DEPARTMENT OF COMPUTER SCIENCE

FACULTY OF SCIENCE

PALACKÝ UNIVERSITY, OLOMOUC

INTRODUCTION TO FORMAL CONCEPT ANALYSIS

RADIM BĚLOHLÁVEK

VÝVOJ TOHOTO UČEBNÍHO TEXTU JE SPOLUFINANCOVÁN
EVROPSKÝM SOCIÁLNÍM FONDEM A STÁTNÍM ROZPOČTEM ČESKÉ REPUBLIKY

Olomouc 2008

Preface

This text develops fundamental concepts and methods of formal concept analysis. The text is meant as an introduction to formal concept analysis. The presentation is rigorous—we include definitions and theorems with proofs. On the other hand, we pay attention to the motivation and explanation of the presented material in informal terms and by means of numerous illustrative examples of the concepts, methods, and their practical meaning. Our goal in writing this text was to make the text accessible not only to mathematically educated people such as mathematicians, computer scientists, and engineers, but also to all potential users of formal concept analysis.

The text can be used for a graduate course on formal concept analysis. In addition, the text can be used as an introductory text to the topic of formal concept analysis for researchers and practitioners.

Contents

1 Introduction . . . 4
  1.1 What is Formal Concept Analysis? . . . 4
  1.2 First Example . . . 4
  1.3 Historical Roots and Development . . . 5
2 Concept Lattices . . . 6
  2.1 Input data . . . 6
  2.2 Concept-Forming Operators . . . 6
  2.3 Formal Concepts and Concept Lattice . . . 7
  2.4 Formal Concepts as Maximal Rectangles . . . 9
  2.5 Basic Mathematical Structures Behind FCA: Galois Connections and Closure Operators . . . 10
  2.6 Main Theorem of Concept Lattices . . . 15
  2.7 Clarification and Reduction of Formal Concepts . . . 17
  2.8 Basic Algorithm For Computing Concept Lattices . . . 23
3 Attribute Implications . . . 27
  3.1 Basic Notions Regarding Attribute Implications . . . 27
  3.2 Armstrong Rules and Reasoning With Attribute Implications . . . 30
  3.3 Models of Attribute Implications . . . 37
  3.4 Non-Redundant Bases of Attribute Implications . . . 40

1 Introduction

1.1 What is Formal Concept Analysis?

Formal concept analysis (FCA) is a method of data analysis with growing popularity across various domains. FCA analyzes data which describe relationship between a particular set of objects and a particular set of attributes. Such data commonly appear in many areas of human activities. FCA produces two kinds of output from the input data. The first is a concept lattice. A concept lattice is a collection of formal concepts in the data which are hierarchically ordered by a subconcept-superconcept relation. Formal concepts are particular clusters which represent natural human-like concepts such as "organism living in water", "car with all wheel drive system", "number divisible by 3 and 4", etc. The second output of FCA is a collection of so-called attribute implications. An attribute implication describes a particular dependency which is valid in the data such as "every number divisible by 3 and 4 is divisible by 6", "every respondent with age over 60 is retired", etc.

A distinguishing feature of FCA is an inherent integration of three components of conceptual processing of data and knowledge, namely, the discovery and reasoning with concepts in data, discovery and reasoning with dependencies in data, and visualization of data, concepts, and dependencies with folding/unfolding capabilities. Integration of these components makes FCA a powerful tool which has been applied to various problems. Examples include hierarchical organization of web search results into concepts based on common topics, gene expression data analysis, information retrieval, analysis and understanding of software code, debugging, data mining, and design in software engineering, internet applications including analysis and organization of documents and e-mail collections, annotated taxonomies, and further various data analysis projects described in the literature. Interesting applications in counterterrorism, in particular in analysis and visualization of data related to terrorist activities, have been reported in a recent article "The N.S.A.'s Math Problem" in the 2006/05/16 edition of The New York Times.

1.2 First Example

A table with logical attributes can be represented by a triplet $\langle X, Y, I\rangle$ where I is a binary relation between X and Y. Elements of X are called objects and correspond to table rows, elements of Y are called attributes and correspond to table columns, and for $x \in X$ and $y \in Y$, $\langle x, y\rangle \in I$ indicates that object x has attribute y while $\langle x, y\rangle \notin I$ indicates that x does not have y. For instance, Fig. 1.2 (left) depicts a table with logical attributes. The corresponding triplet $\langle X, Y, I\rangle$ is given by $X = \{x_1, x_2, x_3, x_4\}$, $Y =$

	y_1	y_2	y_3	$\cdots$
x_1	$\times$	$\times$	$\times$	
x_2	$\times$	$\times$		$\vdots$
x_3		$\times$	$\times$	
$\vdots$		$\cdots$		$\ddots$

	y_1	y_2	y_3	$\cdots$
x_1	1	1	0.7	
x_2	0.8	0.6	0.1	$\vdots$
x_3	0	0.9	0.9	
$\vdots$		$\cdots$		$\ddots$

Figure 1: Tables with logical attributes: crisp attributes (left), fuzzy attributes (right).

$\{y_1, y_2, y_3\}$, and we have $\langle x_1, y_1\rangle \in I$, $\langle x_2, y_3\rangle \notin I$, etc. Since representing tables with logical attributes by triplets is common in FCA, we say just "table $\langle X, Y, I\rangle$" instead of "triplet $\langle X, Y, I\rangle$ representing a given table". FCA aims at obtaining two outputs out of a given table. The first one, called a concept lattice, is a partially ordered collection of particular clusters of objects and attributes. The second one consists of formulas, called attribute implications, describing particular attribute dependencies which are true in the table. The clusters, called formal concepts, are pairs $\langle A, B\rangle$ where $A \subseteq X$ is a set of objects and $B \subseteq Y$ is a set of attributes such that A is a set of all objects which have all attributes from B, and B is the set of all attributes which are common to all objects from A. For instance, $\langle\{x_1, x_2\}, \{y_1, y_2\}\rangle$ and $\langle\{x_1, x_2, x_3\}, \{y_2\}\rangle$ are examples of formal concepts of the (visible part of) left table in Fig. 1.2. An attribute implication is an expression $A \Rightarrow B$ with A and B being sets of attributes. $A \Rightarrow B$ is true in table $\langle X, Y, I\rangle$ if each object having all attributes from A has all attributes from B as well. For instance, $\{y_3\} \Rightarrow \{y_2\}$ is true in the (visible part of) left table in Fig. 1.2, while $\{y_1, y_2\} \Rightarrow \{y_3\}$ is not (x_2 serves as a counterexample).

1.3 Historical Roots and Development

Although some previous attempts exist, it is generally agreed and FCA started by Wille's seminal paper [10]. Cautious development of mathematical foundations which later proved useful when developing computational foundations is one strong feature of FCA. Another is its reliance on a simple and robust notion of a concept inspired by a traditional approach to concepts as developed in traditional logic. Introduction and applications of FCA are described in [2], mathematical foundations are covered in [5], the state of the art is surveyed in [6].

There are three international conferences devoted to FCA, namely, ICFCA (International Conference on Formal Concept Analysis), CLA (Concept Lattices and Their Applications), and ICCS (International Conference on Conceptual Structures). In addition, further papers on FCA can be found in journals and proceedings of other conferences.

2 Concept Lattices

Goals: This chapter introduces basic notions of formal concept analysis, among which are the fundamental notions of a formal context, formal concept, and concept lattice. The chapter introduces these notions and related mathematical structures such as Galois connections and closure operators and their basic properties as well as basic properties of concept lattices. The chapter also presents NextClosure—a basic algorithm for computing a concept lattice.

Keywords: formal context, formal concept, concept lattice, concept-deriving operator, Galois connection.

2.1 Input data

In the basic setting, the input data to FCA is in the form of a table (called a cross-table) which describes a relationship between objects (represented by table rows) and attributes (represented by table columns). An example of such table is shown in Fig. 2. A table entry containing $\times$indicates that the corresponding object has the correspond-

I	y_1	y_2	y_3	y_4
x_1	$\times$	$\times$	$\times$	$\times$
x_2	$\times$		$\times$	$\times$
x_3		$\times$	$\times$	$\times$
x_4		$\times$	$\times$	$\times$
x_5	$\times$			

Figure 2: Cross-table.

ing attribute. For example, if objects are products such as cars and attributes are car attributes such as "has ABS", $\times$indicates that a particular car has ABS (anti-block braking system). A table entry containing a blank symbol (empty entry) indicates that the object does not have the attribute (a particular car does not have ABS). Thus, in Fig. 2, object x_2 has attribute y_1 but does not have attribute y_2.

Formally, a (cross-)table is represented by a so-called formal context.

Definition 2.1 (formal context)**.** A *formal context* is a triplet $\langle X, Y, I\rangle$ where X and Y are non-empty sets and I is a binary relation between X and Y, i.e., $I \subseteq X \times Y$.

For a formal context, elements x from X are called objects and elements y from Y are called attributes. $\langle x, y\rangle \in I$ indicates that object x has attribute y. For a given a cross-table with n rows and m columns, a corresponding formal context $\langle X, Y, I\rangle$ consists of a set $X = \{x_1, \ldots, x_n\}$, a set $Y = \{y_1, \ldots, y_m\}$, and a relation I defined by: $\langle x_i, y_j\rangle \in I$ if and only if the table entry corresponding to row i and column j contains $\times$.

