedge y} = \psi^{\Rightarrow x} = \psi$.

(4) By (6) of Theorem 1, and (3) just proved, we see that $(\psi \otimes \eta)^{\Rightarrow x \vee y} = \psi^{\Rightarrow x \vee y} \otimes \eta^{\Rightarrow x \vee y} = \psi \otimes \eta$. □

If (Ψ, D) is a domain-free information, then define Ψ^* to be the set of all pairs (ψ, x), where $\psi \in \Psi$ and x is a support of ψ. We define then the following operations:

1. *Labeling:* $d(\psi, x) = x$.
2. *Combination:* $(\psi, x) \otimes (\eta, y) = (\psi \otimes \eta, x \vee y)$.
3. *Projection:* $(\psi, x)^{\downarrow y} = (\psi^{\Rightarrow y}, y)$ for $y \leq x$.

It is easy to verify that the two-sorted algebra (Ψ^*, D) with these operations forms a labeled information algebra. It has been shown elsewhere [9] that its domain-free version $(\Psi^*/\equiv, D)$ is then essentially identical to the original algebra (Ψ, D). Conversely, if $(\Psi, D) = (\Phi/\equiv, D)$ for a labeled algebra (Φ, D), then (Ψ^*, D) is essentially identical to (Φ, D) (in fact, *isomorph*, [9]). Thus labeled and domain-free algebras are different versions of the same structure. We may switch at our convenience between the two forms.

3 Some Examples

3.1 Propositional Logic

At the beginning we have shown that relational algebra is an example of a (labeled) information algebra. In this section we want to discuss further examples, especially systems related to *logic*. In the view proposed here, logic offers a *language* to describe information which refers to models or structures. We illustrate this first with *propositional logic* as a prototype case.

The vocabulary of propositional logic is formed by a countable set of variables $P = \{p_1, p_2, \ldots\}$, the constants $\bot, \top$ and the the connectors $\neg, \wedge$. Formulae of the language are:

1. Each element of P, $\bot$ and $\top$ are formulae (*atomic formulae*).
2. If f and g are formulae, then so are $\neg f$, $f \wedge g$.
3. All formulae are generated from atomic formulae by finitely often applying rule 2.

A valuation is a mapping $v : P \rightarrow \{\mathbf{f}, \mathbf{t}\}$ which assigns each propositional variable a truth value $\mathbf{f}$ (false) or $\mathbf{t}$ (true). A valuation assigns a truth value $\hat{v}(f)$ to any formula f by the following inductively defined process:

1. If f is a propositional variable, then $\hat{v}(f) = v(f)$.
2. $\hat{v}(\bot) = \mathbf{f}$ and $\hat{v}(\top) = \mathbf{t}$.
3. $\hat{v}(\neg f) = \begin{cases} \mathbf{f} \text{ if } \hat{v}(f) = \mathbf{t}, \\ \mathbf{t} \text{ if } \hat{v}(f) = \mathbf{f}. \end{cases}$
4. $\hat{v}(f \wedge g) = \begin{cases} \mathbf{t} \text{ if } \hat{v}(f) = \hat{v}(g) = \mathbf{t}, \\ \mathbf{f} \text{ otherwise.} \end{cases}$

A valuation v, under which a formula f evaluates to true, i.e for which $\hat{v}(f) = \mathbf{t}$, is said to *satisfy* the formula, or to be a *model* of the formula, which is denoted as $v \models f$. Let $M(f)$ be the set of all models of a propositional formula f. Since a valuation can also be seen as a sequence $v_1, v_2, \ldots$ of elements of $\{\mathbf{f}, \mathbf{t}\}$, the set of models $M(f)$ can be considered to be a subset of $\{\mathbf{f}, \mathbf{t}\}^\infty$.

In a given problem or context one may assume that there is some true but unknown truth assignment in the real world. The elements of $\{\mathbf{f}, \mathbf{t}\}^\infty$ are then *possible worlds*. A formula f of propositional logic can then be seen as an information about the unknown real world in that it postulates that the real world must be among its models $M(f)$. We are now going to associate an information algebra of models to propositional formulae.

Let D be the lattice of *finite subsets* of $\omega = \{1, 2, \ldots\}$. For any valuation $v \in \{\mathbf{f}, \mathbf{t}\}^\infty$ and any finite subset $x \in D$, we define $v^{\downarrow x}$ to be the x-tuple $v(i), i \in x$. We define an x-equivalence between two valuations v and w by $v \equiv_x w$ if $v^{\downarrow x} = w^{\downarrow x}$. The equivalence classes of this x-equivalence are denoted by $[v]_x$. For any subset A of $\{\mathbf{f}, \mathbf{t}\}^\infty$ let

$$A^{\Rightarrow x} = \bigcup_{v \in A} [v]_x.$$

A subset $\phi \subseteq \{\mathbf{f}, \mathbf{t}\}^\infty$ is called *cylindric* over x, if $\phi = \phi^{\Rightarrow x}$. Let then Φ_x be the family of x-cylindric subsets of $\{\mathbf{f}, \mathbf{t}\}^\infty$ and

$$\Phi = \bigcup_{x \in D} \Phi_x.$$

We claim then that (Φ, D) with intersection as combination $\otimes$ and the focussing operation $\Rightarrow$ defined above is a (domain-free) information algebra.

Let f be a propositional formula and $var(f)$ the set of propositional variables occurring in f. Then its set of models $M(f)$ belongs to $\Phi_{var(f)}$. So, any propositional formula f determines an element $\phi = M(f)$ of the information algebra Φ, its set of models $M(f)$ is the information it describes. Note that $M(f \wedge g) = M(f) \cap M(g)$, conjunction corresponds to combination. Focussing is more complicated. If g is a formula such that $M(g) = M(f)^{\Rightarrow x}$, then g is obtained form f by variable forgetting or existential quantification, we refer to [11] for more details on this algebra. Two formulae f and g are logically equivalent, if $M(f) = M(g)$. Equivalent formulae describe the same information. Below, in Subsection 3.4, it will also be shown to be an instance of a more general logic system related to information algebras.

3.2 Quantifier Algebras

If Φ is a *Boolean algebra* with minimal element $\perp$, then an *existential quantifier* is a mapping $\exists : \Phi \to \Phi$ subject to the following conditions:

1. $\exists \perp = \perp$,
2. $\phi \wedge \exists \phi = \phi$,
3. $\exists(\phi \wedge \exists \psi) = \exists \phi \wedge \exists \psi$.

More generally, let D be a lattice of subsets of some set I. Assume that there is an existential quantifier $\exists(J)$ for every subset $J \in D$ on the Boolean algebra Φ, and that

1. $\exists(\emptyset)\phi = \phi$,
2. if $J, K \in D$, then $\exists(J \cup K)\phi = \exists(J)(\exists(K)\phi)$.

Then (Φ, D) is termed a *quantifier algebra* over D. If we take the meet operation of the Boolean algebra for combination and define $\phi^{\Rightarrow I-J} = \exists(J)\phi$, then (Φ, D^c) is a *domain-free information algebra.* Here D^c is the lattice of subsets $I - J$ for *all* $J \in D$.

If, for all $i \in I$, $\exists(i)$ is a an existential quantifier, and $\exists(i)\exists(j) = \exists(j)\exists(i)$, if $i \neq j$, one can define $\exists(J) = \exists(i_1) \cdots \exists(i_m)$ if $J = \{1, \ldots, m\}$ and $\exists(\emptyset)(\phi) = \phi$. Then (Φ, D) is a quantifier algebra.

For instance, let Φ be the powerset of some cartesian product of a family of sets U_i for $i \in I$ and D a lattice of subsets of I. The mapping $\exists(J)$ is defined by

$$\exists(J)A = \{b \in \prod_{i \in I} U_i : \exists a \in A \text{ such that } b_i = a_i, \forall i \notin J\}.$$

It can be shown that this is an existential quantifier and (Φ, D) forms a quantifier algebra [17]. It is clear that the operation $\exists(\{i\})$ is similar to *variable elimination.* More generally, existential quantification is related to focussing, as we be seen in the next example (Section 3.3). Note also that it is sufficient for Φ to be a *semilattice* in order to define existential quantification and then a quantifier algebra (Φ, D).

3.3 Predicate Logic

Another information algebra is associated with predicate logic. The vocabulary of predicate logic consists of a countable set of variables $X_1, X_2, \ldots$ and a countable set of predicate symbols $P_1, P_2, \ldots$, the logical constants $\perp, \top$ and $\wedge, \neg, \exists$. Each predicate symbol has a definite rank $\rho = 0, 1, 2, \ldots$. We refer to a predicate with rank ρ as a ρ-place predicate. Formulae of predicate logic are built using the following rules:

1. $P_i X_{i_1} \ldots X_{i_\rho}$, where ρ is the rank of P_i, $\perp$ and $\top$ are (atomic) formulae.
2. If f is a formula, then $\neg f$ and $\exists X_i f$ are formulae.
3. If f and g are formulae, then $f \wedge g$ is a formula.

The predicate language $\mathcal{L}$ consists of all formulae which are obtained by applying a finite number of times these rules.

