

PART I

LOCAL COMPUTATION

COPYRIGHTED MATERIAL

CHAPTER 1

VALUATION ALGEBRAS

The valuation algebra framework provides the algebraic foundation for the application of all generic inference mechanisms introduced in this book and therefore marks the beginning of our studies. Comparable to the total order that is required for the application of sorting procedures, all formalisms must satisfy the structure of a valuation algebra in order to be processed by these generic inference tools. Further, this framework mirrors all the essential properties we naturally associate with the rather imprecise notions of knowledge and information. Let us engross ourselves in this thought and put our daily perception of information into words: information exists in pieces and comes from different sources. A piece of information refers to some specific questions that we later call the domain of an information piece. Also, there are two principal operations to manipulate information: we may combine or aggregate pieces of information to a new information piece and we may project a piece of information to some specific question which corresponds to information extraction. Depending on the operation, we either get a broader or more focused information. In the following section, we give a formal definition of the valuation algebra framework consisting of its operations and axioms. Along the way, we further bear on our idea of valuations as pieces of knowledge or information to clarify the rather abstract and formal structures.

1.1 OPERATIONS AND AXIOMS

The basic elements of a valuation algebra are so-called *valuations* that we subsequently denote by lower-case Greek letters ϕ, ψ, \dots . Intuitively, a valuation can be regarded as a representation of information about the possible values of a finite set of variables. We use Roman capitals X, Y, \dots to refer to variables and lower-case letters s, t, \dots for sets of variables. Thus, we assume that each valuation ϕ refers to a finite set of variables $d(\phi)$, called its *domain*. For an arbitrary, finite set of variables s , Φ_s denotes the set of all valuations ϕ with $d(\phi) = s$. With this notation, the set of all possible valuations for a countable set of variables r can be defined as

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

Let $D = \mathcal{P}(r)$ be the powerset (the set of all subsets) of r and Φ the set of valuations with their domains in D . We assume the following operations defined in $\langle \Phi, D \rangle$:

1. *Labeling*: $\Phi \rightarrow D; \phi \mapsto d(\phi)$;
2. *Combination*: $\Phi \times \Phi \rightarrow \Phi; (\phi, \psi) \mapsto \phi \otimes \psi$;
3. *Projection*: $\Phi \times D \rightarrow \Phi; (\phi, x) \mapsto \phi^{\perp x}$ for $x \subseteq d(\phi)$.

These are the three basic operations of a valuation algebra. If we readopt our idea of valuations as pieces of information, the labeling operation tells us to which questions (variables) such a piece refers. Combination can be understood as aggregation of information and projection as the extraction of the part we are interested in. Sometimes this operation is also called *focusing* or *marginalization*. We now impose the following set of axioms on Φ and D :

(A1) *Commutative Semigroup*: Φ is associative and commutative under \otimes .

(A2) *Labeling*: For $\phi, \psi \in \Phi$,

$$d(\phi \otimes \psi) = d(\phi) \cup d(\psi). \quad (1.1)$$

(A3) *Projection*: For $\phi \in \Phi, x \in D$ and $x \subseteq d(\phi)$,

$$d(\phi^{\perp x}) = x. \quad (1.2)$$

(A4) *Transitivity*: For $\phi \in \Phi$ and $x \subseteq y \subseteq d(\phi)$,

$$(\phi^{\perp y})^{\perp x} = \phi^{\perp x}. \quad (1.3)$$

(A5) *Combination*: For $\phi, \psi \in \Phi$ with $d(\phi) = x$, $d(\psi) = y$ and $z \in D$ such that $x \subseteq z \subseteq x \cup y$,

$$(\phi \otimes \psi)^{\downarrow z} = \phi \otimes \psi^{\downarrow z \cap y}. \quad (1.4)$$

(A6) *Domain*: For $\phi \in \Phi$ with $d(\phi) = x$,

$$\phi^{\downarrow x} = \phi. \quad (1.5)$$

These axioms require natural properties regarding knowledge or information. The first axiom indicates that Φ is a commutative semigroup under combination. If information comes in pieces, the sequence of their aggregation does not influence the overall result. The labeling axiom tells us that the combination of valuations yields knowledge about the union of the involved domains. Variables do not vanish, nor do new ones appear. The projection axiom expresses the natural functioning of focusing. Transitivity says that projection can be performed in several steps. For the combination axiom, let us assume that we have some information about a domain in order to answer a certain question. Then, the combination axiom states how the answer is affected if a new information piece arrives. We can either combine the new piece to the given information and project afterwards to the specified domain, or first remove the uninteresting parts of the new information and combine it afterwards. Both approaches lead to the same result. In fact, we are going to see in Section 5.3 that this axiom correlates with some generalized distributive law. Finally, the domain axiom ensures that information is not influenced by projection to its own domain, which expresses some kind of stability with respect to trivial projection.

Definition 1.1 A system $\langle \Phi, D \rangle$ together with the operations of labeling, projection and combination satisfying these axioms is called a valuation algebra.

In the first appearance of the valuation algebra axioms (Shafer & Shenoy, 1988; Shenoy & Shafer, 1990) only Axioms (A1), (A4) and a simpler version of (A5) were listed. Axiom (A2) was simply assumed in the definition of combination. (Shafer, 1991) mentioned Property (A3) for the first time and also remarked that Axiom (A6) cannot be derived from the others. (Kohlas, 2003) then assembled these early results to a complete and sufficient axiomatic system for generic inference and local computation. But in addition to the above system, all former approaches contained so-called neutral valuations that express vacuous information with respect to a certain domain. Here, we will introduce valuation algebras with neutral elements in Section 3.3 as a special case of the definition given here.

Before we turn towards the first concrete examples of valuation algebras, we list a few elementary properties that are derived directly from the above set of axioms.

Lemma 1.1

1. If $\phi, \psi \in \Phi$ with $d(\phi) = x$ and $d(\psi) = y$, then

$$(\phi \otimes \psi)^{\downarrow x \cap y} = \phi^{\downarrow x \cap y} \otimes \psi^{\downarrow x \cap y}. \quad (1.6)$$

2. If $\phi, \psi \in \Phi$ with $d(\phi) = x$, $d(\psi) = y$ and $z \subseteq x$, then

$$(\phi \otimes \psi)^{\downarrow z} = (\phi \otimes \psi^{\downarrow x \cap y})^{\downarrow z}. \quad (1.7)$$

Proof:

1. By the transitivity and combination axiom:

$$(\phi \otimes \psi)^{\downarrow x \cap y} = ((\phi \otimes \psi)^{\downarrow x})^{\downarrow x \cap y} = (\phi \otimes \psi^{\downarrow x \cap y})^{\downarrow x \cap y} = \phi^{\downarrow x \cap y} \otimes \psi^{\downarrow x \cap y}.$$

2. By the transitivity and combination axiom:

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

■

1.2 FIRST EXAMPLES

Our first encounter with the valuation algebra framework took place on a very abstract level. To reward the reader for this formal effort, we now consider a first catalogue of concrete formalisms that satisfy the valuation algebra structure. Such formalisms are subsequently called *valuation algebra instances*. For the moment, we content ourselves with the observation of how these formalisms fit in the valuation algebra framework. Later sections then inform about their exact purpose and semantics. However, these instances are all based on the important concepts of *configurations*, *tuples* or *vectors* that shall first be introduced separately.

Frames, Tuples, Configurations and Vectors:

Consider a countable set r of variables where each variable $X_i \in r$ is referenced by its index $i \in \mathbb{N}$. Conversely, every index $i \in \mathbb{N}$ refers to a unique variable $X_i \in r$. This one-to-one correspondence between variables and indexes allows us subsequently to identify the two concepts. Further, we assume for every variable $X \in r$ a set Ω_X of possible values, called its *frame*. Such variable frames are sometimes assumed to be finite or countable, but this is not a general requirement. If a frame contains exactly two elements, the corresponding variable is said to be *binary*. Moreover, if the two elements represent the states *true* and *false*, the variable is called *propositional* or *Boolean*. A *tuple* or *configuration* with finite domain $s \subseteq r$ is a function

$$\mathbf{x} : s \rightarrow \Omega_s = \prod_{X \in s} \Omega_X$$

that associates a value $x(X) \in \Omega_X$ with each variable $X \in s$. By convention, the single tuple with empty domain is identified by the diamond symbol \diamond , and we use bold-face, lower-case letters $\mathbf{x}, \mathbf{y}, \dots$ to refer to tuples. Subsequently, we write $\mathbf{x} \in \Omega_s$ to mark \mathbf{x} as a tuple with domain s , for which we also use the shorthand term *s-tuple* or *s-configuration*.

It is often convenient to decompose tuples according to some variable partition. This operation is also called *projection*, although it is not directly related to valuation algebras. Given an s -tuple \mathbf{x} and $t \subseteq s$, the projection of \mathbf{x} to t is defined by a t -tuple \mathbf{y} such that $\mathbf{y}(X) = \mathbf{x}(X)$ for all $X \in t$. We subsequently write $\mathbf{x}^{\downarrow t}$ for the projection of \mathbf{x} to t . Note that \diamond denotes the projection of any tuple to the empty set. Further, this notation allows us to write $\mathbf{x} = (\mathbf{x}^{\downarrow t}, \mathbf{x}^{\downarrow s-t})$ for the decomposition of \mathbf{x} with respect to t and $s - t$. Observe also that $(\mathbf{x}, \diamond) = (\diamond, \mathbf{x}) = \mathbf{x}$. Similar to frames of single variables, the set Ω_s represents all possible s -tuples and is therefore called the *frame* of the variable or index set s . In particular, $\Omega_\emptyset = \{\diamond\}$. Then, a *tuple set* with domain s is simply a subset $S \subseteq \Omega_s$. If a tuple set consists of only one element, then it is also called a *singleton*.