2.2 Concept-Forming Operators

Every formal context induces a pair of operators, so-called concept-forming operators.

Definition 2.2 (concept-forming operators)**.** For a formal context $\langle X, Y, I\rangle$, operators $^{\uparrow} : 2^X \to 2^Y$ and $^{\downarrow} : 2^Y \to 2^X$ are defined for every $A \subseteq X$ and $B \subseteq Y$ by

$$\begin{aligned} A^{\uparrow} &= \{y \in Y \mid \text{for each } x \in A : \langle x, y\rangle \in I\}, \\ B^{\downarrow} &= \{x \in X \mid \text{for each } y \in B : \langle x, y\rangle \in I\}. \end{aligned}$$

Remark 2.3.
- Operator $^\uparrow$ assigns subsets of Y to subsets of X. $A^\uparrow$ is just the set of all attributes shared by all objects from A.
- Dually, operator $^\downarrow$ assigns subsets of X to subsets of Y. $B^\uparrow$ is the set of all objects sharing all attributes from B.
- To emphasize that $^\uparrow$ and $^\downarrow$ are induced by $\langle X, Y, I\rangle$, we use $^{\uparrow_I}$ and $^{\downarrow_I}$.

Example 2.4 (concept-forming operators)**.** For table

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

we have:

- $\{x_2\}^\uparrow = \{y_1, y_3, y_4\}$, $\{x_2, x_3\}^\uparrow = \{y_3, y_4\}$,
- $\{x_1, x_4, x_5\}^\uparrow = \emptyset$,
- $X^\uparrow = \emptyset$, $\emptyset^\uparrow = Y$,
- $\{y_1\}^\downarrow = \{x_1, x_2, x_5\}$, $\{y_1, y_2\}^\downarrow = \{x_1\}$,
- $\{y_2, y_3\}^\downarrow = \{x_1, x_3, x_4\}$, $\{y_2, y_3, y_4\}^\downarrow = \{x_1, x_3, x_4\}$,
- $\emptyset^\downarrow = X$, $Y^\downarrow = \{x_1\}$.

2.3 Formal Concepts and Concept Lattice

The notion of a formal concept is fundamental in FCA. Formal concepts are particular clusters in cross-tables, defined by means of attribute sharing.

Definition 2.5 (formal concept)**.** A *formal concept* in $\langle X, Y, I\rangle$ is a pair $\langle A, B\rangle$ of $A \subseteq X$ and $B \subseteq Y$ such that $A^\uparrow = B$ and $B^\downarrow = A$.

For a formal concept $\langle A, B\rangle$ in $\langle X, Y, I\rangle$, A and B are called the extent and intent of $\langle A, B\rangle$, respectively. Note the following verbal description of formal concepts: $\langle A, B\rangle$ is a formal concept if and only if A contains just objects sharing all attributes from B and B contains just attributes shared by all objects from A. Mathematically, $\langle A, B\rangle$ is a formal concept if and only if $\langle A, B\rangle$ is a fixpoint of the pair $\langle ^\uparrow, ^\downarrow\rangle$ of concept-forming operators.

Example 2.6 (formal concept)**.** For table

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

the highlighted rectangle represents formal concept

$$\langle A_1, B_1\rangle = \langle \{x_1, x_2, x_3, x_4\}, \{y_3, y_4\}\rangle$$

because

$$\{x_1, x_2, x_3, x_4\}^\uparrow = \{y_3, y_4\} \text{ and } \{y_3, y_4\}^\downarrow = \{x_1, x_2, x_3, x_4\}.$$

But there are further formal concepts. Three of them are represented by the following highlighted rectangles:

I	y_1	y_2	y_3	y_4
x_1	×	**×**	**×**	**×**
x_2	×		×	×
x_3		**×**	**×**	**×**
x_4		**×**	**×**	**×**
x_5	×			

I	y_1	y_2	y_3	y_4
x_1	**×**	×	**×**	**×**
x_2	**×**		**×**	**×**
x_3		×	×	×
x_4		×	×	×
x_5	×			

I	y_1	y_2	y_3	y_4
x_1	**×**	×	×	×
x_2	**×**		×	×
x_3		×	×	×
x_4		×	×	×
x_5	**×**			

Namely, $\langle A_2, B_2\rangle = \langle\{x_1, x_3, x_4\}, \{y_2, y_3, y_4\}\rangle$, $\langle A_3, B_3\rangle = \langle\{x_1, x_2\}, \{y_1, y_3, y_4\}\rangle$, and $\langle A_4, B_4\rangle = \langle\{x_1, x_2, x_5\}, \{y_1\}\rangle$.

The notion of a formal concept can be seen as a simple mathematization of a well-known notion of a concept, which goes back to Port-Royal logic. According to Port-Royal, a concept is determined by a collection of objects (extent) which fall under the concept and a collection of attributes (intent) covered by the concepts. Concepts are naturally ordered using a subconcept-superconcept relation. The subconcept-superconcept relation is based on inclusion relation on objects and attributes. Formally, the subconcept-superconcept relation is defined as follows.

Definition 2.7 (subconcept-superconcept ordering)**.** For formal concepts $\langle A_1, B_1\rangle$ and $\langle A_2, B_2\rangle$ of $\langle X, Y, I\rangle$, put $\langle A_1, B_1\rangle \leq \langle A_2, B_2\rangle$ iff $A_1 \subseteq A_2$ (iff $B_2 \subseteq B_1$).

Remark 2.8.
- $\leq$ represents the subconcept-superconcept ordering.
- $\langle A_1, B_1\rangle \leq \langle A_2, B_2\rangle$ means that $\langle A_1, B_1\rangle$ is more specific than $\langle A_2, B_2\rangle$ ($\langle A_2, B_2\rangle$ is more general).
- $\leq$ captures the intuition behind DOG $\leq$ MAMMAL (the concept of a dog is more specific than the concept of a mammal).

Example 2.9. Consider the following formal concepts from Example 2.6: $\langle A_1, B_1\rangle = \langle\{x_1, x_2, x_3, x_4\}, \{y_3, y_4\}\rangle$, $\langle A_2, B_2\rangle = \langle\{x_1, x_3, x_4\}, \{y_2, y_3, y_4\}\rangle$, $\langle A_3, B_3\rangle = \langle\{x_1, x_2\}, \{y_1, y_3, y_4\}\rangle$, $\langle A_4, B_4\rangle = \langle\{x_1, x_2, x_5\}, \{y_1\}\rangle$. Then:
$\langle A_3, B_3\rangle \leq \langle A_1, B_1\rangle$, $\langle A_3, B_3\rangle \leq \langle A_2, B_2\rangle$, $\langle A_3, B_3\rangle \leq \langle A_4, B_4\rangle$, $\langle A_2, B_2\rangle \leq \langle A_1, B_1\rangle$, $\langle A_1, B_1\rangle || \langle A_4, B_4\rangle$ (incomparable), $\langle A_2, B_2\rangle || \langle A_4, B_4\rangle$.

The collection of all formal concepts of a given formal contxt is called a concept lattice, another fundamental notion in FCA.

Definition 2.10 (concept lattice)**.** Denote by $\mathcal{B}(X, Y, I)$ the collection of all formal concepts of $\langle X, Y, I\rangle$, i.e.

$$\mathcal{B}(X, Y, I) = \{\langle A, B\rangle \in 2^X \times 2^Y \mid A^{\uparrow} = B, B^{\downarrow} = A\}.$$

$\mathcal{B}(X, Y, I)$ equipped with the subconcept-superconcept ordering $\leq$ is called a *concept lattice* of $\langle X, Y, I\rangle$.

- $\mathcal{B}(X, Y, I)$ represents all (potentially interesting) clusters which are "hidden" in data $\langle X, Y, I\rangle$.
- We will see that $\langle \mathcal{B}(X, Y, I), \leq\rangle$ is indeed a lattice later.

Denote

$$\mathrm{Ext}(X, Y, I) = \{A \in 2^X \mid \langle A, B\rangle \in \mathcal{B}(X, Y, I) \text{ for some } B\} \text{ (extents of concepts)}$$

and

$$\mathrm{Int}(X, Y, I) = \{B \in 2^Y \mid \langle A, B\rangle \in \mathcal{B}(X, Y, I) \text{ for some } A\} \text{ (intents of concepts).}$$

Example 2.11. Consider the following cross-table (input data, taken from [5]):

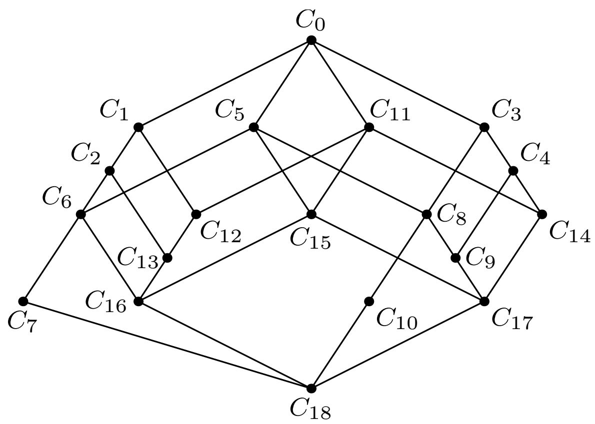
		a	b	c	d	e	f	g	h	i
leech	1	×	×					×		
bream	2	×	×					×	×	
frog	3	×	×	×				×	×	
dog	4	×		×				×	×	×
spike-weed	5	×	×		×		×			
reed	6	×	×	×	×		×			
bean	7	×		×	×	×				
maize	8	×		×	×		×			

a: needs water to live, b: lives in water,
c: lives on land, d: needs chlorophyll to produce food,
e: two seed leaves, f: one seed leaf,
g: can move around, h: has limbs,
i: suckles its offspring.