In order to define an interpretation of formulae of predicate logic, we choose a *relational structure* $\mathcal{R} = (U, R_1, R_2, \ldots)$ where U is a non-empty set, the *universe*, and R_i are relations among elements of U with the arity equal to the rank ρ of P_i, i.e. subsets of U^ρ. A *valuation* is a mapping $v : \omega \to U$, which assigns each variable X_i a value $v(i) \in U$ for $i \in \omega = \{1, 2, \ldots\}$. The set of valuations is U^ω, i.e. the set of sequences $v(1), v(2), \ldots$. We define for a valuation v and an index $i \in \omega$

$$v^{\Rightarrow i} = \{u \in U^\omega : u(j) = v(j) \text{ for } j \neq i\}.$$

Valuations are used to assign a *truth value* $\hat{v}(f)$ to each formula $f \in \mathcal{L}$. This truth assignment is defined inductively as follows:

1. $\hat{v}(\bot) = \mathbf{f}$, $\hat{v}(\top) = \mathbf{t}$
2. $\hat{v}(P_i X_{i_1} \dots X_{i_\rho}) = \mathbf{t}$, if $(v(i_1), \dots, v(i_\rho)) \in R_i$, and $\hat{v}(P_i X_{i_1} \dots X_{i_\rho}) = \mathbf{f}$ otherwise.
3. $\hat{v}(\neg f) = \mathbf{f}$, if $\hat{v}(f) = \mathbf{t}$, and $\hat{v}(\neg f) = \mathbf{t}$, if $\hat{v}(f) = \mathbf{f}$.
4. $\hat{v}(\exists X_i f) = \mathbf{t}$, if there is a valuation $u \in v^{\Rightarrow i}$ such that $\hat{u}(f) = \mathbf{t}$, and $\hat{v}(\exists X_i f) = \mathbf{f}$ otherwise.
5. $\hat{v}(f \wedge g) = \mathbf{t}$, if $\hat{v}(f) = \hat{v}(g) = \mathbf{t}$, and $\hat{v}(f \wedge g) = \mathbf{f}$ otherwise.

A valuation v is called a *model* of a formula f in the structure $\mathcal{R}$, if $\hat{v}(f) = \mathbf{t}$. We write then $v \models_{\mathcal{R}} f$. Given a structure $\mathcal{R}$, we assign finally to each formula $f \in \mathcal{L}$ the set of its models,

$$\hat{r}_{\mathcal{R}}(f) = \{v \in U^\omega : v \models_{\mathcal{R}} f\}.$$

We consider this set as the information relative to the unknown values of the variables $X_1, X_2, \dots$ expressed by the formula f. Let Φ be the family of all sets $\hat{r}_{\mathcal{R}}(f)$ for all $f \in \mathcal{L}$. If we define as usual $f \vee g = \neg(\neg f \wedge \neg g)$, then it is easy to see that

$$\begin{aligned} \hat{r}_{\mathcal{R}}(f \wedge g) &= \hat{r}_{\mathcal{R}}(f) \cap \hat{r}_{\mathcal{R}}(g), \\ \hat{r}_{\mathcal{R}}(f \vee g) &= \hat{r}_{\mathcal{R}}(f) \cup \hat{r}_{\mathcal{R}}(g), \\ \hat{r}_{\mathcal{R}}(\neg f) &= (\hat{r}_{\mathcal{R}}(f))^c. \end{aligned}$$

The family Φ is thus a *Boolean algebra*. Further we see that

$$\hat{r}_{\mathcal{R}}(\exists X_i f) = \bigcup_{v \in \hat{r}_{\mathcal{R}}(f)} v^{\Rightarrow i}.$$

We may denote the right hand side as $\exists(i)\hat{r}_{\mathcal{R}}(f)$. Clearly, for all $i \in \omega$ this is a quantifier on the Boolean algebra Φ in the sense of the previous example. Hence we may derive an existential quantifier $\exists(J)$ for any finite subset of ω. If D is the lattice of finite subsets of ω, then (Φ, D) is a *quantifier algebra* and so a domain-free information algebra. Combination is intersection, focussing is related to existential quantification, as explained in the previous example.

Two formulae f and g of predicate logic are said to be equivalent relative to the structure $\mathcal{R}$, written $f \equiv_{\mathcal{R}} g$, if $\hat{r}_{\mathcal{R}}(f) = \hat{r}_{\mathcal{R}}(g)$. So, equivalent formulae describe the same information. This induces an equivalence relation on $\mathcal{L}$. We may then introduce combination and existential quantification in $\mathcal{L}/\equiv_{\mathcal{R}}$ as follows: if $[f]_{\mathcal{R}}$ denotes the equivalence classes,

$$\begin{aligned} [f]_{\mathcal{R}} \otimes [g]_{\mathcal{R}} &= [f \wedge g]_{\mathcal{R}}, \\ \exists(J)[f]_{\mathcal{R}} &= [\exists(J)f]_{\mathcal{R}}, \end{aligned}$$

where $\exists(J)f = \exists X_{i_1}(\dots \exists X_{i_k}) \dots)$ if $J = \{i_1, \dots, i_k\}$. Then $(\mathcal{L}/\equiv_{\mathcal{R}}, D)$ inherits the properties of an information algebra from (Φ, D). So, the information algebra of structures is reflected in a corresponding information algebra of formulae. These algebras are reducts of *cylindric algebras* [8] or *polyadic* or also *Halmos algebras* [6, 17] introduced for the algebraic study of predicate logic.

3.4 Contexts

Here we consider a general system, which captures the two previous logic examples as well as many other logic and related systems. It is also closely related to the work of [1] on information flow. A *context* is a triple $(\mathcal{L}, \mathcal{M}, \models)$, where $\mathcal{L}$ can be thought of as a set of sentences, a language, $\mathcal{M}$ a set of structures or models and $\models \subseteq \mathcal{L} \times \mathcal{M}$ is a binary relation between sentences and models. This corresponds to *classifications* in [1], where the terms *types* and *tokens* are used instead of sentences and models. Finally, in formal concept analysis the elements are considered as *attributes* and *objects* [3]. We write $m \models s$ instead of $(s, m) \in \models$. The idea is of course that models m *satisfy* sentences s, and thus give some semantics to the language $\mathcal{L}$. An example is provided by *propositional logic*, where $\mathcal{L}$ is a propositional language, the elements of $\mathcal{M}$ are valuations, and $m \models s$ means that m satisfies s or m is a model of s. Similarly, *predicate logic*, together with a structure to interpret the formulae, provides another example of a context.

In a context a set of sentences $X \subseteq \mathcal{L}$ determines a set of possible models, namely the set of models satisfying all sentences of X,

$$\hat{r}(X) = \{m \in \mathcal{M} : \forall s \in X, m \models s\}.$$

If we define also similarly for a subset A of $\mathcal{M}$,

$$\check{r}(A) = \{s \in \mathcal{L} : \forall m \in A, m \models s\},$$

then $\check{r}(A)$ is the set of all sentences whose models contain A.

The following dual pairs of properties of these operators are well known [3]:

$$\begin{gathered}
X \subseteq \check{r}(\hat{r}(X)), \;\; A \subseteq \hat{r}(\check{r}(A)), \\
X \subseteq Y \Rightarrow \hat{r}(X) \supseteq \hat{r}(Y), \;\; A \subseteq B \Rightarrow \check{r}(A) \supseteq \check{r}(B), \\
\hat{r}(X) = \hat{r}(\check{r}(\hat{r}(X))), \;\; \check{r}(A) = \check{r}(\hat{r}(\check{r}(A))), \\
\hat{r}(\bigcup_{j \in J} X_j) = \bigcap_{j \in J} \hat{r}(X_j), \;\; \check{r}(\bigcup_{j \in J} A_j) = \bigcap_{j \in J} \check{r}(A_j).
\end{gathered}$$

We define further for $X \subseteq \mathcal{L}$ and $A \subseteq \mathcal{M}$,

$$C_{\models}(X) = \check{r}(\hat{r}(X)), \quad C^{\models}(A) = \hat{r}(\check{r}(A)).$$

It follows from the properties above that $C_{\models}$ and $C^{\models}$ are *closure* or *consequence operators*, i.e.

1. $X \subseteq C_{\models}(X)$,
2. $C_{\models}(C_{\models}(X)) = C_{\models}(X)$,
3. If $X \subseteq Y$, then $C_{\models}(X) \subseteq C_{\models}(Y)$,

and similarly for $C^{\models}$. Sets $X \subseteq \mathcal{L}$ and $A \subseteq \mathcal{M}$ are called $\models$-*closed* if $X = C_{\models}(X)$ or $A = C^{\models}(A)$ respectively. We obtain then

$$\hat{r}(X) = C^{\models}(\hat{r}(X)), \quad \check{r}(A) = C_{\models}(\check{r}(A)).$$

So, any set of sentences X determines a $\models$-closed set $\hat{r}(X)$ of models as information. In the same way, any set of models A determines a $\models$-closed set $\check{r}(A)$ of sentences, which could be called the *theory* of A. In particular, $\models$-closed sets of models and theories are in a one-to-one relation, i.e. if $A = \hat{r}(X)$ and $X = \check{r}(A)$, then both A and X must be $\models$-closed.

In the case of *propositional logic*, $C_{\models}(X)$ is the set of all *logical consequences* of X or the *theory* of X. In this case, as in predicate logic and in many other cases, all subsets of models are closed. This is not the case in the following example: Let $X_i, i \in \omega = \{1, 2, \ldots\}$, be a countable family of variables, $\mathcal{F}$ a field and let $\mathcal{L}$ be the family of *linear equations* of the form

$$\sum_{i \in I} a_i X_i = a_0, \; I \text{ a finite subset of } \omega \text{ and } a_0, a_i \in \mathcal{F}.$$

Further, let $\mathcal{M} = \mathcal{F}^\omega$. Define $m \models s$, for $m \in \mathcal{M}$ and $s \in \mathcal{L}$, if m satisfies the linear equations s, i.e. if

$$\sum_{i \in I} a_i m_i = a_0.$$

Then, for a subset X of $\mathcal{L}$ the closed set $\hat{r}(X)$ is the linear solution manifold of the system of equations X in $\mathcal{M}$. So here, $\models$-closed sets are linear manifolds, and $C^{\models}(A)$ is the linear manifold spanned by $A \subseteq \mathcal{M}$. If *linear inequalities* in an ordered field, instead of linear equations are considered, then, in the same way, the $\models$-closed sets are *convex polyhedra.*

Consider the set of all $\models$-closed subsets of $\mathcal{M}$. For two elements $\phi = \hat{r}(X)$ and $\psi = \hat{r}(Y)$ we define then a combination operation

$$\phi \otimes \psi = \hat{r}(X \cup Y) = \hat{r}(X) \cap \hat{r}(Y) = \phi \cap \psi. \tag{4}$$

In fact, this operation could be defined for arbitrary families of sets $X_i \in \mathcal{L}$. So, information is combined either by the union of the sentences which define the information or by intersection of their model sets.