Example 1.1 To describe the attributes color, speed and prize of a car, we assume the set of variables $r = \{C, S, P\}$ with variable frames: $\Omega_C = \{\text{red, blue, black}\}$, $\Omega_S = \{\text{slow, fast}\}$ and $\Omega_P = \mathbb{N}$. Two possible r -tuples are $\mathbf{x} = (\text{blue, fast, } 30'000)$ and $\mathbf{y} = (\text{black, slow, } 10'000)$. The tuple \mathbf{x} can for example be decomposed into $\mathbf{x}^{\downarrow\{C,P\}} = (\text{blue, } 30'000)$ and $\mathbf{x}^{\downarrow\{S\}} = (\text{fast})$. Further, the variable set $\{C, S\}$ possesses the frame $\Omega_{\{C,S\}} = \{(\text{red, slow}), (\text{blue, slow}), \dots, (\text{black, fast})\}$. Observe also that the frame Ω_r contains infinitely many tuples.

If $X \in r$ is a variable that takes real numbers, we have $\Omega_X = \mathbb{R}$, and for a set $s \subseteq r$ of real variables the *linear space* \mathbb{R}^s corresponds to the frame of all s -tuples. In this case, s -tuples are also called *s-vectors*. It is possible to introduce an *extension operator* that lifts an s -vector \mathbf{x} to some larger domain $t \subseteq s$ by assigning zeros to all variables in $t - s$. We thus have $\mathbf{x}^{\uparrow t}(X) = \mathbf{x}(X)$ if $X \in s$ and $\mathbf{x}^{\uparrow t}(X) = 0$ otherwise. Clearly, the introduction of this operation is not only possible for real number but for all algebraic structures that contain a *zero element*. These definitions of frames, tuples and tuple sets allow for a uniform notation in the following tour through a first collection of valuation algebra instances.

■ 1.1 Indicator Functions - Boolean Functions - Crisp Constraints

An *indicator function* with domain $s \subseteq r$ identifies a subset $S \subseteq \Omega_s$ by specifying for each tuple $\mathbf{x} \in \Omega_s$ whether \mathbf{x} belongs to S or not. If we adopt the usual interpretation of 0 for \mathbf{x} not being an element of S and 1 for \mathbf{x} being in S , an indicator function i is defined by

$$i(\mathbf{x}) = \begin{cases} 0 & \text{if } \mathbf{x} \notin S, \\ 1 & \text{if } \mathbf{x} \in S. \end{cases}$$

Thus, an indicator i with domain $d(i) = s$ is a function that maps every tuple $\mathbf{x} \in \Omega_s$ onto a value $i(\mathbf{x}) \in \{0, 1\}$, i.e. $i : \Omega_s \rightarrow \{0, 1\}$. These are the valuations in the valuation algebra of indicator functions. As introduced in Section 1.1, we usually write Φ_s for the set of all indicator functions with domain s and Φ for the set of all possible indicator functions over subsets of r . Combination of indicator functions is defined by multiplication. If i_1 and i_2 are indicator functions with domain s and t respectively, we define for $\mathbf{x} \in \Omega_{s \cup t}$

$$i_1 \otimes i_2(\mathbf{x}) = i_1(\mathbf{x}^{\downarrow s}) \cdot i_2(\mathbf{x}^{\downarrow t}).$$

Alternatively, we may also define combination in terms of minimization:

$$i_1 \otimes i_2(\mathbf{x}) = \min\{i_1(\mathbf{x}^{\downarrow s}), i_2(\mathbf{x}^{\downarrow t})\}.$$

Projection, on the other hand, corresponds to maximization. For an indicator i with domain s and $t \subseteq s$ we define for all $\mathbf{x} \in \Omega_t$

$$i^{\downarrow t}(\mathbf{x}) = \max_{\mathbf{y} \in \Omega_{s-t}} i(\mathbf{x}, \mathbf{y}).$$

If we consider only variables with finite frames, an indicator function with domain s can be represented by an $|s|$ -dimensional table with $|\Omega_s|$ zero-one entries. Below, we simply use a *relational* or *tabular representation*, but emphasize that indicator functions are not limited to this inefficient representation. In order to prove that indicator functions form a valuation algebra, the reader may verify that all axioms are satisfied. However, in Section 5.3 this proof will be made for a whole family of formalisms that also includes indicator functions. Otherwise, a direct proof can be found in (Kohlas, 2003). Depending on its application field, the formalism of indicator functions has different names. In the context of constraint reasoning, they are usually called *crisp constraints*. Otherwise, indicator functions for propositional variables are also called *Boolean functions*.

For a concrete example of computing with indicator functions, consider a set of variables $r = \{A, B, C\}$ with frames $\Omega_A = \{a, \bar{a}\}$, $\Omega_B = \{b, \bar{b}\}$ and $\Omega_C = \{c, \bar{c}\}$. Then, assume two indicator functions i_1 and i_2 with domains $d(i_1) = \{A, B\}$ and $d(i_2) = \{B, C\}$:

$$i_1 = \begin{array}{|c|c|c|} \hline \mathbf{A} & \mathbf{B} & \\ \hline a & b & 0 \\ a & \bar{b} & 1 \\ \bar{a} & b & 0 \\ \bar{a} & \bar{b} & 0 \\ \hline \end{array} \quad i_2 = \begin{array}{|c|c|c|} \hline \mathbf{B} & \mathbf{C} & \\ \hline b & c & 1 \\ b & \bar{c} & 0 \\ \bar{b} & c & 1 \\ \bar{b} & \bar{c} & 1 \\ \hline \end{array}$$

We combine i_1 and i_2 and project the result to $\{A\}$:

$$i_3 = i_1 \otimes i_2 =$$

A	B	C	
a	b	c	0
a	b	\bar{c}	0
a	\bar{b}	c	1
a	\bar{b}	\bar{c}	1
\bar{a}	b	c	0
\bar{a}	b	\bar{c}	0
\bar{a}	\bar{b}	c	0
\bar{a}	\bar{b}	\bar{c}	0

$$i_3^{\downarrow\{A\}} =$$

A	
a	1
\bar{a}	0

■ 1.2 Relational Algebra

The second instance we are going to study is closely related to the first one. It is a subset of the *relational algebra* (Maier, 1983; Ullman, 1988), which traditionally belongs to the most fundamental formalisms for representing knowledge and information. In its usual extent, at least six operations are provided to manipulate knowledge represented as sets of tuples. Respecting the language of relational database theory, variables are called *attributes* and sets of tuples are called *relations*. A relation over $s \subseteq \tau$ is therefore simply a tuple set $R \subseteq \Omega_s$. It is important to note that variable frames do not need to be finite and consequently, relations can also be infinite sets. Combination is defined by *natural join*: If R_1 and R_2 are relations with domain s and t respectively, we define

$$R_1 \bowtie R_2 = \{\mathbf{x} \in \Omega_{s \cup t} : \mathbf{x}^{\downarrow s} \in R_1, \mathbf{x}^{\downarrow t} \in R_2\}. \quad (1.8)$$

Projection is defined for a relation R with domain s and $t \subseteq s$ as

$$R^{\downarrow t} = \{\mathbf{x}^{\downarrow t} : \mathbf{x} \in R\}$$

Alternatively, the notation $\pi_t(R) = R^{\downarrow t}$ is also used in the context of relational algebras. If we consider only finite variables in τ , then the relational algebra over τ is isomorphic to the valuation algebra of indicator functions over τ . In fact, relations are just another representation of indicators. We obtain the relation R_ϕ associated to the indicator function ϕ with domain $d(\phi) = s$ by

$$R_\phi = \{\mathbf{x} \in \Omega_s : \phi(\mathbf{x}) = 1\}. \quad (1.9)$$

Conversely, we derive the indicator function ϕ associated to a relation R with domain $s \subseteq \tau$ by $\phi(\mathbf{x}) = 1$ if $\mathbf{x} \in R$, and $\phi(\mathbf{x}) = 0$ otherwise, for all $\mathbf{x} \in \Omega_s$. Hence, we directly conclude that relations form a valuation algebra in the case of finite variables. This statement also holds for the general case as is will be shown in Chapter 7.3. Let us write the combination axiom for $z = x$ in the notation of relational algebra: For $R_1, R_2 \in \Phi$ with $d(R_1) = x$ and $d(R_2) = y$,

$$\pi_x(R_1 \bowtie R_2) = R_1 \bowtie \pi_{x \cap y}(R_2).$$

The right-hand side is called *semi-join* in database theory (Maier, 1983).