The corresponding formal context $\langle X, Y, I\rangle$ contains the following formal concepts:

$C_0 = \langle\{1,2,3,4,5,6,7,8\},\{a\}\rangle$, $C_1 = \langle\{1,2,3,4\},\{a,g\}\rangle$, $C_2 = \langle\{2,3,4\},\{a,g,h\}\rangle$,
$C_3 = \langle\{5,6,7,8\},\{a,d\}\rangle$, $C_4 = \langle\{5,6,8\},\{a,d,f\}\rangle$, $C_5 = \langle\{3,4,6,7,8\},\{a,c\}\rangle$,
$C_6 = \langle\{3,4\},\{a,c,g,h\}\rangle$, $C_7 = \langle\{4\},\{a,c,g,h,i\}\rangle$, $C_8 = \langle\{6,7,8\},\{a,c,d\}\rangle$,
$C_9 = \langle\{6,8\},\{a,c,d,f\}\rangle$, $C_{10} = \langle\{7\},\{a,c,d,e\}\rangle$, $C_{11} = \langle\{1,2,3,5,6\},\{a,b\}\rangle$,
$C_{12} = \langle\{1,2,3\},\{a,b,g\}\rangle$, $C_{13} = \langle\{2,3\},\{a,b,g,h\}\rangle$, $C_{14} = \langle\{5,6\},\{a,b,d,f\}\rangle$,
$C_{15} = \langle\{3,6\},\{a,b,c\}\rangle$, $C_{16} = \langle\{3\},\{a,b,c,g,h\}\rangle$, $C_{17} = \langle\{6\},\{a,b,c,d,f\}\rangle$,
$C_{18} = \langle\{\},\{a,b,c,d,e,f,g,h,i\}\rangle$.

The corresponding concept lattice $\mathcal{B}(X, Y, I)$ is depicted in the following figure:



2.4 Formal Concepts as Maximal Rectangles

Alternatively, formal concepts can conveniently be defined as maximal rectangles in the cross-table.

Definition 2.12 (rectangles in $\langle X, Y, I\rangle$)**.** A rectangle in $\langle X, Y, I\rangle$ is a pair $\langle A, B\rangle$ such that $A \times B \subseteq I$, i.e.: for each $x \in A$ and $y \in B$ we have $\langle x, y\rangle \in I$. For rectangles $\langle A_1, B_1\rangle$ and $\langle A_2, B_2\rangle$, put $\langle A_1, B_1\rangle \sqsubseteq \langle A_2, B_2\rangle$ iff $A_1 \subseteq A_2$ and $B_1 \subseteq B_2$.

Example 2.13. Consider

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

In this table, $\langle\{x_1, x_2, x_3\}, \{y_3, y_4\}\rangle$ is a rectangle which is not maximal w.r.t. $\sqsubseteq$. $\langle\{x_1, x_2, x_3, x_4\}, \{y_3, y_4\}\rangle$ is a rectangle which is maximal w.r.t. $\sqsubseteq$.

Theorem 2.14 (formal concepts as maximal rectangles). *$\langle A, B\rangle$ is a formal concept of $\langle X, Y, I\rangle$ iff $\langle A, B\rangle$ is a maximal rectangle in $\langle X, Y, I\rangle$.*

Proof. Left as an exercise (by direct verification). □

We will see that a "geometrical reasoning" in FCA based on the idea of formal concepts as rectangles is important.

2.5 Basic Mathematical Structures Behind FCA: Galois Connections and Closure Operators

In this section, we present the basic mathematical structures behind FCA and their properties. We start with the concept of Galois connections. Namely, as we will see, the concept-forming operators form a representative case of Galois connections.

Definition 2.15 (Galois connection). A *Galois connection* between sets X and Y is a pair $\langle f, g\rangle$ of $f : 2^X \to 2^Y$ and $g : 2^Y \to 2^X$ satisfying for $A, A_1, A_2 \subseteq X$, $B, B_1, B_2 \subseteq Y$:

$$A_1 \subseteq A_2 \Rightarrow f(A_2) \subseteq f(A_1), \tag{2.1}$$

$$B_1 \subseteq B_2 \Rightarrow g(B_2) \subseteq g(B_1), \tag{2.2}$$

$$A \subseteq g(f(A)), \tag{2.3}$$

$$B \subseteq f(g(B). \tag{2.4}$$

Definition 2.16 (fixpoints of Galois connections). For a Galois connection $\langle f, g\rangle$ between sets X and Y, the set

$$\mathrm{fix}(\langle f, g\rangle) = \{\langle A, B\rangle \in 2^X \times 2^Y \mid f(A) = B, g(B) = A\}$$

is called a set of *fixpoints* of $\langle f, g\rangle$.

The following theorem shows a fundamental property of concept-forming operators.

Theorem 2.17 (concept-forming operators form a Galois connection). *For a formal context $\langle X, Y, I\rangle$, the pair $\langle {}^{\uparrow_I}, {}^{\downarrow_I}\rangle$ of operators induced by $\langle X, Y, I\rangle$ is a Galois connection between X and Y.*

Proof. Left as an exercise (by direct verification). □

We have the following direct consequence.

Lemma 2.18 (chaining of Galois connection). *For a Galois connection $\langle f, g\rangle$ between X and Y we have $f(A) = f(g(f(A)))$ and $g(B) = g(f(g(B)))$ for any $A \subseteq X$ and $B \subseteq Y$.*

Proof. We prove only $f(A) = f(g(f(A)))$, $g(B) = g(f(g(B)))$ is dual:
"$\subseteq$": $f(A) \subseteq f(g(f(A)))$ follows from (2.4) by putting $B = f(A)$.

"$\supseteq$": Since $A \subseteq g(f(A))$ by (2.3), we get $f(A) \supseteq f(g(f(A)))$ by application of (2.1). □

Another important notion related to FCA is that of a closure operator.

Definition 2.19 (closure operator)**.** A *closure operator* on a set X is a mapping $C : 2^X \to 2^X$ satisfying for each $A, A_1, A_2 \subseteq X$

$$A \subseteq C(A), \tag{2.5}$$
$$A_1 \subseteq A_2 \Rightarrow C(A_1) \subseteq C(A_2), \tag{2.6}$$
$$C(A) = C(C(A)). \tag{2.7}$$

Definition 2.20 (fixpoints of closure operators)**.** For a closure operator $C : 2^X \to 2^X$, the set

$$\text{fix}(C) = \{A \subseteq X \mid C(A) = A\}$$

is called a set of *fixpoints* of C.

Closure operators result from the concept-forming operators by their composition:

Theorem 2.21 (from Galois connection to closure operators)**.** *If $\langle f, g\rangle$ is a Galois connection between X and Y then $C_X = f \circ g$ is a closure operator on X and $C_Y = g \circ f$ is a closure operator on Y.*

Proof. We show that $f \circ g : 2^X \to 2^X$ is a closure operator on X:
(2.5) is $A \subseteq g(f(A))$ which is true by definition of a Galois connection.
(2.6): $A_1 \subseteq A_2$ impies $f(A_2) \subseteq f(A_1)$ which implies $g(f(A_1)) \subseteq g(f(A_2))$.
(2.7): Since $f(A) = f(g(f(A)))$, we get $g(f(A)) = g(f(g(f(A))))$. □

The next theorem shows that extents and intents are just the images under the concept-forming operators.

Theorem 2.22 (extents and intents)**.**

$$\begin{aligned} \text{Ext}(X, Y, I) &= \{B^{\downarrow} \mid B \subseteq Y\}, \\ \text{Int}(X, Y, I) &= \{A^{\uparrow} \mid A \subseteq X\}. \end{aligned}$$

Proof. We prove only the part for $\text{Ext}(X, Y, I)$, part for $\text{Int}(X, Y, I)$ is dual.

"$\subseteq$": If $A \in \text{Ext}(X, Y, I)$, then $\langle A, B\rangle$ is a formal concept for some $B \subseteq Y$. By definition, $A = B^{\downarrow}$, i.e. $A \in \{B^{\downarrow} \mid B \subseteq Y\}$.

"$\supseteq$": Let $A \in \{B^{\downarrow} \mid B \subseteq Y\}$, i.e. $A = B^{\downarrow}$ for some B. Then $\langle A, A^{\uparrow}\rangle$ is a formal concept. Namely, $A^{\uparrow\downarrow} = B^{\downarrow\uparrow\downarrow} = B^{\downarrow} = A$ by chaining, and $A^{\uparrow} = A^{\uparrow}$ for free. That is, A is the extent of a formal concept $\langle A, A^{\uparrow}\rangle$, whence $A \in \text{Ext}(X, Y, I)$. □

Closures of sets are the least extents and intents:

Theorem 2.23 (least extent containing A, least intent containing B)**.** *The least extent containing $A \subseteq X$ is $A^{\uparrow\downarrow}$. The least intent containing $B \subseteq Y$ is $B^{\downarrow\uparrow}$.*

Proof. For extents:
1. $A^{\uparrow\downarrow}$ is an extent (by previous theorem).