If we want to extend this semigroup to an information algebra, we must add a *domain structure* and a corresponding *focussing operation.* Let D be a *lattice* and, for any $x \in D$, let $\equiv_x$ be an *equivalence relation* in $\mathcal{M}$ such that

$$x \leq y \Rightarrow \equiv_x \supseteq \equiv_y . \tag{5}$$

A triple $(\mathcal{M}, D, \equiv_{x \in D})$, where D is a lattice and $\equiv_x$ are equivalence relations in $\mathcal{M}$ satisfying the condition above, is called a *similarity model structure* in [23]. For any model $m \in \mathcal{M}$ and $x \in D$, define

$$m^{\Rightarrow x} = \{n \in \mathcal{M} : n \equiv_x m\}.$$

Further, for a subset A of $\mathcal{M}$ let

$$A^{\Rightarrow x} = \bigcup_{m \in A} m^{\Rightarrow x}. \tag{6}$$

A set of models A such that $A = A^{\Rightarrow x}$ is called *cylindric* over x or x-closed. We require now two additional conditions:

1. *Closure:* If A is $\models$-closed and $x \in D$, then $A^{\Rightarrow x}$ is $\models$-closed.
2. *Independence:* If for two models m, n it holds that $m \equiv_{x \wedge y} n$, then there is a model l such that $l \equiv_x m$ and $l \equiv_y n$.

The closure property implies that the family of $\models$-closed sets is closed under the *focussing* operation $\Rightarrow$. The independence property guarantees that indeed a (domain-free) *information algebra* can be associated with contexts, as we shall show below.

In propositional and predicate logic, as in many other cases, the language $\mathcal{L}$ is defined over a set of *variables* $X_1, X_2, \ldots$ with domains $U_1, U_2, \ldots$. Models are valuations $v(i) \in U_i$. A similarity model structure is then defined for instance for finite subsets x of variables by $v \equiv_x u$ if $v(i) = u(i)$ for all $i \in x$. It can be verified that this structure satisfies the closure and independence properties above. This corresponds essentially to a *multivariate domain.*

Let Φ be the set of all cylindric sets which are $\models$-closed. The following lemma collects two important properties of cylindric, $\models$-closed sets.

Lemma 4. *For $x, y \in D$ and $\phi, \psi \in \Phi$, the following holds:*

1. *If ϕ is x-closed and ψ is y-closed, then $\phi \otimes \psi$ is $x \vee y$-closed.*
2. *If ϕ is x-closed, then $\phi^{\Rightarrow y}$ is $x \wedge y$-closed.*

Proof. (1) We claim that if $x \leq y$, then ϕ is x-closed implies ϕ is y-closed. In fact, suppose $\phi = \phi^{\Rightarrow x} = \cup_{n \in \phi} n^{\Rightarrow x}$. Then

$$\begin{aligned}\phi^{\Rightarrow y} &= \bigcup_{m \in \phi} m^{\Rightarrow y} = \bigcup_{n \in \phi} \bigcup_{m \in n^{\Rightarrow x}} m^{\Rightarrow y} \\ &= \bigcup_{n \in \phi} n^{\Rightarrow x} = \phi^{\Rightarrow x} = \phi.\end{aligned}$$

Therefore, if ϕ and ψ are x- and y-closed respectively, both are $x \vee y$-closed and so is $\phi \otimes \psi = \phi \cap \psi$.

(2) We claim that $(m^{\Rightarrow x})^{\Rightarrow y} = m^{\Rightarrow x \wedge y}$, which then implies property 2 immediately. In fact, if $n \equiv_{x \wedge y} m$, then, by the independence property above, there is an l such that $l \in m^{\Rightarrow x}$ and $n \in l^{\Rightarrow y}$. But this means that $n \in (m^{\Rightarrow x})^{\Rightarrow y}$. Conversely, if $n \in (m^{\Rightarrow x})^{\Rightarrow y}$, then there is a l such that $n \equiv_y l \equiv_x m$. By the monotonicity property (5) it follows that $n \equiv_{x \wedge y} l \equiv_{x \wedge y} m$, hence $n \in m^{\Rightarrow x \wedge y}$. $\square$

After this preparation it can be shown that (Φ, D) forms an information algebra.

Theorem 2. *The two-sorted algebra (Φ, D) with combination $\otimes$ and focussing $\Rightarrow$ defined above by (4) and (6) respectively, is a domain-free information algebra if the closure and independence properties are satisfied.*

Proof. We verify properties (1) to (7) of Theorem 1 above. The semigroup properties holds for intersection, hence for combination and $\mathcal{M}$ is the neutral element of combination, whereas the empty set is the null element. Transitivity follows from property 2 of Lemma 4 above, because $\phi^{\Rightarrow x}$ is x-closed. Combination is verified as follows:

$$\begin{aligned}(\phi^{\Rightarrow x} \otimes \psi)^{\Rightarrow x} &= \bigcup_{m \in (\phi^{\Rightarrow x} \cap \psi)} m^{\Rightarrow x} \\ &= \left(\bigcup_{m \in \phi^{\Rightarrow x}} m^{\Rightarrow x} \right) \cap \left(\bigcup_{m \in \psi} m^{\Rightarrow x} \right) \\ &= \phi^{\Rightarrow x} \otimes \psi^{\Rightarrow x}.\end{aligned}$$

The support axiom holds since the elements of Φ are cylindric, and the idempotency axioms is evident. □

We have represented the information algebra associated with a context in terms of models. But we could also represent it in terms of theories. If $\phi = \hat{r}(X)$ and $\psi = \hat{r}(Y)$, then we could consider the associated theories $C_{\models}(X)$ and $C_{\models}(Y)$ and define combination by

$$\begin{aligned}C_{\models}(X) \otimes C_{\models}(Y) &= \check{r}(\phi \otimes \psi) \\ &= C_{\models}(X \cup Y) \\ &= C_{\models}(C_{\models}(X) \cup C_{\models}(Y)).\end{aligned}$$

Further focussing could be defined as follows:

$$C_{\models}(X)^{\Rightarrow x} = \check{r}(\hat{r}(X)^{\Rightarrow x}).$$

This gives then a domain-free information algebra of theories associated to the algebra of models. Predicate logic provides an example with the algebra of structures and the algebra of formulae.

For any $x \in D$ we define

$$\mathcal{M}_x = \{m^{\Rightarrow x} : m \in \mathcal{M}\}, \quad \mathcal{L}_x = \{s \in \mathcal{L} : \hat{r}(\{s\}) = (\hat{r}(\{s\}))^{\Rightarrow x}\}.$$

Furthermore, we define a relation $\models_x$ between $\mathcal{M}_x$ and $\mathcal{L}_x$ by $m^{\Rightarrow x} \models_x s$ if $m \models s$ for all $m \in m^{\Rightarrow x}$. Then $(\mathcal{L}_x, \mathcal{M}_x, \models_x)$ is a context. Note that cylindric sets $A^{\Rightarrow x}$ over x can, in a natural way, also be considered as a subset of $\mathcal{M}_x$, namely the set consisting of elements $m^{\Rightarrow x}$ for all $m \in A$. Further, by the closure property, if A is $\models$-closed, then $A^{\Rightarrow x}$ is $\models_x$-closed.

Consider two elements $x, y \in D$ such that $x \leq y$. Then it follows from (5) that $m^{\Rightarrow x} \supseteq m^{\Rightarrow y}$ and $\mathcal{L}_x \subseteq \mathcal{L}_y$. We define now a contravariant pair of mappings

$$\begin{aligned}g : \mathcal{M}_y &\rightarrow \mathcal{M}_x \\ \mathcal{L}_y &\leftarrow \mathcal{L}_x : f\end{aligned}$$

by $g(m^{\Rightarrow y}) = m^{\Rightarrow x}$, and $f(s) = s$. It can be verified, that this pair of mappings satisfies the following condition

$$g(m^{\Rightarrow y}) \models_x s \Leftrightarrow m^{\Rightarrow y} \models_y f(s).$$

A contravariant pair of mappings between two contexts $(\mathcal{L}_x, \mathcal{M}_x, \models_x)$ and $(\mathcal{L}_y, \mathcal{M}_y, \models_y)$ satisfying this condition is termed a *context morphism*. It corresponds in essence to the *infomorphism* introduced in [1]. If we consider contexts $(\mathcal{L}_x, \mathcal{M}_x, \models_x)$ and $(\mathcal{L}_y, \mathcal{M}_y, \models_y)$ together with the context $(\mathcal{L}_{x\vee y}, \mathcal{M}_{x\vee y}, \models_{x\vee y})$, and add the context morphisms between the first two contexts and the third one, then we have what is called a *channel* in [1]. In fact, it allows to *transport information* (in the sense discussed in Section 2.4) from $(\mathcal{L}_x, \mathcal{M}_x, \models_x)$ to $(\mathcal{L}_y, \mathcal{M}_y, \models_y)$ and vice versa.