To give an example of computing in the relational algebra, we consider two relations R_1 and R_2 with domain $d(R_1) = \{\textit{continent}, \textit{country}\}$ and $d(R_2) = \{\textit{country}, \textit{city}\}$:

$R_1 =$	<table border="1" style="border-collapse: collapse; text-align: center;"><thead><tr><th>continent</th><th>country</th></tr></thead><tbody><tr><td>Africa</td><td>Kenya</td></tr><tr><td>Asia</td><td>China</td></tr><tr><td>Asia</td><td>Japan</td></tr><tr><td>Europe</td><td>Germany</td></tr><tr><td>Europe</td><td>France</td></tr></tbody></table>	continent	country	Africa	Kenya	Asia	China	Asia	Japan	Europe	Germany	Europe	France
	continent	country											
	Africa	Kenya											
	Asia	China											
	Asia	Japan											
Europe	Germany												
Europe	France												

$R_2 =$	<table border="1" style="border-collapse: collapse; text-align: center;"><thead><tr><th>country</th><th>city</th></tr></thead><tbody><tr><td>France</td><td>Paris</td></tr><tr><td>France</td><td>Lyon</td></tr><tr><td>Germany</td><td>Berlin</td></tr><tr><td>Italy</td><td>Rome</td></tr><tr><td>Kenya</td><td>Nairobi</td></tr></tbody></table>	country	city	France	Paris	France	Lyon	Germany	Berlin	Italy	Rome	Kenya	Nairobi
	country	city											
	France	Paris											
	France	Lyon											
	Germany	Berlin											
Italy	Rome												
Kenya	Nairobi												

Combining R_1 and R_2 gives

$R_3 = R_1 \bowtie R_2 =$	<table border="1" style="border-collapse: collapse;"><thead><tr><th>continent</th><th>country</th><th>city</th></tr></thead><tbody><tr><td>Africa</td><td>Kenya</td><td>Nairobi</td></tr><tr><td>Europe</td><td>Germany</td><td>Berlin</td></tr><tr><td>Europe</td><td>France</td><td>Paris</td></tr><tr><td>Europe</td><td>France</td><td>Lyon</td></tr></tbody></table>	continent	country	city	Africa	Kenya	Nairobi	Europe	Germany	Berlin	Europe	France	Paris	Europe	France	Lyon
continent	country	city														
Africa	Kenya	Nairobi														
Europe	Germany	Berlin														
Europe	France	Paris														
Europe	France	Lyon														

and we obtain for the projection of R_3 to $\{\textit{continent}\}$:

$\pi_{\{\textit{continent}\}}(R_3) = R_3 \downarrow^{\{\textit{continent}\}} =$	<table border="1" style="border-collapse: collapse;"><thead><tr><th>continent</th></tr></thead><tbody><tr><td>Africa</td></tr><tr><td>Europe</td></tr></tbody></table>	continent	Africa	Europe
continent				
Africa				
Europe				

■ 1.3 Arithmetic Potentials - Probability Potentials

Probability potentials are perhaps the most cited example of a valuation algebra, and their common algebraic structure with belief functions was originally the guiding theme for the abstraction process that lead to the valuation algebra framework (Shafer & Shenoy, 1988). However, at this point we yet ignore their usual interpretation as discrete probability mass functions and refer to this formalisms as *arithmetic potentials*. These are simple mappings that associate a non-negative real value with each tuple, i.e. $p : \Omega_s \rightarrow \mathbb{R}_{\geq 0}$. Also, we consider from the outset only variables with finite frames and identify $d(p) = s \subseteq r$ to be the domain of the potential p . The tabular representation can again be used for arithmetic potentials. Combination of two potentials p_1 and p_2 with domain s and t is defined by

$$p_1 \otimes p_2(\mathbf{x}) = p_1(\mathbf{x}^{\downarrow s}) \cdot p_2(\mathbf{x}^{\downarrow t}) \tag{1.10}$$

for $\mathbf{x} \in \Omega_{s \cup t}$. Projection consists in summing up all variables to be eliminated. If arithmetic potentials are used to express discrete probability mass functions,

then this corresponds to the operation of *marginalization* in probability theory. For a potential p with domain s , $t \subseteq s$ and $\mathbf{x} \in \Omega_t$, we define

$$p^{\downarrow t}(\mathbf{x}) = \sum_{\mathbf{y} \in \Omega_{s-t}} p(\mathbf{x}, \mathbf{y}). \tag{1.11}$$

Arithmetic potentials belong to the same family of instances as indicator functions. Section 5.3 provides a generic proof that all formalisms of this family satisfy the valuation algebra axioms. Moreover, we will also conclude from this result that other sets (such as integers, rational or complex numbers) may be used instead of real numbers to define arithmetic potentials.

We pointed out in the introduction of this book that the distributive law from arithmetics is used to make the computational process of inference more efficient. Also, we said that this property is contained in the valuation algebra axioms. This becomes evident by writing the combination axiom in the notation of arithmetic potentials: For $p_1, p_2 \in \Phi$ with $d(p_1) = s$, $d(p_2) = t$ and $\mathbf{x} \in \Omega_s$

$$\sum_{\mathbf{y} \in \Omega_{t-s}} p_1(\mathbf{x}) \cdot p_2(\mathbf{x}^{\downarrow s \cap t}, \mathbf{y}) = p_1(\mathbf{x}) \cdot \sum_{\mathbf{y} \in \Omega_{t-s}} p_2(\mathbf{x}^{\downarrow s \cap t}, \mathbf{y}).$$

For an example of how to compute with arithmetic potentials, consider a set $r = \{A, B, C\}$ of three variables with finite frames $\Omega_A = \{a, \bar{a}\}$, $\Omega_B = \{b, \bar{b}\}$ and $\Omega_C = \{c, \bar{c}\}$. We define two arithmetic potentials p_1 and p_2 with domain $d(p_1) = \{A, B\}$ and $d(p_2) = \{B, C\}$:

$p_1 =$	<table border="1" style="border-collapse: collapse; text-align: center;"><tr><th>A</th><th>B</th><th></th></tr><tr><td>a</td><td>b</td><td>0.6</td></tr><tr><td>a</td><td>\bar{b}</td><td>0.4</td></tr><tr><td>\bar{a}</td><td>b</td><td>0.3</td></tr><tr><td>\bar{a}</td><td>\bar{b}</td><td>0.7</td></tr></table>	A	B		a	b	0.6	a	\bar{b}	0.4	\bar{a}	b	0.3	\bar{a}	\bar{b}	0.7
	A	B														
	a	b	0.6													
	a	\bar{b}	0.4													
\bar{a}	b	0.3														
\bar{a}	\bar{b}	0.7														

$p_2 =$	<table border="1" style="border-collapse: collapse; text-align: center;"><tr><th>B</th><th>C</th><th></th></tr><tr><td>b</td><td>c</td><td>0.2</td></tr><tr><td>b</td><td>\bar{c}</td><td>0.8</td></tr><tr><td>\bar{b}</td><td>c</td><td>0.9</td></tr><tr><td>\bar{b}</td><td>\bar{c}</td><td>0.1</td></tr></table>	B	C		b	c	0.2	b	\bar{c}	0.8	\bar{b}	c	0.9	\bar{b}	\bar{c}	0.1
	B	C														
	b	c	0.2													
	b	\bar{c}	0.8													
\bar{b}	c	0.9														
\bar{b}	\bar{c}	0.1														

We combine p_1 and p_2 and project the result to $\{A, C\}$:

$p_3 = p_1 \otimes p_2 =$	<table border="1" style="border-collapse: collapse; text-align: center;"><tr><th>A</th><th>B</th><th>C</th><th></th></tr><tr><td>a</td><td>b</td><td>c</td><td>0.12</td></tr><tr><td>a</td><td>b</td><td>\bar{c}</td><td>0.48</td></tr><tr><td>a</td><td>\bar{b}</td><td>c</td><td>0.36</td></tr><tr><td>a</td><td>\bar{b}</td><td>\bar{c}</td><td>0.04</td></tr><tr><td>\bar{a}</td><td>b</td><td>c</td><td>0.06</td></tr><tr><td>\bar{a}</td><td>b</td><td>\bar{c}</td><td>0.24</td></tr><tr><td>\bar{a}</td><td>\bar{b}</td><td>c</td><td>0.63</td></tr><tr><td>\bar{a}</td><td>\bar{b}</td><td>\bar{c}</td><td>0.07</td></tr></table>	A	B	C		a	b	c	0.12	a	b	\bar{c}	0.48	a	\bar{b}	c	0.36	a	\bar{b}	\bar{c}	0.04	\bar{a}	b	c	0.06	\bar{a}	b	\bar{c}	0.24	\bar{a}	\bar{b}	c	0.63	\bar{a}	\bar{b}	\bar{c}	0.07
	A	B	C																																		
	a	b	c	0.12																																	
	a	b	\bar{c}	0.48																																	
	a	\bar{b}	c	0.36																																	
	a	\bar{b}	\bar{c}	0.04																																	
	\bar{a}	b	c	0.06																																	
	\bar{a}	b	\bar{c}	0.24																																	
\bar{a}	\bar{b}	c	0.63																																		
\bar{a}	\bar{b}	\bar{c}	0.07																																		

$p_3^{\downarrow \{A, C\}} =$	<table border="1" style="border-collapse: collapse; text-align: center;"><tr><th>A</th><th>C</th><th></th></tr><tr><td>a</td><td>c</td><td>0.48</td></tr><tr><td>a</td><td>\bar{c}</td><td>0.52</td></tr><tr><td>\bar{a}</td><td>c</td><td>0.69</td></tr><tr><td>\bar{a}</td><td>\bar{c}</td><td>0.31</td></tr></table>	A	C		a	c	0.48	a	\bar{c}	0.52	\bar{a}	c	0.69	\bar{a}	\bar{c}	0.31
	A	C														
	a	c	0.48													
	a	\bar{c}	0.52													
\bar{a}	c	0.69														
\bar{a}	\bar{c}	0.31														