2. If C is an extent such that $A \subseteq C$, then $A^{\uparrow\downarrow} \subseteq C^{\uparrow\downarrow}$ because ${}^{\uparrow\downarrow}$ is a closure operator. Therefore, $A^{\uparrow\downarrow}$ is the least extent containing A. □

The next theorem provides a useful description of a system of extents, intents, and a concept lattice.

Theorem 2.24. *For any formal context $\langle X, Y, I\rangle$:*

$$\begin{aligned} \text{Ext}(X, Y, I) &= \text{fix}({}^{\uparrow\downarrow}), \\ \text{Int}(X, Y, I) &= \text{fix}({}^{\downarrow\uparrow}), \\ \mathcal{B}(X, Y, I) &= \{\langle A, A^{\uparrow}\rangle \mid A \in \text{Ext}(X, Y, I)\}, \\ \mathcal{B}(X, Y, I) &= \{\langle B^{\downarrow}, B\rangle \mid B \in \text{Int}(X, Y, I)\}. \end{aligned}$$

Proof. For $\mathrm{Ext}(X, Y, I)$:
We need to show that A is an extent iff $A = A^{\uparrow\downarrow}$.
"$\Rightarrow$": If A is an extent then for the corresponding formal concept $\langle A, B\rangle$ we have $B = A^{\uparrow}$ and $A = B^{\downarrow} = A^{\uparrow\downarrow}$. Hence, $A = A^{\uparrow\downarrow}$.
"$\Leftarrow$": If $A = A^{\uparrow\downarrow}$ then $\langle A, A^{\uparrow}\rangle$ is a formal concept. Namely, denoting $\langle A, B\rangle = \langle A, A^{\uparrow}\rangle$, we have both $A^{\uparrow} = B$ and $B^{\downarrow} = A^{\uparrow\downarrow} = A$. Therefore, A is an extent.

For $\mathcal{B}(X, Y, I) = \{\langle A, A^{\uparrow}\rangle \mid A \in \mathrm{Ext}(X, Y, I)\}$:
If $\langle A, B\rangle \in \mathcal{B}(X, Y, I)$ then $B = A^{\uparrow}$ and, obviously, $A \in \mathrm{Ext}(X, Y, I)$.

If $A \in \mathrm{Ext}(X, Y, I)$ then $A = A^{\uparrow\downarrow}$ (above claim) and, therefore, $\langle A, A^{\uparrow}\rangle \in \mathcal{B}(X, Y, I)$.

For $\mathcal{B}(X, Y, I) = \{\langle A, A^{\uparrow}\rangle \mid A \in \mathrm{Ext}(X, Y, I)\}$:
If $\langle A, B\rangle \in \mathcal{B}(X, Y, I)$ then $B = A^{\uparrow}$ and, obviously, $A \in \mathrm{Ext}(X, Y, I)$.

If $A \in \mathrm{Ext}(X, Y, I)$ then $A = A^{\uparrow\downarrow}$ (above claim) and, therefore, $\langle A, A^{\uparrow}\rangle \in \mathcal{B}(X, Y, I)$. □

Remark 2.25. The previous theorem says that in order to obtain $\mathcal{B}(X, Y, I)$, we can:

1. compute $\mathrm{Ext}(X, Y, I)$,
2. for each $A \in \mathrm{Ext}(X, Y, I)$, output $\langle A, A^{\uparrow}\rangle$.

There is a single condition which is equivalent to the four conditions from definition of a Galois connection:

Theorem 2.26. *$\langle f, g\rangle$ is a Galois connection between X and Y iff for every $A \subseteq X$ and $B \subseteq Y$:*

$$A \subseteq g(B) \quad \textit{iff} \quad B \subseteq f(A) \tag{2.8}$$

Proof. "$\Rightarrow$":
Let $\langle f, g\rangle$ be a Galois connection.
If $A \subseteq g(B)$ then $f(g(B)) \subseteq f(A)$ and since $B \subseteq f(g(B))$, we get $B \subseteq f(A)$. In similar way, $B \subseteq f(A)$ implies $A \subseteq g(B)$.

"$\Leftarrow$":
Let $A \subseteq g(B)$ iff $B \subseteq f(A)$. We check that $\langle f, g\rangle$ is a Galois connection.
Due to duality, it suffices to check (a) $A \subseteq g(f(A))$, and (b) $A_1 \subseteq A_2$ implies $f(A_2) \subseteq f(A_1)$.
(a): Due to our assumption, $A \subseteq g(f(A))$ is equivalent to $f(A) \subseteq f(A)$ which is evidently true.
(b): Let $A_1 \subseteq A_2$. Due to (a), we have $A_2 \subseteq g(f(A_2))$, therefore $A_1 \subseteq g(f(A_2))$. Using assumption, the latter is equivalent to $f(A_2) \subseteq f(A_1)$. □

Basic behavior of Galois connections with respect to union and intersection is described by the following theorem.

Theorem 2.27. *$\langle f, g\rangle$ is a Galois connection between X and Y then for $A_j \subseteq X$, $j \in J$, and $B_j \subseteq Y$, $j \in J$ we have*

$$f(\bigcup_{j\in J} A_j) = \bigcap_{j\in J} f(A_j), \tag{2.9}$$

$$g(\bigcup_{j\in J} B_j) = \bigcap_{j\in J} g(B_j). \tag{2.10}$$

Proof. (2.9):
For any $D \subseteq Y$: $D \subseteq f(\bigcup_{j\in J} A_j)$ iff $\bigcup_{j\in J} A_j \subseteq g(D)$ iff for each $j \in J$: $A_j \subseteq g(D)$ iff for each $j \in J$: $D \subseteq f(A_j)$ iff $D \subseteq \bigcap_{j\in J} f(A_j)$.
Since D is arbitrary, it follows that $f(\bigcup_{j\in J} A_j) = \bigcap_{j\in J} f(A_j)$.

(2.10): dual. □

Not only every pair of concept-forming operators forms a Galois, every Galois connection is a concept-forming operator of a particular formal context:

Theorem 2.28. *Let $\langle f, g\rangle$ be a Galois connection between X and Y. Consider a formal context $\langle X, Y, I\rangle$ such that I is defined by*

$$\langle x, y\rangle \in I \quad \text{iff} \quad y \in f(\{x\}) \qquad \text{or, equivalently, iff } x \in g(\{y\}), \tag{2.11}$$

for each $x \in X$ and $y \in Y$. Then $\langle {}^{\uparrow_I}, {}^{\downarrow_I}\rangle = \langle f, g\rangle$, i.e., the arrow operators $\langle {}^{\uparrow_I}, {}^{\downarrow_I}\rangle$ induced by $\langle X, Y, I\rangle$ coincide with $\langle f, g\rangle$.

Proof. First, we show $y \in f(\{x\})$ iff $x \in g(\{y\})$:
From $y \in f(\{x\})$ we get $\{y\} \subseteq f(\{x\})$ from which, using (2.8), we get $\{x\} \subseteq g(\{y\})$, i.e. $x \in g(\{y\})$.
In a similar way, $x \in g(\{y\})$ implies $y \in f(\{x\})$. This establishes $y \in f(\{x\})$ iff $x \in g(\{y\})$.

Now, for each $A \subseteq X$ we have $f(A) = f(\cup_{x\in A}\{x\}) = \cap_{x\in A} f(\{x\}) = \cap_{x\in A}\{y \in Y \mid y \in f(\{x\})\} = \cap_{x\in A}\{y \in Y \mid \langle x, y\rangle \in I\} = \{y \in Y \mid \text{ for each } x \in A : \langle x, y\rangle \in I\} = A^{\uparrow_I}$.
Dually, for $B \subseteq Y$ we get $g(B) = B^{\downarrow_I}$.

Now, using (2.9), for each $A \subseteq X$ we have

$$\begin{aligned} f(A) &= f(\cup_{x\in A}\{x\}) = \cap_{x\in A} f(\{x\}) = \\ &= \cap_{x\in A}\{y \in Y \mid y \in f(\{x\})\} = \cap_{x\in A}\{y \in Y \mid \langle x, y\rangle \in I\} = \\ &= \{y \in Y \mid \text{ for each } x \in A : \langle x, y\rangle \in I\} = A^{\uparrow_I}. \end{aligned}$$

Dually, for $B \subseteq Y$ we get $g(B) = B^{\downarrow_I}$. □

Remark 2.29.
- Relation I induced from $\langle f, g\rangle$ by (2.11) will be denoted by $I_{\langle f,g\rangle}$.
- Therefore, we have established two mappings:
  $I \mapsto \langle {}^{\uparrow_I}, {}^{\downarrow_I}\rangle$ assigns a Galois connection to a binary relation I.
  $\langle {}^{\uparrow}, {}^{\downarrow}\rangle \mapsto I_{\langle \uparrow, \downarrow\rangle}$ assigns a binary relation to a Galois connection.