3.5 Lattice Induced Algebras

Let A be a *distributive, complete lattice* with supremum (join) and infimum (meet) denoted as usual by $\vee$ and $\wedge$. Let further r denote a finite set of variables $X_1, X_2, \ldots$ and U_i the domain of variable X_i. To a set $s \subseteq r$ of variables the cartesian product

$$U_s = \prod_{i \in r} U_i$$

is assigned as domain. The elements of U_s are tuples with domain s. We adopt the convention that the domain of the empty set of variable $U_\emptyset$ consists of a single tuple, denoted by $\diamond$. We use lower case, bold-face letters such as $\mathbf{x}, \mathbf{y}, \ldots$ to denote tuples. In order to emphasize the decomposition of a tuple $\mathbf{x}$ with domain s into components belonging to two disjoint subsets t and $s - t$ of s, we write $\mathbf{x} = (\mathbf{x}^{\downarrow t}, \mathbf{x}^{\downarrow s-t})$. A *valuation* ϕ with domain s is a mapping $\phi : U_s \rightarrow A$. The domain of a valuation ϕ is denoted by $d(\phi)$. The set of all valuations with domain s is denoted by Φ_s. Let then

$$\Phi = \bigcup_{s \subseteq r} \Phi_s.$$

Further let D be the lattice of subsets of r. We now use the lattice operations in A to define two operations in the pair (Φ, D):

1. *Combination:* $\otimes : \Phi \times \Phi \rightarrow \Phi$ defined for $\mathbf{x} \in U_{d(\phi) \cup d(\psi)}$ by
$$\phi \otimes \psi(\mathbf{x}) = \phi(\mathbf{x}^{\downarrow d(\phi)}) \wedge \psi(\mathbf{x}^{\downarrow d(\psi)}).$$
2. *Projection:* $\downarrow : \Phi \times D \rightarrow \Phi$ defined for all $\phi \in \Phi$ and $t \subseteq d(\phi)$ for $\mathbf{x} \in U_t$ by
$$\phi^{\downarrow t}(\mathbf{x}) = \bigvee_{\mathbf{z} \in U_{d(\phi)} : \mathbf{z}^{\downarrow t} = \mathbf{x}} \phi(\mathbf{z}).$$

It has been shown elsewhere that (Φ, D) with the two operations defined above is a (labeled) information algebra [13]. Examples for the lattice A include the Boolean lattice $\{0, 1\}$ (in which case valuations describe *constraints* or *subsets*), or the interval $[0, 1]$ with $\max, \min$ as lattice operations. This is used in fuzzy set theory. More general distributive lattices can be used to express qualitative membership of elements to *fuzzy sets*.

4 Order and Measure of Information

4.1 Partial Orders of Information Content

In an information algebra, the elements may be ordered by information content. The idea is that a piece of information is more informative than another one, if their combination yields the first. More precisely, let (Φ, D) be a *domain-free information algebra*. Then, for $\phi, \psi \in \Phi$ we define $\phi \geq \psi$ if $\phi \otimes \psi = \phi$, i.e. ϕ is *more informative* than ψ, if combining the latter with the first one gives nothing new. It can easily be verified that this relation is a *partial order* in Φ. Here are are a few elementary properties of this order which are proven in [9]:

1. $e \leq \phi$,
2. $\phi, \psi \leq \phi \otimes \psi$,
3. $\phi^{\Rightarrow x} \leq \phi$,
4. $\phi \leq \psi$ implies $\phi^{\Rightarrow x} \leq \psi^{\Rightarrow x}$,
5. $\phi \leq \psi$ implies $\phi \otimes \eta \leq \psi \otimes \eta$,
6. $\phi^{\Rightarrow x} \otimes \psi^{\Rightarrow x} \leq (\phi \otimes \psi)^{\Rightarrow x}$,
7. $x \leq y$ implies $\phi^{\Rightarrow x} \leq \psi^{\Rightarrow y}$.

In particular, it can also be verified that $\phi \otimes \psi = \sup\{\phi, \psi\}$. Therefore, Φ is also a *semilattice* and we write sometimes $\phi \otimes \psi = \phi \vee \psi$, if we want to stress order-theoretic issues.

This order reflects the absolute information content of the elements of Φ. It is also interesting to compare the information contents of the elements of Φ with respect to a determined question, i.e. a given domain $x \in D$. For this purpose we define $\phi \leq_x \psi$, if $\phi^{\Rightarrow x} \leq \psi^{\Rightarrow x}$. So, ϕ is less informative than ψ, *relative to a domain* x, if its part relating to x is less informative than the part of ψ relating to x. The relation $\leq_x$ is a *preorder* (reflexive and transitive, but not antisymmetric) on Φ. This is equivalent to a similar order defined on *labeled information algebras*, where again $\phi \geq \psi$ if $\phi \otimes \psi = \phi$. Then, for $x \in D$, we define $\phi \leq_x \psi$, if $\phi^{\rightarrow x} \leq \psi^{\rightarrow x}$.

In the case of *propositional logic*, a propositional formula f is more informative than a formula g, if $M(f) \subseteq M(g)$, since combination is intersection. This means that f is more informative than g, if, and only if, the latter is a *logical consequence* of the former, i.e. if $f \models g$. Similarly in predicate logic, a predicate formula f is more informative than g, relative to a structure $\mathcal{R}$, if $\hat{r}_{\mathcal{R}}(f) \subseteq \hat{r}_{\mathcal{R}}(g)$, i.e. again if g is a logical consequence of f, i.e. $f \models_{\mathcal{R}} g$. In a lattice-induced information algebra, a valuation v with domain x is more informative than another valuation u with the same domain, if $v(\mathbf{x}) \leq u(\mathbf{x})$ for all $\mathbf{x} \in U_x$. This is a kind of fuzzy subset relation, generalizing the ordinary subset relation.

These partial orders describe qualitative comparisons of information content between pieces of information. We may also try to measure quantitatively the content of an information. This is discussed below in Section 4.4.

4.2 Boolean Information Algebras

In the case of a relational algebra, for two relations R and S with the same domain x, R is more informative than S, if $R \subseteq S$. This makes sense in many cases: For instance if somebody is expected in Zurich on a flight from London, and information on possible flights is given by a list of flights from London to Zurich, then the smaller the list, the more information is obtained, the less uncertainty remains. If however we look for a flight we could take from London to Zurich, then obviously we feel to dispose of more information the longer the list of possible flights we obtain. So, information content of a relation seems to depend on the question we are interested in. This is related to the Boolean nature of the relational algebra. In order to elucidate this issue we introduce Boolean information algebra in this section.

Let (Φ, D) be a domain-free information algebra, such that in particular Φ is a semilattice relative to the partial order on information content. We assume now in addition that Φ is not only a semilattice but a *Boolean algebra*. This means that Φ has a bottom and a top element e and z and is a *distributive lattice*, where not only the supremum $\phi \vee \psi$ exists relative to the order, but also the infimum $\phi \wedge \psi$ and the distributive laws hold between these two operations. Further there is a complement ϕ^c for each element $\phi \in \Phi$ such that $\phi \wedge \phi^c = e$ and $\phi \vee \phi^c = z$. Then (Φ, D) is called a *Boolean information algebra.* For instance the information algebras associated with propositional and predicate logic are Boolean.

In a Boolean algebra there exists a well known *duality* which carries over to Boolean information algebras. If (Φ, D) is a Boolean information algebra, then we define the following *dual operations* of combination and focussing:

1. *Dual Combination:* $\phi \otimes_d \psi = (\phi^c \otimes \psi^c)^c$,
2. *Dual Focussing:* $\phi^{\Rightarrow_d x} = ((\phi^c)^{\Rightarrow x})^c$.

Note that by de Morgan's law $\phi \otimes_d \psi = \phi \wedge \psi$. Similar relations hold also in the labeled version of the Boolean information algebra. It can be verified that (Φ, D) with these dual operations is still a Boolean information algebra and the mapping $\phi \rightarrow \phi^c$ is an isomorphism between dual Boolean information algebras.

Now, in the dual algebra, the partial order $\leq_d$ is defined as usual. Then, clearly, $\phi \leq_d \psi$ if, and only if, $\phi \geq \psi$.

From a domain-free Boolean algebra we may derive in the usual way (see Section 2.5) the associated labeled information algebra. This algebra as a whole is no more a Boolean algebra. Only the elements associated with a support $x \in D$ form still Boolean algebras. More precisely, a *labeled* information algebra (Φ, D) is called *Boolean*, if the following two properties hold:

1. $\forall x \in D$, the semilattice Φ_x is Boolean.
2. $\forall x, y \in D$ and $\phi, \psi \in \Phi_x$ is holds that

$$(\phi \wedge \psi) \otimes e_y = ((\phi \otimes e_y) \wedge (\psi \otimes e_y)).$$

The labeled algebra derived from a domain-free Boolean algebra certainly satisfies these properties. So does for example relational algebra, seen as a labeled information algebra.

Although Φ itself is not a Boolean algebra, it is still possible to define a *dual* algebra, using duality within the Boolean algebras Φ_x. So dual combination is defined for $\phi, \psi \in \Phi$ as

$$\phi \otimes_d \psi = (\phi^c \otimes \psi^c)^c.$$

Similarly, dual marginalization is defined for $\phi \in \Phi$ and $x \leq d(\phi)$ as

$$\phi^{\downarrow_d x} = ((\phi^c)^{\downarrow x})^c.$$

Note that the dual neutral elements are the original null elements z_x. So, dual vacuous extension is defined as follows for $x \geq d(\phi)$,

$$\phi^{\uparrow_d x} = \phi \otimes_d z_x = (\phi^c \otimes e_x)^c = ((\phi^c)^{\uparrow x})^c.$$

This allows finally the introduction of a dual transport operation, for $\phi \in \Phi$, with $d(\phi) = y$,

$$\phi^{\to_d x} = (\phi^{\uparrow_d x \vee y})^{\downarrow_d x} = ((\phi^c)^{\uparrow x \vee y})^{\downarrow x})^c = ((\phi^c)^{\to x})^c.$$

This results in a dual labeled information algebra, which is isomorph to the original one by the mapping $\phi \mapsto \phi^c$. We warn that the dual partial order $\leq_d$ induced in this dual algebra is *not* the inverse of the order $\leq$ in the original algebra. However, the dual order accounts for the issue addressed at the beginning of the section: according to the question one is interested in, one should either consider the one or the other of the dual algebras.