■ 1.4 Set Potentials - Belief Functions

As we already mentioned, *belief functions* (Shafer, 1976; Kohlas & Monney, 1995) rank among the few instances that originally initiated the abstraction process culminating in the valuation algebra framework (Shafer & Shenoy, 1988). Belief functions are special cases of *set potentials*. Therefore, we first focus on this slightly more general formalism and refer to Instance D.9 and D.6 for a discussion of belief functions and its alternative representations. The theory is again restricted to variables with finite frames. A set potential $m : \mathcal{P}(\Omega_s) \rightarrow \mathbb{R}_{\geq 0}$ with domain $d(m) = s \subseteq r$ is a function that assigns non-negative real numbers to all subsets of the variable set frame Ω_s . Combination of two set potentials m_1 and m_2 with domains s and t is defined for all tuple sets $A \subseteq \Omega_{s \cup t}$ by

$$m_1 \otimes m_2(A) = \sum_{B \bowtie C = A, B \subseteq \Omega_s, C \subseteq \Omega_t} m_1(B) \cdot m_2(C). \tag{1.12}$$

This is a simplified version of Dempster’s rule of combination (Dempster, 1968; Shafer, 1991). Projection of a set potential m with domain s to $t \subseteq s$ is given for $A \subseteq \Omega_t$ by

$$m^{\downarrow t}(A) = \sum_{\pi_t(B) = A, B \subseteq \Omega_s} m(B). \tag{1.13}$$

Observe that these definitions use the language from the relational algebra to deal with tuple sets. This particular notation will later be helpful. Since relations satisfy the valuation algebra axioms, we may call in its properties to derive important results about set potentials. Note also that the formalism of set potentials essentially differs from the other instances introduced beforehand that map single tuples to some value. Indeed, it belongs to a second family of valuation algebras that will be introduced in Section 5.7 where we also provide a generic proof that all formalisms of this family satisfy the valuation algebra axioms. Before presenting a concrete example, it is important to remark that although all variable frames are finite, a simple enumeration of all assignments as proposed for indicator functions or arithmetic potentials is beyond question since set potentials assign values to all elements of the powerset of Ω_s . Instead, only those entries are listed whose values differ from zero. These tuple sets are usually called *focal sets*.

Consider a set of two variables $r = \{A, B\}$ with finite frames $\Omega_A = \{a, \bar{a}\}$ and $\Omega_B = \{b, \bar{b}\}$. We define two set potentials m_1 and m_2 with domains $d(m_1) = \{A, B\}$ and $d(m_2) = \{A\}$:

$$m_1 = \begin{array}{|l|l} \{(a, b)\} & 0.7 \\ \{(\bar{a}, b), (a, \bar{b})\} & 0.1 \\ \{(a, \bar{b}), (\bar{a}, \bar{b})\} & 0.2 \end{array} \quad m_2 = \begin{array}{|l|l} \{(\bar{a})\} & 0.6 \\ \{(a)\} & 0.4 \end{array}$$

The task of computing the combination of m_1 and m_2 is simplified by constructing the following table as an intermediate step. The first column contains m_1 and the top row m_2 , both extended to the union domain $d(m_1) \cup d(m_2) = \{A, B\}$. Then, every internal cell contains the intersection between the two corresponding tuple sets and the product of the corresponding values. This corresponds to the natural join in equation (1.12).

	$\{(\bar{a}, \bar{b}), (\bar{a}, b)\}, 0.6$	$\{(a, \bar{b}), (a, b)\}, 0.4$
$\{(a, b)\}, 0.7$	$\emptyset, 0.42$	$\{(a, b)\}, 0.28$
$\{(\bar{a}, b), (a, \bar{b})\}, 0.1$	$\{(\bar{a}, b)\}, 0.06$	$\{(a, \bar{b})\}, 0.04$
$\{(a, b), (\bar{a}, \bar{b})\}, 0.2$	$\{(\bar{a}, \bar{b})\}, 0.12$	$\{(a, b)\}, 0.08$

To complete the combination, it is then sufficient to add the values of all internal cells with equal tuple set. The result of the combination is projected afterwards to $\{A\}$ using equation (1.13)

$$m_3 = m_1 \otimes m_2 = \begin{array}{|l|l|} \hline \emptyset & 0.42 \\ \hline \{(a, b)\} & 0.36 \\ \{(a, \bar{b})\} & 0.04 \\ \{(\bar{a}, b)\} & 0.06 \\ \{(\bar{a}, \bar{b})\} & 0.12 \\ \hline \end{array} \quad m_3^{\downarrow\{A\}} = \begin{array}{|l|l|} \hline \emptyset & 0.42 \\ \{(\bar{a})\} & 0.18 \\ \{(a)\} & 0.40 \\ \hline \end{array}$$

■ 1.5 Density Functions

All instances discussed so far are based on variables with finite (or at least countable) frames. Although such a setting is typical for a large number of formalisms used in computer science, it is in no way mandatory. Thus, we next introduce the valuation algebra of *density functions* which are defined over variables taking values from the set of real numbers. For a set of variables $s \subseteq r$, a density function $f : \mathbb{R}^s \rightarrow \mathbb{R}_{\geq 0}$ with domain s is a continuous, non-negative valued function with finite Riemann integral,

$$\int_{-\infty}^{\infty} f(\mathbf{x}) d\mathbf{x} < \infty. \quad (1.14)$$

Note that \mathbf{x} denotes a vector in the linear space $\Omega_s = \mathbb{R}^s$. The combination $f \otimes g$ of two density functions f and g with domain s and t is defined for $\mathbf{x} \in \mathbb{R}^{s \cup t}$ by simple multiplication

$$f \otimes g(\mathbf{x}) = f(\mathbf{x}^{\downarrow s}) \cdot g(\mathbf{x}^{\downarrow t}). \quad (1.15)$$

The projection $f^{\downarrow t}$ of a density f with $d(f) = s$ to $t \subseteq s$ is given by the integral

$$f^{\downarrow t}(\mathbf{x}) = \int_{-\infty}^{\infty} f(\mathbf{x}, \mathbf{y}) d\mathbf{y} \quad (1.16)$$

where $x \in \mathbb{R}^t$ and $y \in \mathbb{R}^{s-t}$. Density functions together with the above definitions of combination and projection form a valuation algebra. The proof of this statement even for the more general case of Lebesgue measurable density functions can be found in (Kohlas, 2003). The perhaps most interesting example of a density function in the context of valuation algebras is the *Gaussian density*. In fact, Instance 1.6 below shows that Gaussian densities are closed under combination and projection and therefore establish a subalgebra of the valuation algebra of density functions.

Labeled Matrices

These first examples were all based on tuples and tuple sets. Next, we are going to introduce two further valuation algebra instances based on *labeled matrices*. For finite sets $s, t \subseteq r$ a labeled, real-valued matrix is a mapping $\mathbf{M} : s \times t \rightarrow \mathbb{R}$ such that for $X \in s$ and $Y \in t$ we have $\mathbf{M}(X, Y) \in \mathbb{R}$. We write $\mathcal{M}(\mathbb{R}, s \times t)$ for the set of matrices of this kind. The sum of $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{M}(\mathbb{R}, s \times t)$ is defined by

$$(\mathbf{M}_1 + \mathbf{M}_2)(X, Y) = \mathbf{M}_1(X, Y) + \mathbf{M}_2(X, Y) \quad (1.17)$$

for $X \in s$ and $Y \in t$. If $\mathbf{M}_1 \in \mathcal{M}(\mathbb{R}, s \times t)$ and $\mathbf{M}_2 \in \mathcal{M}(\mathbb{R}, t \times u)$ for $s, t, u \subseteq r$, $X \in s$ and $Y \in u$ we define the product $\mathbf{M}_1 \cdot \mathbf{M}_2$ by

$$(\mathbf{M}_1 \cdot \mathbf{M}_2)(X, Y) = \sum_{Z \in t} \mathbf{M}_1(X, Z) \cdot \mathbf{M}_2(Z, Y). \quad (1.18)$$

A special element of the set $\mathcal{M}(\mathbb{R}, s \times t)$ is the *zero matrix* defined for $X \in s$ and $Y \in t$ by $\mathbf{O}(X, Y) = 0$. The *projection* of a labeled matrix $\mathbf{M} \in \mathcal{M}(\mathbb{R}, s \times t)$ to $u \subseteq s$ and $v \subseteq t$ is defined by $\mathbf{M}^{\downarrow u, v}(X, Y) = \mathbf{M}(X, Y)$ if $X \in u$ and $Y \in v$, which simply corresponds to dropping the unrequested rows and columns. Similarly to the extension of vectors, we *extend* a labeled matrices $\mathbf{M} \in \mathcal{M}(\mathbb{R}, s \times t)$ to $u \supseteq s$ and $v \supseteq t$ with $X \in u$ and $Y \in v$ by

$$\mathbf{M}^{\uparrow u, v}(X, Y) = \begin{cases} \mathbf{M}(X, Y) & \text{if } X \in s \text{ and } Y \in t, \\ 0, & \text{otherwise.} \end{cases} \quad (1.19)$$

Of special interest are often square matrices $\mathbf{M} \in \mathcal{M}(\mathbb{R}, s \times s)$. We then say that the *domain* of \mathbf{M} is $d(\mathbf{M}) = s$ and use the abbreviation $\mathcal{M}(\mathbb{R}, s)$ for the set of all matrices with domain s . Also, the projection of a matrix \mathbf{M} with domain s to $t \subseteq s$ is abbreviated by $\mathbf{M}^{\downarrow s}$ and accordingly, we write $\mathbf{M}^{\uparrow u}$ for its extension to the larger domain $u \supseteq s$. Finally, we define the *identity matrix* for the domain s by $\mathbf{I}(X, Y) = 1$ if $X = Y$ and $\mathbf{I}(X, Y) = 0$ otherwise. In later chapters, we will also consider labeled matrices with values from other algebraic structures than real numbers. But for the introduction of the following instances, the limitation to real numbers is sufficient.