Therefore, we get:

Theorem 2.30 (representation theorem). *$I \mapsto \langle {}^{\uparrow_I}, {}^{\downarrow_I}\rangle$ and $\langle {}^{\uparrow}, {}^{\downarrow}\rangle \mapsto I_{\langle \uparrow, \downarrow\rangle}$ are mutually inverse mappings between the set of all binary relations between X and Y and the set of all Galois connections between X and Y.*

Proof. Using the results established above, it remains to check that $I = I_{\langle \uparrow_I, \downarrow_I\rangle}$:
We have

$$\langle x, y\rangle \in I_{\langle \uparrow_I, \downarrow_I\rangle} \text{ iff } y \in \{x\}^{\uparrow_I} \text{ iff } \langle x, y\rangle \in I,$$

finishing the proof. □

Remark 2.31. In particular, previous theorem assures that (2.1)–(2.4) fully describe all the properties of our arrow operators induced by data $\langle X, Y, I\rangle$.

Having established properties of $\langle {}^{\uparrow}, {}^{\downarrow}\rangle$, we can see the duality relationship between extents and intents:

Theorem 2.32. *For* $\langle A_1, B_1\rangle, \langle A_2, B_2\rangle \in \mathcal{B}(X,Y,I)$,

$$A_1 \subseteq A_2 \quad \textit{iff} \quad B_2 \subseteq B_1. \tag{2.12}$$

Proof. By assumption, $A_i = B_i^{\downarrow}$ and $B_i = A_i^{\uparrow}$. Therefore, using (2.1) and (2.2), we get $A_1 \subseteq A_2$ implies $A_2^{\uparrow} \subseteq A_1^{\uparrow}$, i.e., $B_2 \subseteq B_1$, which implies $B_1^{\downarrow} \subseteq B_2^{\downarrow}$, i.e. $A_1 \subseteq A_2$. □

Therefore, the definition of a partial order $\leq$ on $\mathcal{B}(X,Y,I)$ is correct.

An immediate consequence of the above properties is the following theorem:

Theorem 2.33 (extents, intents, and formal concepts)**.** *1.* $\langle \mathrm{Ext}(X,Y,I), \subseteq\rangle$ *and* $\langle \mathrm{Int}(X,Y,I), \subseteq\rangle$ *are partially ordered sets.*

2. $\langle \mathrm{Ext}(X,Y,I), \subseteq\rangle$ *and* $\langle \mathrm{Int}(X,Y,I), \subseteq\rangle$ *are dually isomorphic, i.e., there is a mapping* $f : \mathrm{Ext}(X,Y,I) \to \mathrm{Int}(X,Y,I)$ *satisfying* $A_1 \subseteq A_2$ *iff* $f(A_2) \subseteq f(A_1)$.

3. $\langle \mathcal{B}(X,Y,I), \leq\rangle$ *is isomorphic to* $\langle \mathrm{Ext}(X,Y,I), \subseteq\rangle$.

4. $\langle \mathcal{B}(X,Y,I), \leq\rangle$ *is dually isomorphic to* $\langle \mathrm{Int}(X,Y,I), \subseteq\rangle$.

Proof. 1.: Obvious because $\mathrm{Ext}(X,Y,I)$ is a collection of subsets of X and $\subseteq$ is set inclusion. Same for $\mathrm{Int}(X,Y,I)$.

2.: Just take $f = {}^{\uparrow}$ and use previous results.

3.: Obviously, mapping $\langle A, B\rangle \mapsto A$ is the required isomorphism.

4.: Mapping $\langle A, B\rangle \mapsto B$ is the required dual isomorphism. □

We know that $\mathcal{B}(X,Y,I)$ (set of all formal concepts) equipped with $\leq$ (subconcept-superconcept hierarchy) is a partially ordered set. Now, the question is:

What is the structure of $\langle \mathcal{B}(X,Y,I), \leq\rangle$?

It turns out that $\langle \mathcal{B}(X,Y,I), \leq\rangle$ is a complete lattice (we will see this as a part of Main theorem of concept lattices). The fact that $\langle \mathcal{B}(X,Y,I), \leq\rangle$ is a lattice is a "welcome property". Namely, it says that for any collection $K \subseteq \mathcal{B}(X,Y,I)$ of formal concepts, $\mathcal{B}(X,Y,I)$ contains both the "direct generalization" $\bigvee K$ of concepts from K (supremum of K), and the "direct specialization" $\bigvee K$ of concepts from K (infimum of K). In this sense, $\langle \mathcal{B}(X,Y,I), \leq\rangle$ is a complete conceptual hierarchy. Let us now look at details.

We start with the following abstract theorem.

Theorem 2.34 (system of fixpoints of closure operators)**.** *For a closure operator* C *on* X, *the partially ordered set* $\langle \mathrm{fix}(C), \subseteq\rangle$ *of fixpoints of* C *is a complete lattice with infima and suprema given by*

$$\bigwedge_{j\in J} A_j = \bigcap_{j\in J} A_j, \tag{2.13}$$

$$\bigvee_{j\in J} A_j = C(\bigcup_{j\in J} A_j). \tag{2.14}$$

Proof. Evidently, $\langle \mathrm{fix}(C), \subseteq\rangle$ is a partially ordered set.

(2.13): First, we check that for $A_j \in \mathrm{fix}(C)$ we have $\bigcap_{j\in J} A_j \in \mathrm{fix}(C)$ (intersection of fixpoints is a fixpoint). We need to check $\bigcap_{j\in J} A_j = C(\bigcap_{j\in J} A_j)$.
"$\subseteq$": $\bigcap_{j\in J} A_j \subseteq C(\bigcap_{j\in J} A_j)$ is obvious (property of closure operators).
"$\supseteq$": We have $C(\bigcap_{j\in J} A_j) \subseteq \bigcap_{j\in J} A_j$ iff for each $j \in J$ we have $C(\bigcap_{j\in J} A_j) \subseteq A_j$

which is true. Indeed, we have $\bigcap_{j\in J} A_j \subseteq A_j$ from which we get $C(\bigcap_{j\in J} A_j) \subseteq C(A_j) = A_j$.

Now, since $\bigcap_{j\in J} A_j \in \mathrm{fix}(C)$, it is clear that $\bigcap_{j\in J} A_j$ is the infimum of A_j's: first, $\bigcap_{j\in J} A_j$ is less of equal to every A_j; second, $\bigcap_{j\in J} A_j$ is greater or equal to any $A \in \mathrm{fix}(C)$ which is less or equal to all A_j's; that is, $\bigcap_{j\in J} A_j$ is the greatest element of the lower cone of $\{A_j \mid j \in J\}$).

(2.14): We verify $\bigvee_{j\in J} A_j = C(\bigcup_{j\in J} A_j)$. Note first that since $\bigvee_{j\in J} A_j$ is a fixpoint of C, we have $\bigvee_{j\in J} A_j = C(\bigvee_{j\in J} A_j)$.

"$\subseteq$": $C(\bigcup_{j\in J} A_j)$ is a fixpoint which is greater or equal to every A_j, and so $C(\bigcup_{j\in J} A_j)$ must be greater or equal to the supremum $\bigvee_{j\in J} A_j$, i.e. $\bigvee_{j\in J} A_j \subseteq C(\bigcup_{j\in J} A_j)$.

"$\supseteq$": Since $\bigvee_{j\in J} A_j \supseteq A_j$ for any $j \in J$, we get $\bigvee_{j\in J} A_j \supseteq \bigcup_{j\in J} A_j$, and so $\bigvee_{j\in J} A_j = C(\bigvee_{j\in J} A_j) \supseteq C(\bigcup_{j\in J} A_j)$.

To sum up, $\bigvee_{j\in J} A_j = C(\bigcup_{j\in J} A_j)$. □

2.6 Main Theorem of Concept Lattices

The previous results enable us to formulate the following theorem characterizing the structure of concept lattices.

Theorem 2.35 (Main theorem of concept lattices, Wille (1982)). *(1) $\mathcal{B}(X, Y, I)$ is a complete lattice with infima and suprema given by*

$$\bigwedge_{j\in J} \langle A_j, B_j\rangle = \langle \bigcap_{j\in J} A_j, (\bigcup_{j\in J} B_j)^{\downarrow\uparrow}\rangle \, , \bigvee_{j\in J} \langle A_j, B_j\rangle = \langle (\bigcup_{j\in J} A_j)^{\uparrow\downarrow}, \bigcap_{j\in J} B_j\rangle \, . \tag{2.15}$$

(2) Moreover, an arbitrary complete lattice $\mathbf{V} = (V, \leq)$ is isomorphic to $\mathcal{B}(X, Y, I)$ iff there are mappings $\gamma : X \to V$, $\mu : Y \to V$ such that

(i) $\gamma(X)$ is $\bigvee$-dense in V, $\mu(Y)$ is $\bigwedge$-dense in V;

(ii) $\gamma(x) \leq \mu(y)$ iff $\langle x, y\rangle \in I$.

Remark 2.36. (1) Note that $K \subseteq V$ is supremally dense in V iff for each $v \in V$ there exists $K' \subseteq K$ such that $v = \bigvee K'$ (i.e., every element v of V is a supremum of some elements of K).
Dually for infimal density of K in V (every element v of V is an infimum of some elements of K).

(2) Supremally (infimally) dense sets canbe considered building blocks of V.