4.3 Atomic Algebras

In many cases there are for every domain x *most informative information pieces* representing the finest possible answers to the question posed by the domain. In relational algebra for example the one-tuple relations over a domain x represent such *atomic* information. In this section we study more generally information algebras with atomic information pieces.

For this purpose it is more convenient to work with a *labeled* information algebra (Φ, D). Remember now that the algebra has *null elements*, i.e. for all $x \in D$ there is a (necessarily unique) element z_x such that $\phi^{\to x} \otimes z_x = z_x$ for all $\phi \in \Phi$. We further have $z_x^{\to y} = z_y$. These null elements represent *contradictory information.* In fact, if $\phi \otimes \psi = z_x$, the combination of this pieces of information with further pieces yields again the contradiction. In relational algebra these null elements are represented by the empty relations, in propositional and predicate logic the logical constant $\perp$ (falsity), which has no models, represents contradiction.

Now, an *atom* in a domain x is a *maximal element* different form z_x among the elements Φ_x with domain x:

Definition 1. *An element $\alpha \in \Phi_x$ is called an atom on x if*

1. *$\alpha \neq z_x$,*
2. *for all $\phi \in \Phi_x$, $\alpha \leq \phi$ implies either $\alpha = \phi$ or $\phi = z_x$.*

Here are a few elementary properties of atoms which are proven in [9]:

1. If α is an atom on x, and $y \leq x$, then $\alpha^{\downarrow y}$ is an atom on y.
2. If α is an atom on x, and $d(\phi) = x$, then either $\phi \leq \alpha$ or $\alpha \otimes \phi = z_x$.
3. If α and β are atoms on x, then either $\alpha = \beta$ or $\alpha \otimes \beta = z_x$.

Denote the set of all atoms in Φ_x by $At_x(\Phi)$ and the set of all atoms in Φ by $At(\Phi)$. Furthermore let for any $\phi \in \Phi$

$$At(\phi) = \{\alpha \in At(\Phi) : d(\alpha) = d(\phi), \phi \leq \alpha\}.$$

If $\alpha \in At(\phi)$ we say also that α is an atom of ϕ or contained in ϕ. This terminology will be justified below.

We are now especially interested in information algebras, where each element is *composed* by all the atoms it contains. The following definition gives a more precise meaning to this idea:

Definition 2. *A labeled information algebra* (Φ, D) *is called atomic, if for all* $\phi \in \Phi$, $\phi \neq z_{d(\phi)}$,

$$\phi = \wedge At(\phi),$$

i.e. each information is the infimum of the atoms it contains.

The labeled versions of the information algebras associated with propositional logic and predicate logic are atomic: In the case of propositional logic, the elements of Φ_x can be considered as subsets of the Boolean cube $\{\mathbf{t}, \mathbf{f}\}^{|x|}$ and the atoms are tuples $t : x \to \{\mathbf{t}, \mathbf{f}\}$. Therefore each element of Φ_x is simply the set of the tuples it contains. Similarly, in the case of predicate logic, the elements of Φ_x can be considered as subsets of the cartesian product $U^{|x|}$ and the atoms are tuples $t : x \to U$. In the case of information algebras related to contexts, atoms exist, if $m^{\Rightarrow x}$ is $\models$-closed for all $m \in \mathcal{M}$ and $x \in D$. Then, if a cylindric set A is $\models$-closed,

$$A = A^{\Rightarrow x} = \bigcup_{m \in A} m^{\Rightarrow x}.$$

Hence, again, each element of Φ_x is simply the set of the atoms it contains. The example of linear manifolds shows however that not every set of atoms forms necessarily an element of Φ.

These examples reflect in fact a more general situation: We claim that the set $At(\Phi)$ of all atoms of an atomic information algebra (Φ, D) forms itself an information algebra, very similar to a relational algebra. We note first, that atoms behave with respect to projection like ordinary tuples in relational algebra. In fact, the following lemma summarizes the basic properties of atoms:

Lemma 5. *If a labeled information algebra* (Φ, D) *is atomic, then its atoms* α, β *in* $At(\Phi)$ *satisfy the following properties:*

1. *If* $x \leq d(\alpha)$*, then* $\alpha^{\downarrow x} \in At(\Phi)$ *and* $d(\alpha^{\downarrow x}) = x$*.*
2. *If* $x \leq y \leq d(\alpha)$*, then* $(\alpha^{\downarrow y})^{\downarrow x} = \alpha^{\downarrow x}$*.*
3. *If* $d(\alpha) = x$*, then* $\alpha^{\downarrow x} = \alpha$*.*
4. *If* $d(\alpha) = x$*,* $d(\beta) = y$ *and* $\alpha^{\downarrow x \wedge y} = \beta^{\downarrow x \wedge y}$*, then there exists a* $\gamma \in At(\Phi)$ *with* $d(\gamma) = x \vee y$ *and* $\gamma^{\downarrow x} = \alpha$*,* $\gamma^{\downarrow y} = \beta$*.*
5. *If* $d(\alpha) = x$ *and* $x \leq y$*, then there exists a* $\beta \in At(\Phi)$ *such that* $d(\beta) = y$ *and* $\beta^{\downarrow x} = \alpha$*.*

Proof. Properties (1) to (3) follow from the axioms of an information algebra, since atoms are elements of the algebra.

Let $\gamma = \alpha \otimes \beta$. Then $\gamma^{\downarrow x} = \alpha$ by the combination and idempotency axioms, considering that $\alpha^{\downarrow x \wedge y} = \beta^{\downarrow x \wedge y}$. Similarly, $\gamma^{\downarrow y} = \beta$. Assume that $\gamma = z_{x \vee y}$. But then $\alpha = z_x$, which is excluded, since α is an atom. Hence we conclude that $\gamma \neq z_{x \vee y}$. Therefore, since (Φ, D) is atomic, $At(\gamma)$ is not empty. Let $\eta \in At(\gamma)$. Then it follows from $\gamma \leq \eta$, that $\alpha = \gamma^{\downarrow x} \leq \eta^{\downarrow x}$. But since α is an atom, either $\alpha = \eta^{\downarrow x}$ or $\eta^{\downarrow x} = z_x$. The latter case is excluded, since η is an atom. Similarly $\beta = \eta^{\downarrow y}$. So property (4) is satisfied by η.

Further, $At(\alpha^{\uparrow y})$ is not empty either. Thus, let $\beta \in At(\alpha^{\uparrow y})$. Then $d(\beta) = y$ and $\alpha^{\uparrow y} \leq \beta$. This implies $\alpha = (\alpha^{\uparrow y})^{\downarrow x} \leq \beta^{\downarrow x}$. Since α is an atom, it holds that either $\alpha = \beta^{\downarrow x}$ or $\beta^{\downarrow x} = z_x$. But the latter case is excluded because β is an atom. So property (5) is satisfied by β. □

Of course, ordinary tuples in relational algebra satisfy these properties too. That is why we may consider atoms as *generalized tuples*. As with relational algebra, we define *generalized relations* over x to be subsets R of $At(\Phi)$ such that $d(\alpha) = x$ for all $\alpha \in R$. The *domain* of α is supposed to be attached to R. It is denoted by $d(R)$. For a generalized relation R and $x \leq d(R)$, the *projection* of R onto x is defined as

$$\pi_x(R) = \{\alpha^{\downarrow x} : \alpha \in R\}.$$

The *join* of a generalized relation R over x and a generalized relation S over y is defined as follows:

$$R \bowtie S = \{\alpha \in At(\Phi) : d(\alpha) = x \vee y, \alpha^{\downarrow x} \in R, \alpha^{\downarrow y} \in S\}.$$

It is easily possible that the set on the right hand side is empty. We attach the empty set with the domain $x \vee y$ and call it $Z_{x \vee y}$, the *empty* relation on $x \vee y$. We assign it the domain $d(Z_{x \vee y}) = x \vee y$. Finally, for $x \in D$, the full relation over x is

$$E_x = \{\alpha \in At(\Phi) : d(\phi) = x\} = At_x(\Phi).$$

This is the neutral element for the join operation between generalized relations on x. Note that $R \bowtie S = R \cap S$ if R and S are relations over the same domain.

Let $\mathcal{R}_\Phi$ be the set of all generalized relations of atoms of the information algebra (Φ, D). Then these generalized relations form a labeled information algebra.

Theorem 3. *The two-sorted algebra $(\mathcal{R}_\Phi, D)$ with the operations of projection and join defined above forms a labeled information algebra.*

This can easily be verified. In fact, it satisfies the same properties as an ordinary relational algebra summarized in Section 2.1. Furthermore, just as ordinary relational algebra, it forms the labeled version of a *Boolean information algebra*. It turns out that the atomic algebra (Φ, D) is part of its associated generalized relational algebra.

Assume the labeled information algebra (Φ, D) to be atomic.