■ 1.6 Gaussian Densities - Gaussian Potentials

An important family of density functions is provided by *Gaussian distributions*. For a set of real variables $s \subseteq r$ and $\mathbf{x} \in \mathbb{R}^s$ a *Gaussian density* is defined by

$$f(\mathbf{x}) = \sqrt{\frac{|\det(\mathbf{K})|}{(2\pi)^{|s|}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{K}(\mathbf{x}-\boldsymbol{\mu})}, \quad (1.20)$$

where $\boldsymbol{\mu} \in \mathbb{R}^s$ denotes the *mean vector* and $\mathbf{K} : s \times s \rightarrow \mathbb{R}$ the symmetric and positive definite *concentration matrix*. The matrix \mathbf{K}^{-1} is the *variance-covariance matrix* of the Gaussian density. Looking at this definition, we remark that a Gaussian density is completely determined by its mean vector and concentration matrix. Such pairs $(\boldsymbol{\mu}, \mathbf{K})$ with $\boldsymbol{\mu} \in \mathbb{R}^s$ and $\mathbf{K} \in \mathcal{M}(\mathbb{R}, s)$ being symmetric and positive definite are called *Gaussian potentials*. The square root in front of the exponential function in equation (1.20) is merely a *normalization factor* guaranteeing that the integral over the density function equals one.

Let f and g be two Gaussian densities with domain $d(f) = s$ and $d(g) = t$, represented by the Gaussian potentials $(\boldsymbol{\mu}_1, \mathbf{K}_1)$ and $(\boldsymbol{\mu}_2, \mathbf{K}_2)$. When combining these two densities by equation (1.15) we can neglect the normalization constants and simply multiply the exponential functions. This in turn means adding their exponents. Doing so, we may also neglect additive terms without variables, since these terms simply go into the normalization factor. In the following, equality must always be taken up to such terms. We then get for the exponent of the combined density function and $\mathbf{x} \in \mathbb{R}^{s \cup t}$

$$\begin{aligned} & (\mathbf{x}^{\downarrow s} - \boldsymbol{\mu}_1)^T \mathbf{K}_1 (\mathbf{x}^{\downarrow s} - \boldsymbol{\mu}_1) + (\mathbf{x}^{\downarrow t} - \boldsymbol{\mu}_2)^T \mathbf{K}_2 (\mathbf{x}^{\downarrow t} - \boldsymbol{\mu}_2) = \\ & (\mathbf{x} - \boldsymbol{\mu}_1^{\uparrow s \cup t})^T \mathbf{K}_1^{\uparrow s \cup t} (\mathbf{x} - \boldsymbol{\mu}_1^{\uparrow s \cup t}) + (\mathbf{x} - \boldsymbol{\mu}_2^{\uparrow s \cup t})^T \mathbf{K}_2^{\uparrow s \cup t} (\mathbf{x} - \boldsymbol{\mu}_2^{\uparrow s \cup t}) = \\ & \mathbf{x}^T \mathbf{K}_1^{\uparrow s \cup t} \mathbf{x} - \mathbf{x}^T \mathbf{K}_1^{\uparrow s \cup t} \boldsymbol{\mu}_1^{\uparrow s \cup t} - \boldsymbol{\mu}_1^{T \uparrow s \cup t} \mathbf{K}_1^{\uparrow s \cup t} \mathbf{x} + \\ & \mathbf{x}^T \mathbf{K}_2^{\uparrow s \cup t} \mathbf{x} - \mathbf{x}^T \mathbf{K}_2^{\uparrow s \cup t} \boldsymbol{\mu}_2^{\uparrow s \cup t} - \boldsymbol{\mu}_2^{T \uparrow s \cup t} \mathbf{K}_2^{\uparrow s \cup t} \mathbf{x} = \\ & \mathbf{x}^T (\mathbf{K}_1^{\uparrow s \cup t} + \mathbf{K}_2^{\uparrow s \cup t}) \mathbf{x} - \mathbf{x}^T (\mathbf{K}_1^{\uparrow s \cup t} \boldsymbol{\mu}_1^{\uparrow s \cup t} + \mathbf{K}_2^{\uparrow s \cup t} \boldsymbol{\mu}_2^{\uparrow s \cup t}) - \\ & (\boldsymbol{\mu}_1^{T \uparrow s \cup t} \mathbf{K}_1^{\uparrow s \cup t} + \boldsymbol{\mu}_2^{T \uparrow s \cup t} \mathbf{K}_2^{\uparrow s \cup t}) \mathbf{x}. \end{aligned}$$

Define now

$$\mathbf{K} = \mathbf{K}_1^{\uparrow s \cup t} + \mathbf{K}_2^{\uparrow s \cup t}.$$

Note that \mathbf{K} is still symmetric and positive definite. Then, we rewrite the last expression using \mathbf{K} , add a constant, additive term to the last expression

completing it to a quadratic form:

$$\begin{aligned}
& \mathbf{x}^T \mathbf{K} \mathbf{x} - \mathbf{x}^T (\mathbf{K}_1^{\uparrow s U t} \mu_1^{\uparrow s U t} + \mathbf{K}_2^{\uparrow s U t} \mu_2^{\uparrow s U t}) - \\
& \quad (\mathbf{K}_1^{\uparrow s U t} \mu_1^{\uparrow s U t} + \mathbf{K}_2^{\uparrow s U t} \mu_2^{\uparrow s U t})^T \mathbf{x} + \\
& (\mathbf{K}_1^{\uparrow s U t} \mu_1^{\uparrow s U t} + \mathbf{K}_2^{\uparrow s U t} \mu_2^{\uparrow s U t})^T \mathbf{K}^{-1} (\mathbf{K}_1^{\uparrow s U t} \mu_1^{\uparrow s U t} + \mathbf{K}_2^{\uparrow s U t} \mu_2^{\uparrow s U t}) = \\
& \quad (\mathbf{x} - \mathbf{K}^{-1} (\mathbf{K}_1^{\uparrow s U t} \mu_1^{\uparrow s U t} + \mathbf{K}_2^{\uparrow s U t} \mu_2^{\uparrow s U t}))^T \mathbf{K} \\
& \quad (\mathbf{x} - \mathbf{K}^{-1} (\mathbf{K}_1^{\uparrow s U t} \mu_1^{\uparrow s U t} + \mathbf{K}_2^{\uparrow s U t} \mu_2^{\uparrow s U t}))
\end{aligned}$$

This finally means that the multiplication of two Gaussian densities results again in a Gaussian density with concentration matrix

$$\mathbf{K} = \mathbf{K}_1^{\uparrow s U t} + \mathbf{K}_2^{\uparrow s U t}. \quad (1.21)$$

and mean vector

$$\mu = \mathbf{K}^{-1} (\mathbf{K}_1^{\uparrow s U t} \mu_1^{\uparrow s U t} + \mathbf{K}_2^{\uparrow s U t} \mu_2^{\uparrow s U t}). \quad (1.22)$$

Motivated by this result, we may define the combination of two Gaussian potentials (μ_1, \mathbf{K}_1) and (μ_2, \mathbf{K}_2) with $d(\mu_1, \mathbf{K}_1) = s$ and $d(\mu_2, \mathbf{K}_2) = t$ by

$$(\mu_1, \mathbf{K}_1) \otimes (\mu_2, \mathbf{K}_2) = (\mu, \mathbf{K}),$$

where μ is defined by (1.22) and \mathbf{K} by (1.21). We found this combination rule by a purely formal derivation. Its reason and meaning will be given in Section 10.1. The projection of a Gaussian density f with domain $d(f) = s$, mean vector μ and concentration matrix \mathbf{K} to a subset $t \subseteq s$ is again a Gaussian density, as it will be shown in Appendix J.2. Its mean vector is simply $\mu^{\downarrow t}$ and $((\mathbf{K}^{-1})^{\downarrow t})^{-1}$ is its concentration matrix. Consequently, Gaussian densities are closed under combination and projection and form an important subalgebra of the valuation algebra of density functions. The operations of combination and projection can both be expressed in terms of the Gaussian potentials associated with the densities. Gaussian densities and Gaussian potentials will be studied in detail in Chapter 10, which will also clarify the interest in this valuation algebra.