Proof. For part (1) of the Main Theorem only: We check $\bigwedge_{j\in J} \langle A_j, B_j\rangle = \langle \bigcap_{j\in J} A_j, (\bigcup_{j\in J} B_j)^{\downarrow\uparrow}\rangle$:

First, $\langle \mathrm{Ext}(X, Y, I), \subseteq\rangle = \langle \mathrm{fix}(^{\uparrow\downarrow}), \subseteq\rangle$ and $\langle \mathrm{Int}(X, Y, I), \subseteq\rangle = \langle \mathrm{fix}(^{\downarrow\uparrow}), \subseteq\rangle$. That is, $\mathrm{Ext}(X, Y, I)$ and $\mathrm{Int}(X, Y, I)$ are systems of fixpoints of closure operators, and therefore, suprema and infima in $\mathrm{Ext}(X, Y, I)$ and $\mathrm{Int}(X, Y, I)$ obey the formulas from previous theorem.

Second, recall that $\langle \mathcal{B}(X, Y, I), \leq\rangle$ is isomorphic to $\langle \mathrm{Ext}(X, Y, I), \subseteq\rangle$ and dually isomorphic to $\langle \mathrm{Int}(X, Y, I), \subseteq\rangle$.

Therefore, infima in $\mathcal{B}(X,Y,I)$ correspond to infima in $\mathrm{Ext}(X,Y,I)$ and to suprema in $\mathrm{Int}(X,Y,I)$.

That is, since $\bigwedge_{j\in J}\langle A_j,B_j\rangle$ is the infimum of $\langle A_j,B_j\rangle$'s in $\langle\mathcal{B}(X,Y,I),\leq\rangle$: The extent of $\bigwedge_{j\in J}\langle A_j,B_j\rangle$ is the infimum of A_j's in $\langle\mathrm{Ext}(X,Y,I),\subseteq\rangle$ which is, according to (2.13), $\bigcap_{j\in J}A_j$. The intent of $\bigwedge_{j\in J}\langle A_j,B_j\rangle$ is the supremum of B_j's in $\langle\mathrm{Int}(X,Y,I),\subseteq\rangle$ which is, according to (2.14), $(\bigcup_{j\in J}B_j)^{\downarrow\uparrow}$. We just proved

$$\bigwedge_{j\in J}\langle A_j,B_j\rangle=\langle\bigcap_{j\in J}A_j,(\bigcup_{j\in J}B_j)^{\downarrow\uparrow}\rangle.$$

Checking the formula for $\bigvee_{j\in J}\langle A_j,B_j\rangle$ is dual. □

Consider now part (2) of the Main Theorem and take $V:=\mathcal{B}(X,Y,I)$. Since $\mathcal{B}(X,Y,I)$ is isomorphic to $\mathcal{B}(X,Y,I)$, there exist mappings

$$\gamma: X\to\mathcal{B}(X,Y,I)\text{ and }\mu: Y\to\mathcal{B}(X,Y,I)$$

satisfying properties from part (2). How do mappings γ and μ work? One may put

$$\begin{aligned}\gamma(x) &= \langle\{x\}^{\uparrow\downarrow},\{x\}^{\uparrow}\rangle\ldots\text{object concept of }x,\\ \mu(y) &= \langle\{y\}^{\downarrow},\{y\}^{\downarrow\uparrow}\rangle\ldots\text{attribute concept of }y.\end{aligned}$$

Then: (i) says that each $\langle A,B\rangle\in\mathcal{B}(X,Y,I)$ is a supremum of some objects concepts (and, infimum of some attribute concepts). This is true since

$$\langle A,B\rangle=\bigvee_{x\in A}\langle\{x\}^{\uparrow\downarrow},\{x\}^{\uparrow}\rangle\text{ and }\langle A,B\rangle=\bigwedge_{y\in B}\langle\{y\}^{\downarrow},\{y\}^{\downarrow\uparrow}\rangle.$$

(ii) is true, too: $\gamma(x)\leq\mu(y)$ iff $\{x\}^{\uparrow\downarrow}\subseteq\{y\}^{\downarrow}$ iff $\{y\}\subseteq\{x\}^{\uparrow\downarrow\uparrow}=\{x\}^{\uparrow}$ iff $\langle x,y\rangle\in I$.

What does then the Main Theorem say? Part (1) says that $\mathcal{B}(X,Y,I)$ is a lattice and describes its infima and suprema. Part (2) provides a way to label a concept lattice so that no information is lost.

The labeling has two rules:
Since $\gamma(x)=\langle\{x\}^{\uparrow\downarrow},\{x\}^{\uparrow}\rangle$, an object concept of x is labeled by x,
since $\mu(y)=\langle\{y\}^{\downarrow},\{y\}^{\downarrow\uparrow}\rangle$, an attribute concept of y is labeled by y.

Now, how do we see extents and intents in a labeled Hasse diagram? Consider formal concept $\langle A,B\rangle$ corresponding to node c of a labeled diagram of concept lattice $\mathcal{B}(X,Y,I)$. What is then extent and the intent of $\langle A,B\rangle$?

$x\in A$ iff node with label x lies on a path going from c downwards,
$y\in B$ iff node with label y lies on a path going from c upwards.

One can verify correctness of the above labeling procedure.

Example 2.37. (1) Draw a labeled Hasse diagram of a concept lattice associated to formal context

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

(2) Is every formal concept either an object concept or an attribute concept? Can a formal concept be both an object concept and an attribute concept?

2.7 Clarification and Reduction of Formal Concepts

A formal context may be redundant in that one can remove some of its objects or attributes and get a formal context for which the associated concept lattice is isomorphic to that one of the original formal context. Two main notions in this regards are that of a clarified formal context and that of a reduced formal context.

Definition 2.38 (clarified context)**.** A formal context $\langle X, Y, I\rangle$ is called *clarified* if the corresponding table does neither contain identical rows nor identical columns.

That is, if $\langle X, Y, I\rangle$ is clarified then
$\{x_1\}^{\uparrow} = \{x_2\}^{\uparrow}$ implies $x_1 = x_2$ for every $x_1, x_2 \in X$;
$\{y_1\}^{\downarrow} = \{y_2\}^{\downarrow}$ implies $y_1 = y_2$ for every $y_1, y_2 \in Y$.

Clarification can therefore be performed by removing identical rows and columns (only one of several identical rows/columns is left).

Example 2.39. The formal context on the right results by clarification from the formal context on the left.

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

I	y_1	y_2	y_3
x_1	×	×	×
x_2	×		×
x_3		×	×
x_5	×		

Theorem 2.40. *If $\langle X_1, Y_1, I_1\rangle$ is a clarified context resulting from $\langle X_2, Y_2, I_2\rangle$ by clarification, then $\mathcal{B}(X_1, Y_1, I_1)$ is isomorphic to $\mathcal{B}(X_2, Y_2, I_2)$.*

Proof. Let $\langle X_2, Y_2, I_2\rangle$ contain x_1, x_2 s.t. $\{x_1\}^{\uparrow} = \{x_2\}^{\uparrow}$ (identical rows). Let $\langle X_1, Y_1, I_1\rangle$ result from $\langle X_2, Y_2, I_2\rangle$ by removing x_2 (i.e., $X_1 = X_2 - \{x_2\}$, $Y_1 = Y_2$). An isomorphism $f : \mathcal{B}(X_1, Y_1, I_1) \to \mathcal{B}(X_2, Y_2, I_2)$ is given by

$$f(\langle A_1, B_1\rangle) = \langle A_2, B_2\rangle$$

where $B_1 = B_2$ and

$$A_2 = \begin{cases} A_1 & \text{if } x_1 \notin A_1, \\ A_1 \cup \{x_2\} & \text{if } x_1 \in A_1. \end{cases}$$

Namely, one can easily see that $\langle A_1, B_1\rangle$ is a formal concept of $\mathcal{B}(X_1, Y_1, I_1)$ iff $f(\langle A_1, B_1\rangle)$ is a formal concept of $\mathcal{B}(X_2, Y_2, I_2)$ and that for formal concepts $\langle A_1, B_1\rangle, \langle C_1, D_1\rangle$ of $\mathcal{B}(X_1, Y_1, I_1)$ we have

$$\langle A_1, B_1\rangle \leq \langle C_1, D_1\rangle \text{ iff } f(\langle A_1, B_1\rangle) \leq f(\langle C_1, D_1\rangle).$$

Therefore, $\mathcal{B}(X_1, Y_1, I_1)$ is isomorphic to $\mathcal{B}(X_2, Y_2, I_2)$. This justifies the claim for removing one (identical) row. The same is true for removing one column. Repeated application gives the theorem. □

Example 2.41. Find the isomorphism between concept lattices of formal contexts from the previous example.

Another way to simplify the input formal context: removing reducible objects and attributes

Example 2.42. Draw concept lattices of the following formal contexts:

I	y_1	y_2	y_3
x_1			×
x_2	×	×	×
x_3	×		

I	y_1	y_3
x_1		×
x_2	×	×
x_3	×	

Why are they isomorphic?

(Hint: y_2 = intersection of y_1 and y_3 (i.e., $\{y_2\}^{\downarrow} = \{y_1\}^{\downarrow} \cap \{y_3\}^{\downarrow}$).)

This leads us to the following definition.

Definition 2.43 (reducible objects and attributes)**.** For a formal context $\langle X, Y, I\rangle$, an attribute $y \in Y$ is called *reducible* iff there is $Y' \subset Y$ with $y \notin Y'$ such that

$$\{y\}^{\downarrow} = \bigcap_{z \in Y'} \{z\}^{\downarrow},$$

i.e., the column corresponding to y is the intersection of columns corresponding to zs from Y'. An object $x \in X$ is called *reducible* iff there is $X' \subset X$ with $x \notin X'$ such that

$$\{x\}^{\uparrow} = \bigcap_{z \in X'} \{z\}^{\uparrow},$$

i.e., the row corresponding to x is the intersection of rows corresponding to zs from X'.