Theorem 4. *The mapping $At : \Phi \to \mathcal{R}_\Phi$ defined by $\phi \mapsto At(\phi)$ is an embedding of (Φ, D) into $(\mathcal{R}_\Phi, D)$.*

Proof. We have first to show the following:

1. $At(\phi \otimes \psi) = At(\phi) \bowtie At(\psi)$,
2. $At(\phi^{\downarrow x}) = \pi_x(At(\phi))$.
3. $At(e_x) = E_x$.
4. $At(z_x) = Z_x$.

(1) Let $d(\phi) = x$ and $d(\psi) = y$. Consider an atom $\alpha \in At(\phi \otimes \psi)$. Then it follows that $\phi \leq \phi \otimes \psi \leq \alpha$, hence $\phi = \phi^{\downarrow x} \leq \alpha^{\downarrow x}$. Thus $\alpha^{\downarrow x} \in At(\phi)$. Similarly $\alpha^{\downarrow y} \in At(\psi)$, hence $\alpha \in At(\phi) \bowtie At(\psi)$. Conversely, assume $\alpha \in At(\phi) \bowtie At(\psi)$. Then $d(\alpha) = x \vee y$ and $\phi \leq \alpha^{\downarrow x} \leq \alpha$ and $\psi \leq \alpha^{\downarrow y} \leq \alpha$, hence $\phi \otimes \psi \leq \alpha$, which means that $\alpha \in At(\phi \otimes \psi)$. This proves (1).

(2) Let $\alpha \in At(\phi^{\downarrow x})$, such that $d(\alpha) = x$ and $\phi^{\downarrow x} \leq \alpha$. We have $\phi \leq \phi \otimes \alpha$. Suppose that $\phi \otimes \alpha = z_y$, if $d(\phi) = y \geq x$. Then

$$\alpha = \phi^{\downarrow x} \otimes \alpha = (\phi \otimes \alpha)^{\downarrow x} = z_x.$$

But this is excluded, because α is an atom. Therefore $\phi \otimes \alpha \neq z_y$. Since Φ is atomic there is a $\beta \in At(\phi \otimes \alpha)$ with $d(\beta) = y$ and $\phi \leq \beta$, hence $\beta \in At(\phi)$. But we have also $\alpha = (\phi \otimes \alpha)^{\downarrow x} \leq \beta^{\downarrow x}$. Since $\beta^{\downarrow x}$ is also an atom, we must have $\alpha = \beta^{\downarrow x}$ and therefore $\alpha \in \pi_x(At(\phi))$. Conversely, if $\beta \in \pi_x(At(\phi))$, then $\beta = \gamma^{\downarrow x}$ for some atom $\gamma \in At(\phi)$. But $\phi \leq \gamma$, hence $\phi^{\downarrow x} \leq \beta$ and therefore $\beta \in At(\phi^{\downarrow x})$. So (2) holds.

(3), (4) follow directly from the definition of At.

It remains to show that the mapping At is one-to-one. Assume $At(\phi) = At(\psi)$. Then $\phi = \wedge At(\phi) = \wedge At(\psi) = \psi$. □

The information algebras associated with propositional logic, for instance, coincide with their relational version. But this is not the case in general. The information algebras associated with predicate logic are proper subalgebras of the relational information algebra of relations over U.

In the case of an *atomic Boolean information algebra* (Φ, D) there is also a dual notion of the concept of an atom. A dual atom on x is a maximal element on x with respect to the dual order $\leq_d$. Let $At_d(\Phi)$ denote the set of dual atoms. If $\alpha \in At_d(\Phi)$ and $d(\alpha) = x$, then $\alpha \neq z_x^c = e_x$, hence $\alpha^c \neq z_x$. Further, assume $\alpha^c \leq \phi$ for a $\phi \in \Phi_x$. Then $\alpha \leq_d \phi^c$, hence either $\phi^c = \alpha$, i.e. $\phi = \alpha^c$, or $\phi^c = z_x^c$, i.e. $\phi = z_x$. Thus if α is a dual atom, then α^c is an atom.

Lemma 6. *Let (Φ, D) be an atomic Boolean information algebra. Then the following holds:*

1. $At(\phi) \cap At(\phi^c) = \emptyset$.
2. *If* $d(\phi) = x$, *then* $At(\phi) \cup At(\phi^c) = At_x(\Phi)$.

Proof. (1) Suppose there is an atom α on x such that $\phi \leq \alpha$ and $\phi^c \leq \alpha$. Taking the join of both sides we obtain $z_x \leq \alpha$, which is impossible.

(2) If $\alpha \in At_x(\Phi)$, and $d(\phi) = x$, then either $\phi \leq \alpha$ and $\alpha \in At(\phi)$ or $\alpha \vee \phi = z_x$. But in the latter case $\phi^c \leq \alpha$ and $\alpha \in At(\phi^c)$. □

Note further that $\alpha \in At(\phi)$, i.e. $\phi \leq \alpha$, implies $\phi^c \leq_d \alpha^c$, hence $\alpha^c \in At_d(\phi^c)$ and vice versa. Hence the cardinality of the two sets $At(\phi)$ and $At_d(\phi^c)$ are the same. This implies also that the sets of all atoms on x, $At_x(\Phi) = At(e_x)$ and $At_{d,x}(\Phi) = At_d(z_x)$ have the same cardinality.

4.4 Measure of Information Content

Shannon, in his information theory, introduces a *quantitative measure* of information. He measures the information about a transmitted symbol by the reduction of uncertainty, when the transmitted symbol becomes known. So, in our context, we may say that Shannon considers a fixed *question*, namely what symbol out of a (finite) alphabet is selected for transmission. The uncertainty is measured by the *entropy* of the alphabet [19]. Once the symbol to be transmitted is known, the uncertainty is reduced to zero. Therefore the entropy measures the information gained by knowing the symbol.

This basic idea can be applied in our context too, if the labeled information algebra (Φ, D) is *atomic*. The first point to stress is that the information content of an element $\phi \in \Phi$ is measured relative to its domain $d(\phi)$, i.e. relative to the question it refers to. We assume further that for all domains $x \in D$ the total number $At(e_x)$ of atoms of the domain is *finite*. An atom of a domain x is the finest, i.e. the maximal information one may obtain about the domain. Assuming the number of atoms finite means that this information can be coded by a number of bits bounded by $\lceil \log_2 |At(e_x)| \rceil$, whereas an infinite number of atoms would mean that the information in an atom cannot be coded into a finite memory. Then the total uncertainty associated with a domain x can be measured by $\log |At(e_x)|$, the *Hartley measure*. This corresponds to the entropy of $At(e_x)$ under an assumed uniform probability distributions over the atoms. However, we shall avoid here probabilistic considerations, since there is no random experiment involved in our discussion. Usually the logarithm is taken to base 2, but any other base serves our purpose too, since it involves only a shift of scale in the measurement of uncertainty and information. Once information ϕ with $d(\phi) = x$ is given, the uncertainty concerning the possible atoms is reduced to $\log |At(\phi)| \leq \log |At(e_x)|$. So, the information content of ϕ relative to the domain (question) x can be defined as the reduction of uncertainty obtained by ϕ with respect to knowing nothing (i.e. knowing only the vacuous information e_x),

$$i(\phi) = \log |At(e_x)| - \log |At(\phi)| = -\log \frac{|At(\phi)|}{|At(e_x)|}. \tag{7}$$

We may consider

$$p(\phi) = \frac{|At(\phi)|}{|At(e_x)|}$$

as the probability of ϕ, or, more precisely, the probability that an atom in $At(\phi)$ is selected out of the atoms of $At(e_x)$, when all atoms have the same chance to be selected. Then we obtain $i(\phi) = -\log p(\phi)$, which corresponds to an often proposed definition of the information content of an "event" observed in a random experiment. But, once more, we prefer at this place to not refer to probabilistic considerations, since, in our view, information, in the first place at least, has nothing to do with probability, although in applications probability may play an important role (as for example in communication theory). We note that $i(e_x) = 0$, the vacuous information carries no information. Further we obtain $i(z_x) = \infty$ by (7) since $At(z_x) = \emptyset$. Note that z_x is in fact not really an information about a possible atom, since it contains no atom at all. We could as well convene that the information content of z_x is not defined.

More generally, any element $\phi \in \Phi$ contains possibly information about any other domain $y \neq d(\phi)$. In fact, it is natural to define the information content of ϕ relative to domain y by

$$i(\phi; y) = i(\phi^{\to y}).$$

Clearly and consistently we see that $i(\phi; x) = i(\phi)$, if $d(\phi) = x$. Further, if $[\phi]$ is the class of equivalent information elements (see Section 2.5), then all elements of the class have the same information content $i(\phi; y)$ with respect to any domain y. This means that we may assign a measure of information content by defining $i([\phi]; y) = i(\phi; y)$ also to the elements of the domain-free version of an information algebra.

The next theorem shows that our quantitative measure of information respects the qualitative orders of information introduced above (Section 4.1):

Theorem 5. *Let (Φ, D) be an atomic information algebra, with finite sets of atoms $At_x(\Phi)$. Then, for all $x \in D$ and $\phi, \psi \in \Phi$, the inequalities $\phi \leq \psi$, $[\phi] \leq [\psi]$ and $\phi \leq_x \psi$ imply $i(\phi; x) \leq i(\psi; x)$.*

Proof. Both $\phi \leq \psi$ and $[\phi] \leq [\psi]$ imply $\phi \leq_x \psi$. The latter implies $At(\phi^{\to x}) \supseteq At(\psi^{\to x})$, hence $|At(\phi^{\to x})| \geq |At(\psi^{\to x})|$, and therefore $i(\phi; x) = i(\phi^{\to x}) \leq i(\psi^{\to x}) = i(\psi; x)$. □

From this theorem a number of simple results may be derived, which follow from the properties of the partial order: For all $x, y, z \in D$ and $\phi, \psi \in \Phi$:

1. $i(\phi; x), i(\psi; x) \leq i(\phi \otimes \psi; x)$,
2. $i(\phi^{\to y}; x) \leq i(\phi; x)$,
3. $\phi \leq \psi$ implies $i(\phi^{\to y}; x) \leq i(\psi^{\to y}; x)$,
4. $\phi_1 \leq \phi_2$ and $\psi_1 \leq \psi_2$ imply $i(\phi_1 \otimes \psi_1 : x) \leq i(\phi_2 \otimes \psi_2 : x)$,
5. $i(\phi^{\to y} \otimes \psi^{\to y}; x) \leq i((\phi \otimes \psi)^{\to y}; x)$,
6. $x \leq y$ implies $i(\phi^{\to x}; z) \leq i(\phi^{\to y}; z)$.