Many further valuation algebra instances will be introduced throughout the subsequent chapters. At this point we interrupt the study of formalisms for the time being and focus on the computational interest in valuation algebras. The following chapter phrases an abstract computational task called inference problem using the language of valuation algebras and also gives a first impression of the different semantics of this problem under different valuation algebra instances. But first, we propose a more far-reaching discussion of the valuation algebra axiomatics in the appendix of this chapter. The valuation algebra framework presented in Section 1.1 is more general than most of its antecessors. Nevertheless, there is still potential for further generalizations. Two such systems will be presented in the following appendix: the first

introduces valuation algebras that are no more necessarily based on variable systems. Instead, valuations take their domains from a more general lattice structure. The second framework focuses on a generalization of the projection operator. Here, it is no longer the case that projections on arbitrary subsets are possible, but there is a fourth operator that tells us to which domain a valuation is allowed to be projected. For the understanding of the subsequent chapters, these systems are of minor importance. But it is nevertheless interesting to see that further instances can be covered by a more general and abstract definition of the valuation algebra framework, which still provides enough structure for the application of local computation.

1.3 CONCLUSION

This first chapter introduced the valuation algebra framework upon which all following chapters are based. Valuations can be imagined as pieces of information that refer to a question called the domain of the valuation. This domain is returned by the labeling operation. Two further operations called combination and projection are used to manipulate valuations. Combination corresponds to aggregation, and projection to focussing or extraction of knowledge. In addition, the valuation algebra framework consists of a set of six axioms that determine the behaviour of the three operations. Formalisms that satisfy the structure of a valuation algebra are called instances and occur numerous in very different fields of mathematics and computer science. This chapter gave a first selection of instances including crisp constraints, arithmetic and probability potentials, Dempster-Shafer belief functions, density functions and the important family of Gaussian distributions.

Appendix: Generalizations of the Valuation Algebra Framework

We first give a formal definition of lattices with their main variations and refer to (Davey & Priestley, 1990) for an extensive discussion of these and related concepts.

A.1 ORDERED SETS AND LATTICES

Definition A.2 A preorder is a binary relation \leq over a set P which is reflexive and transitive. We have for $a, b, c \in P$:

- $a \leq a$ (reflexivity);
- $a \leq b$ and $b \leq c$ implies that $a \leq c$ (transitivity).

A preorder is called partial order if it is antisymmetric, i.e. if

- $a \leq b$ and $b \leq a$ implies that $a = b$ (antisymmetry).

A set with a partial order is also called a *partially ordered set* or simply an *ordered set*. Chapter 5 starts with an introduction to semiring theory that naturally offers a

large number of examples for preorders and partial orders. We therefore refer to the numerous examples of semirings in Chapter 5 for concrete sets providing preorders and partial orders.

Definition A.3 Let P be an ordered set.

- P has a bottom element if there exists an element $\perp \in P$ such that $\perp \leq x$ for all $x \in P$.
- P has a top element if there exists an element $\top \in P$ such that $x \leq \top$ for all $x \in P$.

Definition A.4 Let P be an ordered set and $S \subseteq P$. An element $u \in P$ is called supremum, least upper bound or join of S if

- $x \leq u$ for all $x \in S$;
- for any $v \in P$ such that $x \leq v$ for all $x \in S$ it holds that $u \leq v$.

Likewise, an element $u \in P$ is called infimum, greatest lower bound or meet of S if

- $u \leq x$ for all $x \in S$;
- for any $v \in P$ such that $v \leq x$ for all $x \in S$ it holds that $v \leq u$.

If join or meet of a subset $S \subseteq P$ exist, then they are always unique, and we write $\bigvee S$ for join and $\bigwedge S$ for meet. Moreover, if $S = \{x, y\}$ consists of two elements, we generally write $x \vee y$ or $\sup\{x, y\}$ for join and $x \wedge y$ or $\inf\{x, y\}$ for meet.

Definition A.5 Let P be a non-empty, ordered set.

- If $x \vee y$ and $x \wedge y$ exist for all $x, y \in P$, then P is called a lattice.
- If $\bigvee S$ and $\bigwedge S$ exist for all subsets $S \subseteq P$, then P is called a complete lattice.

Lattices may satisfy further identities:

Definition A.6 Let L be a lattice.

- L is said to be bounded, if it has a bottom and a top element.
- L is said to be distributive, if for all $a, b, c \in L$

$$\begin{aligned} a \wedge (b \vee c) &= (a \wedge b) \vee (a \wedge c); \\ a \vee (b \wedge c) &= (a \vee b) \wedge (a \vee c). \end{aligned}$$

This definition lists both statements of distributivity, although they are in fact equivalent. The proof is given in (Davey & Priestley, 1990). Note also that every complete lattice is bounded.

Example A.2 (Powerset or Subset Lattice) For any set A we consider the set of all subsets called its powerset or subset lattice $\mathcal{P}(A)$. This set is partially ordered via set inclusion and forms a lattice with set intersection as meet $a \wedge b = a \cap b$ and set union as join $a \vee b = a \cup b$. The powerset lattice is distributive, complete and bounded by A itself and the empty set \emptyset .

Example A.3 (Division Lattice) The set of $\mathbb{N} \cup \{0\}$ also forms a distributive lattice with the least common multiple as join $\text{lcm}(a, b) = a \vee b$ and the greatest common divisor as meet $\text{gcd}(a, b) = a \wedge b$. The order is defined by divisibility, i.e. $a \leq b$ if a divides b . This lattice is also bounded with 1 as bottom and 0 as top element.

An important source of non-distributive lattices are partitions:

A.1.1 Partitions and Partition Lattices

Partition lattices take a universal position among all lattices. (Grätzer, 1978) shows that every lattice is isomorph to some partition lattice.

Definition A.7 A partition $\pi = \{B_i : 1 \leq i \leq n\}$ of a set U called universe consists of a collection of subsets $B_i \subseteq U$ called blocks such that

- $B_i \neq \emptyset$;
- $B_i \cap B_j = \emptyset$ for $i \neq j$ and $1 \leq i, j \leq n$;
- $\bigcup_{i=1}^n B_i = U$.

Let $\text{Part}(U)$ denote the set of all possible partitions of a universe U . It is then possible to introduce a partial order between its elements. For $\pi_1, \pi_2 \in \text{Part}(U)$ we write $\pi_1 \leq \pi_2$ if every block of π_1 is contained in some block of π_2 . This is the case if, and only if, every block of π_2 is a union of blocks from π_1 . We then also say that the partition π_1 is *finer* than π_2 , or conversely that π_2 is *coarser* than π_1 . The blocks of the infimum or meet of an arbitrary collection of partitions $P \subseteq \text{Part}(U)$ corresponds to the non-empty intersections of all blocks contained in the partitions of P . The definition of the supremum or join especially for the case of universes with an infinite number of elements is more involved. We refer to (Grätzer, 1978) for a discussion of this aspect. Here, we directly conclude that the set $\text{Part}(U)$ of all partitions of the universe U forms a complete lattice. The partition $\{\{u\} : u \in U\}$ consisting of all one-element subsets of U is the lower bound, and the partition $\{U\}$ is the upper bound. A lattice whose elements are partitions of a universe U is called *partition lattice*. They are sublattices of $\text{Part}(U)$ and generally not distributive. The above refinement relation is the natural way of introducing an order relation between partitions. However, in the context of algebraic information theory (Kohlas, 2003) it is often useful to consider the inverse of the natural order, i.e. $\pi_2 \leq_c \pi_1$ if, and only if, $\pi_1 \leq \pi_2$ or equivalently, if π_2 is coarser than π_1 . The motivation is that a finer partition expresses more information than a coarser partition.

Example A.4 (Interval Partitions) Let $U = [a, b)$ be a semi-closed interval of real numbers. A sequence $x_0 < x_1 < \dots < x_n$ with $x_0 = a$ and $x_n = b$ establishes a partition $\pi = \{[x_i, x_{i+1}) : 0 \leq i \leq n-1\}$. If an interval partition π_1 contains all elements of the sequence defining the partition π_2 , then every bloc of π_1 is contained in some bloc of π_2 . In other words, π_1 is a refinement of π_2 . The meet of two interval partitions is obtained from the union sequence, and the join from the intersection. If for example $U = [0, 1)$ the partition for $0 < 0.2 < 0.4 < 0.6 < 0.8 < 1$ is a refinement of the partition for $0 < 0.4 < 0.8 < 1$. If furthermore π_1 is given by the sequence $0 < 0.5 < 0.8 < 1$ and π_2 by $0 < 0.2 < 0.8 < 1$, their meet is obtained from the sequence $0 < 0.2 < 0.5 < 0.8 < 1$ and their join from $0 < 0.8 < 1$.

A.2 VALUATION ALGEBRAS ON GENERAL LATTICES

The definition of a valuation algebra given at the beginning of this chapter is based on a particular lattice, namely on the powerset lattice of variables (see Example A.2). This lattice is distributive. However, it turns out that the valuation algebra framework can even be generalized to arbitrary lattices (Shafer, 1991; Kohlas, 2003). Thus, let D be a lattice with a partial order \leq and the two operations meet \wedge and join \vee . We denote by Φ the set of valuations with domains in D and suppose the following three operations defined on Φ and D .

1. *Labeling*: $\Phi \rightarrow D; \phi \mapsto d(\phi)$,
2. *Combination*: $\Phi \times \Phi \rightarrow \Phi; (\phi, \psi) \mapsto \phi \otimes \psi$,
3. *Projection*: $\Phi \times D \rightarrow \Phi; (\phi, x) \mapsto \phi^{\perp x}$ for $x \leq d(\phi)$.

We impose the following set of axioms on Φ and D :

(A1') *Commutative Semigroup*: Φ is associative and commutative under \otimes .