Note the following:

- y_2 from the previous example is reducible ($Y' = \{y_1, y_3\}$).
- Analogy: If a (real-valued attribute) y is a linear combination of other attributes, it can be removed (caution: this depends on what we do with the attributes). Intersection = particular attribute combination.
- (Non-)reducibility in $\langle X, Y, I\rangle$ is connected to so-called $\bigwedge$-(ir)reducibility and $\bigvee$-(ir)reducibility in $\mathcal{B}(X, Y, I)$.
- In a complete lattice $\langle V, \leq\rangle$, $v \in V$ is called $\bigwedge$-irreducible if there is no $U \subset V$ with $v \notin U$ s.t. $v = \bigwedge U$. Dually for $\bigvee$-irreducibility.
- Determine all $\bigwedge$-irreducible elements in $\langle 2^{\{a,b,c\}}, \subseteq\rangle$, in a "pentagon", and in a 4-element chain.
- Verify that in a finite lattice $\langle V, \leq\rangle$: v is $\bigwedge$-irreducible iff v is covered by exactly one element of V; v is $\bigvee$-irreducible iff v covers exactly one element of V.

Furthermore, note the following:

- easily from definition: y is reducible iff there is $Y' \subset Y$ with $y \notin Y'$ s.t.

$$\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle = \bigwedge_{z \in Y'} \langle \{z\}^{\downarrow}, \{z\}^{\downarrow\uparrow}\rangle. \tag{2.16}$$

- Let $\langle X, Y, I\rangle$ be clarified. Then in (2.16), for each $z \in Y'$: $\{y\}^{\downarrow} \neq \{z\}^{\downarrow}$, and so, $\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle \neq \langle \{z\}^{\downarrow}, \{z\}^{\downarrow\uparrow}\rangle$. Thus: y is reducible iff $\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle$ is an infimum of attribute concepts different from $\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle$. Now, since every concept $\langle A, B\rangle$ is an infimum of some attribute concepts (attribute concepts are $\bigwedge$-dense), we get that y is not reducible iff $\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle$ is $\bigwedge$-irreducible in $\mathcal{B}(X, Y, I)$.
- Therefore, if $\langle X, Y, I\rangle$ is clarified, y is not reducible iff $\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle$ is $\bigwedge$-irreducible.
- Suppose $\langle X, Y, I\rangle$ is not clarified due to $\{y\}^{\downarrow} = \{z\}^{\downarrow}$ for some $z \neq y$. Then y is reducible by definition (just put $Y' = \{z\}$ in the definition). Still, it can happen that $\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle$ is $\bigwedge$-irreducible and it can happen that y is $\bigwedge$-reducible, see the next example.

– Example. Two non-clarified contexts. Left: y_2 reducible and $\langle \{y_2\}^{\downarrow}, \{y_2\}^{\downarrow\uparrow}\rangle$ $\bigwedge$-reducible. Right: y_2 reducible but $\langle \{y_2\}^{\downarrow}, \{y_2\}^{\downarrow\uparrow}\rangle$ $\bigwedge$-irreducible.

I	y_1	y_2	y_3	y_4
x_1			×	
x_2	×	×	×	×
x_3	×	×	×	×
x_4	×			

I	y_1	y_2	y_3	y_4	y_5
x_1	×		×		
x_2		×		×	
x_3	×	×	×	×	
x_4	×		×		

– The same for reducibility of objects: If $\langle X, Y, I\rangle$ is clarified, then x is not reducible iff $\langle \{x\}^{\uparrow\downarrow}, \{x\}^{\uparrow}\rangle$ is $\bigvee$-irreducible in $\mathcal{B}(X, Y, I)$.

– Therefore, it is convenient to consider reducibility on clarified contexts (then, reducibility of objects and attributes corresponds to $\bigvee$- and $\bigwedge$-reducibility of object concepts and attribute concepts).

We now get the following theorem regarding reducibility.

Theorem 2.44. *Let $y \in Y$ be reducible in $\langle X, Y, I\rangle$. Then $\mathcal{B}(X, Y - \{y\}, J)$ is isomorphic to $\mathcal{B}(X, Y, I)$ where $J = I \cap (X \times (Y - \{y\}))$ is the restriction of I to $X \times Y - \{y\}$, i.e., $\langle X, Y - \{y\}, J\rangle$ results by removing column y from $\langle X, Y, I\rangle$.*

Proof. Follows from part (2) of Main theorem of concept lattices:
Namely, $\mathcal{B}(X, Y - \{y\}, J)$ is isomorphic to $\mathcal{B}(X, Y, I)$ iff there are mappings $\gamma : X \to \mathcal{B}(X, Y, I)$ and $\mu : Y - \{y\} \to \mathcal{B}(X, Y, I)$ such that (a) $\gamma(X)$ is $\bigvee$-dense in $\mathcal{B}(X, Y, I)$, (b) $\mu(Y - \{y\})$ is $\bigwedge$-dense in $\mathcal{B}(X, Y, I)$, and (c) $\gamma(x) \leq \mu(z)$ iff $\langle x, z\rangle \in J$.
If we define $\gamma(x)$ and $\mu(z)$ to be the object and attribute concept of $\mathcal{B}(X, Y, I)$ corresponding to x and z, respectively, then:
(a) is evident.
(c) is satisfied because for $z \in Y - \{z\}$ we have $\langle x, z\rangle \in J$ iff $\langle x, z\rangle \in I$ (J is a restriction of I).

(b): We need to show that each $\langle A, B\rangle \in \mathcal{B}(X, Y, I)$ is an infimum of attribute concepts different from $\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle$. But this is true because y is reducible: Namely, if $\langle A, B\rangle \in \mathcal{B}(X, Y, I)$ is the infimum of attribute concepts which include $\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle$, then we may replace $\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle$ by the attribute concepts $\langle \{z\}^{\downarrow}, \{z\}^{\downarrow\uparrow}\rangle$, $z \in Y'$ (cf. definition of reducible attribute), of which $\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle$ is the infimum. □

Definition 2.45 (reduced formal context)**.** $\langle X, Y, I\rangle$ is

– *row reduced* if no object $x \in X$ is reducible,
– *column reduced* if no attribute $y \in Y$ is reducible,
– *reduced* if it is both row reduced and column reduced.

Note that

– by the above observation: If $\langle X, Y, I\rangle$ is not clarified, then either some object is reducible (if there are identical rows) or some attribute is reducible (if there are identical columns). Therefore, if $\langle X, Y, I\rangle$ is reduced, it is clarified.

– The relationship between reducibility of objects/attributes and $\bigvee$- and $\bigwedge$-reducibility of object/attribute concepts gives us:

Remark 2.46. A clarified $\langle X, Y, I\rangle$ is

– row reduced iff every object concept is $\bigvee$-irreducible,
– column reduced iff every attribute concept is $\bigwedge$-irreducible.

How to find out which objects and attributes are reducible? A useful way is provided by so-called arrow relations.

Definition 2.47 (arrow relations)**.** For $\langle X, Y, I\rangle$, define relations $\nearrow$, $\swarrow$, and $\updownarrow$ between X and Y by

- $x \swarrow y$ iff $\langle x, y\rangle \notin I$ and if $\{x\}^{\uparrow} \subset \{x_1\}^{\uparrow}$ then $\langle x_1, y\rangle \in I$.
- $x \nearrow y$ iff $\langle x, y\rangle \notin I$ and if $\{y\}^{\downarrow} \subset \{y_1\}^{\downarrow}$ then $\langle x, y_1\rangle \in I$.
- $x \updownarrow y$ iff $x \swarrow y$ and $x \nearrow y$.

Therefore, if $\langle x, y\rangle \in I$ then none of $x \swarrow y$, $x \nearrow y$, $x \updownarrow y$ occurs. The arrow relations can therefore be entered in the table of $\langle X, Y, I\rangle$.

For

I	y_1	y_2	y_3	y_4
x_1	$\times$	$\times$	$\times$	$\times$
x_2	$\times$	$\times$		
x_3		$\times$	$\times$	$\times$
x_4		$\times$		
x_5		$\times$	$\times$	

we get

I	y_1	y_2	y_3	y_4
x_1	$\times$	$\times$	$\times$	$\times$
x_2	$\times$	$\times$	$\updownarrow$	$\swarrow$
x_3	$\updownarrow$	$\times$	$\times$	$\times$
x_4	$\nearrow$	$\times$	$\nearrow$	
x_5	$\nearrow$	$\times$	$\times$	$\updownarrow$

We have the following theorem connecting arrow relations and reducibility.