When the partial order in the domain-free algebra is considered, similar results are obtained. So far we measure the information content of an element ϕ on a domain x by the reduction of the uncertainty with respect to initial ignorance, i.e. vacuous information. More generally, we may also measure the relative information content of a piece of information relative to another, previous piece of information ψ. If information ψ is already given, the remaining uncertainty is $\log|At(\psi)|$. When a new information ϕ arrives, the total information is $\psi \otimes \phi$, and the remaining uncertainty $\log|At(\psi \otimes \phi)|$. If ϕ and ψ have the same domain $d(\phi) = d(\psi) = x$, then the relative information content of ϕ relative to ψ relating to the domain x can then be measured by

$$i(\phi|\psi) = \log|At(\psi)| - \log|At(\psi \otimes \phi)| = -\log\frac{|At(\psi \otimes \phi)|}{|At(\psi)|}.$$

Since by Theorem 4 in this case $At(\psi \otimes \phi) = At(\phi) \cap At(\psi)$, we may also consider $i(\phi|\psi) = -\log p(\phi|\psi)$, i.e. as the negative logarithm of the *conditional probability* of ϕ given ψ, with the usual assumption of uniform probability distribution over the atoms of $At(e_x)$. As before we may extend this definition of relative information measure to any domain y and information elements ψ and ϕ on any domains

$$i(\phi|\psi;x) = \log|At(\psi^{\rightarrow x})| - \log|At((\psi \otimes \phi)^{\rightarrow x})| = -\log\frac{|At((\psi \otimes \phi)^{\rightarrow x})|}{|At(\psi^{\rightarrow x})|}.$$

Note however that in general $i(\phi|\psi;x) \neq i(\phi^{\rightarrow x}|\psi^{\rightarrow x})$. Further, $\psi \leq \psi \otimes \phi$ implies also $\psi^{\rightarrow x} \leq (\psi \otimes \phi)^{\rightarrow x}$; therefore we conclude that $i(\phi|\psi;x) \geq 0$. It may be that $\psi \otimes \phi = z_y$, which means that ϕ and ψ are *incompatible* or *contradictory* pieces of information. Correspondingly we obtain in this case $i(\phi|\psi;x) = \infty$ for all domains x. This is simply the mathematical expression for the fact that such two pieces of information can not hold at the same time. Note further that $i(\phi;x) = i(\phi|e_x;x)$.

The following result shows that the measure of a combined information can be obtained as the sum of the measure of the first information and the relative information of the second relative to the first one. This is called the *chaining theorem.*

Theorem 6. *For all $x \in D$ and $\phi, \psi \in \Phi$ it holds that*

$$i(\phi \otimes \psi; x) = i(\phi; x) + i(\psi|\phi; x).$$

Proof. We have

$$\begin{aligned} i(\phi \otimes \psi; x) &= \log|At(e_x)| - \log|At(\phi \otimes \psi)^{\rightarrow x}| \\ &= (\log|At(e_x)| - \log|At(\phi^{\rightarrow x})|) \\ &\quad +(\log|At(\phi^{\rightarrow x})| - \log|At(\phi \otimes \psi)^{\rightarrow x}|) \\ &= i(\phi; x) + i(\psi|\phi; x). \end{aligned}$$

□

This result can easily be generalized to the combination of $n \geq 2$ information elements.

Corollary 1. *For all $x \in D$ and $\phi_1, \ldots, \phi_n \in \Phi$,*

$$i(\phi_1 \otimes \cdots \otimes \phi_n; x) = i(\phi_1; x) + i(\phi_2|\phi_1; x) + \cdots + i(\phi_n|\phi_1 \otimes \cdots \otimes \phi_{n-1}; x).$$

Here are a few simple results about relative information.

Lemma 7. *For all $x \in D$ and $\phi, \psi \in \Phi$*

1. *If $\phi \geq \psi$, then $i(\phi|\psi; x) \leq i(\phi; x)$.*
2. *$\phi_1 \leq \phi_2$ implies $i(\phi_1|\psi; x) \leq i(\phi_2|\psi; x)$.*
3. *$\phi \leq \psi$ implies $i(\phi|\psi; x) = 0$.*

Proof. (1) We note that $At(e_x) \supseteq At(\psi^{\rightarrow x})$ and $At((\phi \otimes \psi)^{\rightarrow x}) = At(\phi^{\rightarrow x})$. Then

$$\begin{aligned} i(\phi|\psi; x) &= \log |At(\psi^{\rightarrow x})| - \log |At((\phi \otimes \psi)^{\rightarrow x})| \\ &\leq \log |At(e_x)| - \log |At(\phi^{\rightarrow x})| \\ &= i(\phi; x). \end{aligned}$$

(2) follows since $(\phi_1 \otimes \psi)^{\rightarrow x} \leq (\phi_2 \otimes \psi)^{\rightarrow x}$, hence $At((\phi_1 \otimes \psi)^{\rightarrow x}) \supseteq At((\phi_2 \otimes \psi)^{\rightarrow x})$.

(3) follows because in this case $\phi \otimes \psi = \psi$. □

Suppose $i(\phi|\psi; x) = i(\phi; x)$ and $i(\psi|\phi; x) = i(\psi; x)$. In this case, knowing ψ contributes nothing to the information represented by ϕ and, similarly, knowing ϕ contributes nothing to the information represented by ψ. Therefore we say that ϕ and ψ are *independent* pieces of information relative to x and we write $\phi||\psi; x$. In this case, by the Chaining Theorem 6, the following *additivity property* holds,

$$i(\phi \otimes \psi; x) = i(\phi; x) + i(\psi; x).$$

Independent information simply adds up.

An important special case are *atomic Boolean information algebras*. We may define there a *dual information measure* for an element $\phi \in \Phi$ with $d(\phi) = x$,

$$i_d(\phi) = \log |At_d(z_x)| - \log |At_d(\phi)|,$$

since z_x is the dual neutral element, hence the dual vacuous information. This dual measure makes sense: In relational databases for instance, if a relation indicates all the flights by which a person can arrive, then the first measure applies, the smaller the relation the more information is available. When however the relations represents all the flights which a person may select for her trip, then the dual measure applies, the larger the relation, the more information is given. We have seen that $|At_d(\phi)| = |At(\phi^c)|$ (Section 4.3), hence

$$i_d(\phi) = \log |At(e_x)| - \log |At(\phi^c)| = i(\phi^c).$$

This duality relation between the dual information measures holds also for the general information measure relative to a domain

$$i_d(\phi; x) = i_d(\phi^{\rightarrow_d x}) = i((\phi^{\rightarrow_d x})^c) = i((\phi^c)^{\rightarrow x}) = i(\phi^c; x).$$

It holds also for relative information. By the dual chaining theorem,

$$\begin{aligned} i_d(\phi|\psi; x) &= i_d(\phi \otimes_d \psi; x) - i_d(\psi; x) \\ &= i((\phi \otimes_d \psi)^c; x) - i(\psi^c; x) \\ &= i(\phi^c \otimes \psi^c; x) - i(\psi^c; x) \\ &= i(\phi^c|\psi^c; x). \end{aligned}$$

We illustrate these concepts in the important case of information algebras with *multivariate domains*, where the results can be considerably sharpened. Assume thus that (Φ, D) is an atomic information algebra, where D is a lattice of subsets of some set r such that $x, y \in D$ implies $x - y \in D$. This is the case for instance for the lattice of finite subsets of an arbitrary set r. We introduce two further assumptions:

– Every atom $\alpha \in At_{x \cup y}(\Phi)$ on the domain $x \cup y$ has a decomposition of the form

$$\alpha = \alpha^{\downarrow x} \otimes \alpha^{\downarrow y}. \tag{8}$$

– For every $\eta \in \Phi_t$ we have

$$\eta^{\downarrow \emptyset} = \begin{cases} e_\emptyset, & \text{if } \eta \neq z_t, \\ z_\emptyset, & \text{else.} \end{cases} \tag{9}$$

Note that the combination of atoms is, in the general case, not necessarily an atom. This condition is satisfied, whenever Φ contains subsets of cartesian products, i.e. in the case of propositional and predicate logic, relational algebra and linear manifolds. As before we assume that the atom sets $At_x(\Phi)$ are finite for all $x \in D$. This is the case for propositional logic, predicate logic and relational algebra with finite domains and linear manifolds over product spaces of finite (or Galois) fields. In this case the following basic result holds:

Lemma 8. *Let (Φ, D) is an atomic information algebra, where D is a lattice of subsets of some set r, and such that conditions (8) and (9) hold. Then, if for $x, y \in D$ with $x \cap y = \emptyset$, and*

$$\phi = \phi^{\downarrow x} \otimes \phi^{\downarrow y}, \tag{10}$$

it holds that

$$i(\phi) = i(\phi^{\downarrow x}) + i(\phi^{\downarrow y}).$$

Proof. From (10) and Theorem 4 it follows that

$$\begin{aligned} At(\phi) &= At(\phi^{\downarrow x}) \bowtie At(\phi^{\downarrow y}) \\ &= \{\alpha \in At_{x \cup y}(\Phi) : \alpha = \alpha^{\downarrow x} \otimes \alpha^{\downarrow y}, \alpha^{\downarrow x} \in At(\phi^{\downarrow x}), \alpha^{\downarrow y} \in At(\phi^{\downarrow y})\}. \end{aligned}$$