(A2') *Labeling*: For $\phi, \psi \in \Phi$,

$$d(\phi \otimes \psi) = d(\phi) \vee d(\psi). \quad (\text{A.1})$$

(A3') *Projection*: For $\phi \in \Phi, x \in D$ and $x \leq d(\phi)$,

$$d(\phi^{\perp x}) = x. \quad (\text{A.2})$$

(A4') *Transitivity*: For $\phi \in \Phi$ and $x \leq y \leq d(\phi)$,

$$(\phi^{\perp y})^{\perp x} = \phi^{\perp x}. \quad (\text{A.3})$$

(A5') *Combination*: For $\phi, \psi \in \Phi$ with $d(\phi) = x$, $d(\psi) = y$ and $z \in D$ such that $x \leq z \leq x \vee y$,

$$(\phi \otimes \psi)^{\downarrow z} = \phi \otimes \psi^{\downarrow z \wedge y}. \quad (\text{A.4})$$

(A6') *Domain*: For $\phi \in \Phi$ with $d(\phi) = x$,

$$\phi^{\downarrow x} = \phi. \quad (\text{A.5})$$

Clearly, this definition of a valuation algebra is more general than the framework introduced beforehand. It is indeed surprising that no further properties of the lattice are required to enable the application of local computation (Kohlas & Monney, 1995). We must, however, accept that not all properties of the more restricted concept of a valuation algebra are maintained. Although the subsequent chapters of this book are entirely based on variable systems, we nevertheless present an example from diagnostics that is based on a valuation algebra with domains from a more general lattice, i.e. from a lattice of partitions (Kohlas & Monney, 1995; Shafer *et al.*, 1987).



Figure A.1 A tree of diagnoses for stubborn cars.

Example A.5 Suppose that your car does not start anymore and you want to determine the cause of its failure. The tree of diagnoses shown in Figure A.1 provides a structural approach for the identification of the cause of failure. At the beginning, the list of possible diagnoses is fairly coarse and only contains the failure possibilities $\Omega_1 = \{a, b, c, d\}$. Note, however, that these options are mutually exclusive and collectively exhaustive. More precise statements are obtained by partitioning

coarse diagnoses. Doing so, we replace diagnosis a by $\{e, f\}$ and diagnosis c by $\{l, g, h\}$ which leads to the new set of failure possibilities $\Omega_2 = \{e, f, b, l, g, h, d\}$. After a third specialization step we finally obtain $\Omega_3 = \{e, f, b, i, j, k, g, h, d\}$. These step-wise refined sets of diagnoses are called frames of discernment. When a frame Ω_i is replaced by a finer frame Ω_{i+1} the substitution of a diagnosis by a set of more precise diagnoses is described by a mapping

$$\tau : \Omega_i \rightarrow \mathcal{P}(\Omega_{i+1}) \quad (\text{A.6})$$

called refinement mapping or simply refinement (Shafer, 1976). Refinement mappings must satisfy the following requirements:

1. $\tau(\theta) \neq \emptyset$ for all $\theta \in \Omega_i$;
2. $\tau(\theta_1) \cap \tau(\theta_2) = \emptyset$ whenever $\theta_1 \neq \theta_2$;
3. $\cup\{\tau(\theta) : \theta \in \Omega_i\} = \Omega_{i+1}$.

For example, the passage from Ω_1 to Ω_2 is expressed by the mapping

$$\tau(x) = \begin{cases} \{x\} & \text{if } x \in \{b, d\}, \\ \{e, f\} & \text{if } x = a, \\ \{l, g, h\} & \text{if } x = c. \end{cases}$$

From an alternative but equivalent point of view, this family of related frames can be seen as a collection of partitions of the universe Ω_3 . These partitions are:

- $\pi_0 = \{\{e, f, b, i, j, k, g, h, d\}\}$;
- $\pi_1 = \{\{e, f\}, \{b\}, \{i, j, k, g, h\}, \{d\}\}$;
- $\pi_2 = \{\{e\}, \{f\}, \{b\}, \{i, j, k\}, \{g\}, \{h\}, \{d\}\}$;
- $\pi_3 = \{\{e\}, \{f\}, \{b\}, \{i\}, \{j\}, \{k\}, \{g\}, \{h\}, \{d\}\}$.

Note that $\pi_{i+1} \leq \pi_i$, i.e. every block of π_{i+1} is contained in some block of π_i . Because each frame Ω_i corresponds to a partition π_i , we may replace the frames in equation (A.6) by their corresponding partitions. The mapping

$$\delta : \pi_i \rightarrow \mathcal{P}(\pi_{i+1})$$

assigns to each block $B \in \pi_i$ the set of blocks in π_{i+1} whose union is B . The passage from π_1 to π_2 is thus expressed by the mapping

$$\delta(x) = \begin{cases} \{\{x\}\} & \text{if } x = \{b, d\}, \\ \{\{e\}, \{f\}\} & \text{if } x = \{e, f\}, \\ \{\{i, j, k\}, \{g\}, \{h\}\} & \text{if } x = \{i, j, k, g, h\}. \end{cases}$$

This mapping is referred to as decomposition mapping.

Although many valuation algebras with domains from variable lattices could be adapted to partition and other lattices, the modified valuation algebra of set potentials from Dempster-Shafer theory is probably the most important example for practical applications. In fact, the derivation of partitions from a tree of diagnoses takes center stage in the applications of medical diagnostics in (Gordon & Shortliffe, 1985). Other authors studied Dempster-Shafer theory on unspecified lattices (Grabisch, 2009; Shafer, 1991). Here, we follow the algebraic approach from (Kohlas & Monney, 1995) and first extract the necessary requirements from the above example that later enables the definition of set potentials with domains from partition lattices.

Definition A.8 *A collection of frames \mathcal{F} together with a collection of refinements \mathcal{R} forms a family of compatible frames if the following conditions hold:*

1. *For each pair $\Omega_1, \Omega_2 \in \mathcal{F}$ of frames, there is at most one refinement $\tau : \Omega_1 \rightarrow \mathcal{P}(\Omega_2)$ in \mathcal{R} .*
2. *There exists a set U and a mapping $\varphi : \mathcal{F} \rightarrow \text{Part}(U)$ such that for all frames $\Omega_1, \Omega_2 \in \mathcal{F}$ we have $\varphi(\Omega_1) \neq \varphi(\Omega_2)$ whenever $\Omega_1 \neq \Omega_2$, and such that $\varphi(\mathcal{F}) = \{\varphi(\Omega) : \Omega \in \mathcal{F}\}$ forms a bounded sublattice of $\text{Part}(U)$.*
3. *For each frame $\Omega \in \mathcal{F}$ there is a bijective mapping $b : \Omega \rightarrow \varphi(\Omega)$. This mapping identifies for each frame element the corresponding block in the partition that contains this element.*
4. *There is a refinement $\tau : \Omega_1 \rightarrow \mathcal{P}(\Omega_2)$ in \mathcal{R} exactly if $\varphi(\Omega_2) \leq \varphi(\Omega_1)$.*

Condition 2 means that every frame corresponds to a partition and that the corresponding mapping is an embedding of \mathcal{F} into $\text{Part}(U)$. Since the least partition of U is in $\varphi(\mathcal{F})$, there is exactly one frame Ω whose partition $\varphi(\Omega)$ corresponds to the least partition. By the mapping of Condition 3, the elements of the frame Ω are in one-to-one correspondence with the blocks of the least partition that are the singletons of the universe U . From a mathematical point of view, the two sets U and Ω are thus isomorphic. Finally, as a consequence of Condition 4, we can introduce a decomposition mapping for each refinement. Indeed, if $\tau : \Omega_1 \rightarrow \mathcal{P}(\Omega_2)$, we have $\varphi(\Omega_2) \leq \varphi(\Omega_1)$ which means that every block in $\varphi(\Omega_2)$ is contained in some block of $\varphi(\Omega_1)$. We may therefore assume a decomposition mapping

$$\delta : \varphi(\Omega_1) \rightarrow \mathcal{P}(\varphi(\Omega_2)) \quad (\text{A.7})$$

that assigns to each block $B \in \varphi(\Omega_1)$ the set of blocks in $\varphi(\Omega_2)$ whose union is B . This mapping can further be extended to sets of blocks $\delta : \mathcal{P}(\pi_i) \rightarrow \mathcal{P}(\pi_{i+1})$ by

$$\delta(A) = \bigcup \{\delta(B) : B \in A\}. \quad (\text{A.8})$$

for $A \subseteq \pi_i$. Such families of compatible frames provide sufficient structure for the definition of set potentials over partition lattices (Kohlas & Monney, 1995).

■ A.7 Set Potentials on Partition Lattices

Let $\langle \mathcal{F}, \mathcal{R} \rangle$ be a family of compatible frames with universe U and partition lattice $Part(U)$. A set potential $m : \mathcal{P}(\pi) \rightarrow \mathbb{R}_{\geq 0}$ with domain $d(m) = \pi \in \varphi(\mathcal{F}) \subseteq Part(U)$ is defined by a mapping that assigns non-negative real numbers to all subsets of the partition π . Let us again consider the inverse natural order between partitions. According to the labeling axiom, the domain of the combination of m_1 and m_2 with domains π_1 and π_2 must be $\pi_1 \vee \pi_2$, which corresponds to the coarsest partition that is finer than π_1 and π_2 . We then conclude from Condition 2 that frames Ω_1, Ω_2 and Ω exist such that $\pi_1 = \varphi(\Omega_1)$, $\pi_2 = \varphi(\Omega_2)$ and $\pi_1 \vee \pi_2 = \varphi(\Omega)$. Since $\pi_1, \pi_2 \leq \pi_1 \vee \pi_2$ in the natural order, it follows from Condition 4 that the two refinement mappings $\tau_1 : \Omega \rightarrow \mathcal{P}(\Omega_1)$ and $\tau_2 : \Omega \rightarrow \mathcal{P}(\Omega_2)$ exist in \mathcal{R} . We thus obtain the corresponding decomposition functions $\delta_1 : \mathcal{P}(\pi_1 \vee \pi_2) \rightarrow \mathcal{P}(\pi_1)$ and $\delta_2 : \mathcal{P}(\pi_1 \vee \pi_2) \rightarrow \mathcal{P}(\pi_2)$. Altogether, this allows us to define the combination rule: For $A \subseteq \pi_1 \vee \pi_2$ we define

$$m_1 \otimes m_2(A) = \sum_{A=B \cap C} m_1(\delta_1(B)) \cdot m_2(\delta_2(C)) \quad (\text{A.9})$$

For the projection operator, we assume a set potential m with domain π_1 and $\pi_2 \leq \pi_1$. By the same justification we find the decomposition mapping $\delta : \mathcal{P}(\pi_1) \rightarrow \mathcal{P}(\pi_2)$ and define for $A \subseteq \pi_2$ the projection rule as follows:

$$m^{\downarrow \pi_2}(A) = \sum_{B \subseteq \pi_1 : A \cap \delta(B) \neq \emptyset} m(B). \quad (\text{A.10})$$

The proof that set potentials over partition lattices satisfy the axioms of a valuation algebra on general lattices can be found in (Kohlas & Monney, 1995).

A.3 VALUATION ALGEBRAS WITH PARTIAL PROJECTION

The valuation algebra definition given at the beginning of this chapter allows every valuation $\phi \in \Phi$ to be projected to any subset of $d(\phi)$. Hence, we may say that ϕ can be projected to all domains in the *marginal set* $\mathcal{M}(\phi) = \mathcal{P}(d(\phi))$. Valuation algebras with partial projection are more general in the sense that not all projections are necessarily defined. In this view, $\mathcal{M}(\phi)$ may be a strict subset of $\mathcal{P}(d(\phi))$. It is therefore sensible that a fourth operation is needed which produces $\mathcal{M}(d(\phi))$ for all $\phi \in \Phi$. Additionally, all axioms that bear somehow on projection must be generalized to take the corresponding marginal sets into account. Thus, let Φ be a set of valuations over domains $s \subseteq r$ and $D = \mathcal{P}(r)$. We assume the following operations in $\langle \Phi, D \rangle$:

1. *Labeling*: $\Phi \rightarrow D; \phi \mapsto d(\phi)$,
2. *Combination*: $\Phi \times \Phi \rightarrow \Phi; (\phi, \psi) \mapsto \phi \otimes \psi$,
3. *Domain*: $\Phi \rightarrow \mathcal{P}(D); \phi \mapsto \mathcal{M}(\phi)$,

4. *Partial Projection*: $\Phi \times D \rightarrow \Phi$; $(\phi, x) \mapsto \phi^{\perp x}$ defined for $x \in \mathcal{M}(\phi)$.

The set $\mathcal{M}(\phi)$ contains therefore all domains $x \in D$ such that the marginal of ϕ relative to x is defined. We impose now the following set of axioms on Φ and D , pointing out that the two Axioms (A1'') and (A2'') remain identical to the traditional definition of a valuation algebra.

(A1'') *Commutative Semigroup*: Φ is associative and commutative under \otimes .

(A2'') *Labeling*: For $\phi, \psi \in \Phi$,

$$d(\phi \otimes \psi) = d(\phi) \cup d(\psi). \quad (\text{A.11})$$

(A3'') *Projection*: For $\phi \in \Phi$ and $x \in \mathcal{M}(\phi)$,

$$d(\phi^{\perp x}) = x. \quad (\text{A.12})$$

(A4'') *Transitivity*: If $\phi \in \Phi$ and $x \subseteq y \subseteq d(\phi)$, then

$$x \in \mathcal{M}(\phi) \Rightarrow x \in \mathcal{M}(\phi^{\perp y}), y \in \mathcal{M}(\phi) \text{ and } (\phi^{\perp y})^{\perp x} = \phi^{\perp x}. \quad (\text{A.13})$$

(A5'') *Combination*: If $\phi, \psi \in \Phi$ with $d(\phi) = x$, $d(\psi) = y$ and $z \in D$ such that $x \subseteq z \subseteq x \cup y$, then

$$z \cap y \in \mathcal{M}(\psi) \Rightarrow z \in \mathcal{M}(\phi \otimes \psi) \text{ and } (\phi \otimes \psi)^{\perp z} = \phi \otimes \psi^{\perp z \cap y}. \quad (\text{A.14})$$

(A6'') *Domain*: For $\phi \in \Phi$ with $d(\phi) = x$, we have $x \in \mathcal{M}(\phi)$ and

$$\phi^{\perp x} = \phi. \quad (\text{A.15})$$

Definition A.9 A system (Φ, D) together with the operations of labeling, combination, partial projection and domain satisfying these axioms is called a valuation algebra with partial projection.

It is easy to see that this system is indeed a generalization of the traditional valuation algebra, because if $\mathcal{M}(\phi) = \mathcal{P}(d(\phi))$ holds for all $\phi \in \Phi$, the axioms reduce to the system given at the beginning of this chapter. Therefore, the latter is also called valuation algebra with full projection.

■ A.8 Quotients of Density Functions

Let us reconsider the valuation algebra of density functions introduced in Instance 1.5 and assume that f is a positive density function over variables in $s \subseteq r$, i.e. for $\mathbf{x} \in \mathbb{R}^s$ we always have $f(\mathbf{x}) > 0$. We consider the quotient

$$g(\mathbf{x}) = \frac{f(\mathbf{x})}{f^{\downarrow t}(\mathbf{x}^{\downarrow t})}$$

with $t \subseteq s$. For any $t \subseteq s$ these quotients represent the family of *conditional distributions* of $\mathbf{x}^{\downarrow s-t}$ given $\mathbf{x}^{\downarrow t}$ associated with the density f . However, projection is only partially defined for such quotients since

$$g^{\downarrow t}(\mathbf{x}^{\downarrow t}) = \left(\frac{f(\mathbf{x})}{f^{\downarrow t}(\mathbf{x}^{\downarrow t})} \right)^{\downarrow t} = \frac{f^{\downarrow t}(\mathbf{x}^{\downarrow t})}{f^{\downarrow t}(\mathbf{x}^{\downarrow t})} = 1.$$

This is a constant function with an infinite integral and therefore not a density anymore. Nevertheless, conditional density functions are part of a valuation algebra with partial projection. A formal verification of the corresponding axiomatic system can be found in (Kohlas, 2003).

PROBLEM SETS AND EXERCISES

A.1 ★ Verify the valuation algebra axioms for the relational algebra of Instance 1.2 without restriction to variables with finite frames.

A.2 ★ Reconsider the valuation algebra of arithmetic potentials from Instance 1.3. This time, however, we restrict ourselves to the unit interval and replace the operation of addition in the definition of projection in equation (1.11) by maximization. This leads to the formalism of *possibility potentials* or *probabilistic constraints* that will later be discussed in Instance 5.2. Prove that the valuation algebra axioms still hold in this new formalism.

A.3 ★ Again, reconsider the valuation algebra of arithmetic potentials from Instance 1.3. This time, we replace the operation of multiplication in the definition of combination in equation (1.10) by addition and the operation of addition in the definition of projection in equation (1.11) by minimization. This leads to the formalism of *Spohn potentials* or *weighted constraints* which will later be discussed in Instance 5.1. Prove that the valuation algebra axioms still hold in this new formalism. Alternatively, we could also take maximization for the projection rule.

A.4 ★ Cancellativity is an important property in semigroup theory and will be used frequently in later parts of this book. A valuation algebra $\langle \Phi, D \rangle$ is called *cancellative*,

if its semigroup under combination is cancellative, i.e. if for all $\phi \in \Phi$

$$\phi \otimes \psi = \phi \otimes \psi'$$

implies that $\psi = \psi'$. Prove that cancellativity holds in the valuation algebras of arithmetic potentials from Instance 1.3, Gaussian potentials from Instance 1.6 and Spohn potentials from Exercise A.3.

A.5 ★ The domain axiom (A6) expresses that valuations are not affected by trivial projection. Prove that this axiom is not a consequence of the remaining axioms (A1) to (A5), i.e. construct a formalism that satisfies the axioms (A1) to (A5) but not the domain axiom (A6). The basic idea is given in (Shafer, 1991): take any valuation algebra and double the number of elements by distinguishing two versions of each element, one marked and one unmarked. We further define that the result of a projection is always marked and that a combination produces a marked element if, and only if, one of its factors is marked. Prove that all valuation algebra axioms except the domain axiom (A6) still hold in this algebra.

A.6 ★★ Study the approximation of probability density functions by discrete probability distributions and provide suitable definitions of combination and projection.

A.7 ★★★ It was shown in Instance 1.6 that Gaussian potentials form a subalgebra of the valuation algebra of density functions. Look for other parametric classes of densities that establish subalgebras, for example *uniform distributions*, or more general classes like densities with *finite* or *infinite support*. The support of a density corresponds to the part of its range where it adopts a non-zero value.