From this we conclude that $|At(\phi)| = |At(\phi^{\downarrow x})| \cdot |At(\phi^{\downarrow y})|$. Similarly, it follows that $|At(e_{x \cup y})| = |At(e_x)| \cdot |At(e_y)|$. Therefore we obtain

$$\begin{aligned} i(\phi) &= \log |At(e_{x \cup y})| - \log |At(\phi)| \\ &= (\log |At(e_x)| + \log |At(e_y)|) - (\log |At(\phi^{\downarrow x})| + \log |At(\phi^{\downarrow y})|) \\ &= (\log |At(e_x)| - \log |At(\phi^{\downarrow x})|) + (\log |At(e_y)| - \log |At(\phi^{\downarrow y})|) \\ &= i(\phi^{\downarrow x}) + i(\phi^{\downarrow y}). \end{aligned}$$

□

This result allows to introduce an absolute information measure into the domain-free version of the information algebra. In fact, if $\phi \equiv \psi$, then $\phi^{\uparrow x \cup y} = \psi^{\uparrow x \cup y}$, if $d(\phi) = x$ and $d(\psi) = y$. Let $z = x \cup y - x$. Then, since $\phi^{\uparrow x \cup y} = \phi \otimes e_z = (\phi^{\uparrow x \cup y})^{\downarrow x} \otimes (\phi^{\uparrow x \cup y})^{\downarrow z}$, by the previous Lemma 8 $i(\phi^{\uparrow x \cup y}) = i(\phi)$. Similarly we obtain $i(\psi^{\uparrow x \cup y}) = i(\psi)$, hence $i(\phi) = i(\psi)$. Define then the *absolute information measure* $i([\phi]) = i(\phi)$. The absolute information measure respects the partial information order in the domain-free information algebra $\Phi/\equiv$. Indeed, if $[\phi] \leq [\psi]$, then $[\phi] \otimes [\psi] = [\phi \otimes \psi] = [\psi]$. We have then $i([\phi]) = i(\phi^{\uparrow x \cup y}) \leq i(\phi \otimes \psi) = i(\psi^{\uparrow x \cup y}) = i([\psi])$, if $d(\phi) = x$ and $d(\psi) = y$.

In a similar way we define the relative information measure

$$i([\phi]|[\psi]) = \log |At(\psi^{\uparrow x \cup y})| - \log |At(\phi \otimes \psi)| = i([\phi] \otimes [\psi]) - i([\psi]) \geq 0.$$

Thus the absolute chaining theorem holds too,

$$i([\phi] \otimes [\psi]) = i([\phi]) + i([\psi]|[\phi]).$$

As before, we may call $[\phi]$ and $[\psi]$ *independent*, if the addition property

$$i([\phi] \otimes [\psi]) = i([\phi]) + i([\psi])$$

or, equivalently, $i([\phi]) = i([\phi]|[\psi])$ and $i([\psi]) = i([\psi]|[\phi])$ hold. This is the case if there are supports x and y of $[\phi]$ and $[\psi]$ respectively, such that $x \cap y = \emptyset$.

5 Conclusion

Information algebras represent a structure which captures essential features of any concept of "information". The presentation here focuses on its basic theory. There are many more aspects: One is computation: How are pieces of information combined and focussed on the domains of interest? This is the problem of *query processing* where *local computation methods* can be applied. A lot of work has been done with respect to this problem. In particular domains like query processing in relational algebra, solving linear equations, for instance in coding theory, consequence finding in logic, etc have been studied extensively. Whereas each of these domains has its particularities which can and must be explored, information algebra offers a common background on which *generic* methods can be developed [9, 12] . In this respect we refer to [18] which describes a generic software for local computation, permitting instantiations with any information (or rather valuation) algebra.

Another issue is *approximation* of ‘infinite” information by “finite” information. This is modeled by *compact* information algebras [9]. More precisely, the question arises when an information algebra is *effective*, i.e. when the operations of combination and focussing can effectively be computed on a computer. Similar questions arise in domain theory, a theory which has close links to compact information algebras, see for instance [22]. In this same context, one can ask how information and its algebra is related to deduction. One is used that information may allow the inference of further information; in fact projection is a deduction procedure. Particularly in logic, i.e. in contexts, this became clear. In fact, it turns out that, similar to domain theory, information algebra may be equivalently be replaced by a system based on entailment, similar to *information systems* in the sense of Scott [22]. This means that *logic* in a wider sense is a general way to express and treat information. In a similar way, it can also be shown that any information algebra, in some precise sense, is part of *generalized relational algebra*. Thus, there are two general and complementary ways to see information in general, as (generalized) *relations* or as *logic*. These relations are discussed in [9].

Finally comes up the idea of “uncertain information”, a term often used, but rarely, if ever, precisely defined. In the framework of information algebras, uncertain information can be represented by random variables taking values in an information algebra [9, 10]. This is closely related to *probabilistic argumentation systems* [5, 9]. Here, probability theory is combined with logic, in the way that the latter serves to prove hypotheses under certain assumptions and the former permits to compute the probability that those assumptions are valid. This brings the theory more into the realm of Shannon’s entropy based information theory. Also it generalizes the concept of *random sets* [16], which are usually considered as random variables taking values in an algebra of closed sets of some topological space.

References

1. Barwise, J., Seligman, J.: Information Flow: The Logic of Distributed Systems. Cambridge Tracts in Theoretical Computer Science, vol. 44. Cambridge University Press, Cambridge (1997)
2. Beeri, C., Fagin, R., Maier, D., Yannakakis, M.: On the desirability of acyclic database schemes. Journal of the ACM 30(3), 479–513 (1983)
3. Davey, B.A., Priestley, H.A.: Introduction to Lattices and Order. Cambridge University Press, Cambridge (1990)
4. Grätzer, G.: General Lattice Theory. Academic Press, London (1978)

5. Haenni, R., Kohlas, J., Lehmann, N.: Probabilistic argumentation systems. In: Kohlas, J., Moral, S. (eds.) Handbook of Defeasible Reasoning and Uncertainty Management Systems. Algorithms for Uncertainty and Defeasible Reasoning, vol. 5, pp. 221–287. Kluwer, Dordrecht (2000)
6. Halmos, P.R.: Algebraic Logic. Chelsea, New York (1962)
7. Paul, R.: Halmos and Steven Givant. Logic as Algebra. Dolciani Mathematical Expositions No. 21, Mathematical Association of America (1998)
8. Henkin, L., Monk, J.D., Tarski, A.: Cylindric Algebras. North-Holland, Amsterdam (1971)
9. Kohlas, J.: Information Algebras: Generic Structures for Inference. Springer, Heidelberg (2003)
10. Kohlas, J.: Uncertain information: random variables in graded semilattices. Int. J. Approx. Reason. (2007) doi:10.1016/j.ijar.2006.12.005
11. Kohlas, J., Haenni, R., Moral, S.: Propositional information systems. Journal of Logic and Computation 9(5), 651–681 (1999)
12. Kohlas, J., Shenoy, P.P.: Computation in valuation algebras. In: Kohlas, J., Moral, S. (eds.) Handbook of Defeasible Reasoning and Uncertainty Management Systems. Algorithms for Uncertainty and Defeasible Reasoning, vol. 5, pp. 5–39. Kluwer, Dordrecht (2000)
13. Kohlas, J., Wilson, N.: Exact and approximate local computation in semiring induced valuation algebras. Technical Report 06-06, Department of Informatics, University of Fribourg (2006)
14. Lauritzen, S.L., Spiegelhalter, D.J.: Local computations with probabilities on graphical structures and their application to expert systems. J. of Royal Stat. Soc. 50(2), 157–224 (1988)
15. Maier, D.: The Theory of Relational Databases. Pitman, London (1983)
16. Molchanov, I.: Theory of Random Sets. Springer, London (2005)
17. Plotkin, B.I.: Universal Algebra, Algebraic Logic, and Databases. Mathematics and its applications, vol. 272. Kluwer Academic Publishers, Dordrecht (1994)
18. Pouly, M.: Nenok 1.1 user guide. Technical Report 06-02, Department of Informatics, University of Fribourg (2006)
19. Shannon, C.E.: A mathematical theory of communications. The Bell System Technical Journal 27, 379–432 (1948)
20. Shenoy, P.P., Shafer, G.: Axioms for probability and belief-function proagation. In: Shachter, R.D., Levitt, T.S., Kanal, L.N., Lemmer, J.F. (eds.) Uncertainty in Artificial Intelligence 4. Machine intelligence and pattern recognition, vol. 9, pp. 169–198. Elsevier, Amsterdam (1990)
21. Shenoy, P.P., Shafer, G.: Axioms for probability and belief function propagation. In: Lemmer, J.F., Shachter, R.D., Levitt, T.S., Kanal, L.N. (eds.) Uncertainty in Artif. Intell. 4, pp. 169–198. North-Holland, Amsterdam (1990)
22. Stoltenberg-Hansen, V., Lindstroem, I., Griftor, E.: Mathematical Theory of Domains. Cambridge University Press, Cambridge (1994)
23. Wilson, N., Mengin, J.: Logical deduction using the local computation framework. In: Hunter, A., Parsons, S. (eds.) ECSQARU 1999. LNCS, vol. 1638, pp. 386–396. Springer, Heidelberg (1999)