Theorem 2.48 (arrow relations and reducibility)**.** *For any $\langle X, Y, I\rangle$, $x \in X$, $y \in Y$:*

- $\langle \{x\}^{\uparrow\downarrow}, \{x\}^{\uparrow}\rangle$ *is $\bigvee$-irreducible iff there is $y \in Y$ s.t. $x \swarrow y$;*
- $\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle$ *is $\bigwedge$-irreducible iff there is $x \in Y$ s.t. $x \nearrow y$.*

Proof. Due to duality, we verify $\bigwedge$-irreducibility:
$x \nearrow y$ IFF
$x \notin \{y\}^{\downarrow}$ and for every y_1 with $\{y\}^{\downarrow} \subset \{y_1\}^{\downarrow}$ we have $x \in \{y_1\}^{\downarrow}$ IFF
$\{y\}^{\downarrow} \subset \bigcap_{y_1 : \{y\}^{\downarrow} \subset \{y_1\}^{\downarrow}}$ IFF
$\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle$ is not an infimum of other attribute concepts IFF
$\langle \{y\}^{\downarrow}, \{y\}^{\downarrow\uparrow}\rangle$ is $\bigwedge$-irreducible. $\square$

Consider the following problem:
INPUT: (arbitrary) formal context $\langle X_1, Y_1, I_1\rangle$
OUTPUT: a reduced context $\langle X_2, Y_2, I_2\rangle$

This problem can be solved by the following algorithm (verify using the above observations):

1. clarify $\langle X_1, Y_1, I_1\rangle$ to get a clarified context $\langle X_3, Y_3, I_3\rangle$ (removing identical rows and columns),
2. compute arrow relations $\swarrow$ and $\nearrow$ for $\langle X_3, Y_3, I_3\rangle$,
3. obtain $\langle X_2, Y_2, I_2\rangle$ from $\langle X_3, Y_3, I_3\rangle$ by removing objects x from X_3 for which there is no $y \in Y_3$ with $x \swarrow y$, and attributes y from Y_3 for which there is no $x \in X_3$ with $x \nearrow y$. That is:
   $X_2 = X_3 - \{x \mid$ there is no $y \in Y_3$ s. t. $x \swarrow y\}$,
   $Y_2 = Y_3 - \{y \mid$ there is no $x \in X_3$ s. t. $x \nearrow y\}$,
   $I_2 = I_3 \cap (X_2 \times Y_2)$.

Example 2.49. Compute arrow relations $\swarrow$, $\nearrow$, $\updownarrow$ for the following formal context:

I_1	y_1	y_2	y_3	y_4
x_1	$\times$	$\times$	$\times$	$\times$
x_2	$\times$	$\times$		
x_3		$\times$	$\times$	$\times$
x_4		$\times$		
x_5		$\times$	$\times$	

Start with $\nearrow$. We need to go through cells in the table not containing $\times$ and decide whether $\nearrow$ applies.
The first such cell corresponds to $\langle x_2, y_3\rangle$. By definition, $x_2 \nearrow y_3$ iff for each $y \in Y$

such that $\{y_3\}^\downarrow \subset \{y\}^\downarrow$ we have $x_2 \in \{y\}^\downarrow$. The only such y is y_2 for which we have $x_2 \in \{y_2\}^\downarrow$, hence $x_2 \nearrow y_3$.

And so on up to $\langle x_5, y_4\rangle$ for which we get $x_5 \nearrow y_4$.

Compute arrow relations $\swarrow$, $\nearrow$, $\updownarrow$ for the following formal context:

I_1	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×	×		
x_3		×	×	×
x_4		×		
x_5		×	×	

Continue with $\swarrow$. Go through cells in the table not containing × and decide whether $\swarrow$ applies. The first such cell corresponds to $\langle x_2, y_3\rangle$. By definition, $x_2 \swarrow y_3$ iff for each $x \in X$ such that $\{x_2\}^\uparrow \subset \{x\}^\uparrow$ we have $y_3 \in \{x\}^\uparrow$. The only such x is x_1 for which we have $y_3 \in \{x_1\}^\uparrow$, hence $x_2 \swarrow y_3$.

And so on up to $\langle x_5, y_4\rangle$ for which we get $x_5 \swarrow y_4$.

Compute arrow relations $\swarrow$, $\nearrow$, $\updownarrow$ for the following formal context (left):

I_1	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×	×		
x_3		×	×	×
x_4		×		
x_5		×	×	

I_1	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×	×	$\updownarrow$	$\swarrow$
x_3	$\updownarrow$	×	×	×
x_4	$\nearrow$	×	$\nearrow$	
x_5	$\nearrow$	×	×	$\updownarrow$

The arrow relations are indicated in the right table. Therefore, the corresponding reduced context is

I_2	y_1	y_3	y_4
x_2	×		
x_3		×	×
x_5		×	

For a complete lattice $\langle V, \leq\rangle$ and $v \in V$, denote

$$v_* = \bigvee_{u \in V, u < v} u,$$
$$v^* = \bigwedge_{u \in V, v < u} u.$$

Example 2.50. – Show that $x \swarrow y$ iff $\langle \{x\}^{\uparrow\downarrow}, \{x\}^\uparrow\rangle \vee \langle \{y\}^\downarrow, \{y\}^{\downarrow\uparrow}\rangle = \langle \{x\}^{\uparrow\downarrow}, \{x\}^\uparrow\rangle_* < \langle \{y\}^\downarrow, \{y\}^{\downarrow\uparrow}\rangle$,

– Show that $x \nearrow y$ iff $\langle \{x\}^{\uparrow\downarrow}, \{x\}^\uparrow\rangle \wedge \langle \{y\}^\downarrow, \{y\}^{\downarrow\uparrow}\rangle = \langle \{y\}^\downarrow, \{y\}^{\downarrow\uparrow}\rangle^* > \langle \{y\}^\downarrow, \{y\}^{\downarrow\uparrow}\rangle$.

Let $\langle X_1, Y_1, I_1\rangle$ be clarified, $X_2 \subseteq X_1$ and $Y_2 \subseteq Y_1$ be sets of irreducible objects and attributes, respectively, let $I_2 = I_1 \cap (X_2 \times Y_2)$ (restriction of I_1 to irreducible objects and attributes).
How can we obtain from concepts of $\mathcal{B}(X_1, Y_1, I_1)$ from those of $\mathcal{B}(X_2, Y_2, I_2)$? The answer is based on:

1. $\langle A_1, B_1\rangle \mapsto \langle A_1 \cap X_2, B_1 \cap Y_2\rangle$ is an isomorphism from $\mathcal{B}(X_1, Y_1, I_1)$ on $\mathcal{B}(X_2, Y_2, I_2)$.
2. therefore, each extent A_2 of $\mathcal{B}(X_2, Y_2, I_2)$ is of the form $A_2 = A_1 \cap X_2$ where A_1 is an extent of $\mathcal{B}(X_1, Y_1, I_1)$ (same for intents).
3. for $x \in X_1$: $x \in A_1$ iff $\{x\}^{\uparrow\downarrow} \cap X_2 \subseteq A_1 \cap X_2$,
for $y \in Y_1$: $y \in B_1$ iff $\{y\}^{\downarrow\uparrow} \cap Y_2 \subseteq B_1 \cap Y_2$.

Here, $^{\uparrow}$ and $^{\downarrow}$ are operators induced by $\langle X_1, Y_1, I_1\rangle$.

Therefore, given $\langle A_2, B_2\rangle \in \mathcal{B}(X_2, Y_2, I_2)$, the corresponding $\langle A_1, B_1\rangle \in \mathcal{B}(X_1, Y_1, I_1)$ is given by

$$\begin{aligned} A_1 &= A_2 \cup \{x \in X_1 - X_2 \mid \{x\}^{\uparrow\downarrow} \cap X_2 \subseteq A_2\}, && (2.17)\\ B_1 &= B_2 \cup \{y \in Y_1 - Y_2 \mid \{y\}^{\downarrow\uparrow} \cap Y_2 \subseteq B_2\}. && (2.18)\end{aligned}$$

Example 2.51. Left is a clarified formal context $\langle X_1, Y_1, I_1\rangle$, right is a reduced context $\langle X_2, Y_2, I_2\rangle$ (see previous example).

I_1	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×	×		
x_3		×	×	×
x_4		×		
x_5		×	×	

I_2	y_1	y_3	y_4
x_2	×		
x_3		×	×
x_5		×	

Determine $\mathcal{B}(X_1, Y_1, I_1)$ by first computing $\mathcal{B}(X_2, Y_2, I_2)$ and then using the method from the previous slide to obtain concepts $\mathcal{B}(X_1, Y_1, I_1)$ from the corresponding concepts from $\mathcal{B}(X_2, Y_2, I_2)$.

I_1	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×	×		
x_3		×	×	×
x_4		×		
x_5		×	×	

I_2	y_1	y_3	y_4
x_2	×		
x_3		×	×
x_5		×	

$\mathcal{B}(X_2, Y_2, I_2)$ consists of:
$\langle \emptyset, Y_2\rangle$, $\langle \{x_2\}, \{y_1\}\rangle$, $\langle \{x_3\}, \{y_3, y_4\}\rangle$, $\langle \{x_3, x_5\}, \{y_3\}\rangle$, $\langle X_2, \emptyset\rangle$.

We need to go through all $\langle A_2, B_2\rangle \in \mathcal{B}(X_2, Y_2, I_2)$ and determine the corresponding $\langle A_1, B_1\rangle \in \mathcal{B}(X_1, Y_1, I_1)$ using (2.17) and (2.18). Note: $X_1 - X_2 = \{x_1, x_4\}$, $Y_1 - Y_2 = \{y_2\}$.

1. for $\langle A_2, B_2\rangle = \langle \emptyset, Y_2\rangle$ we have
   $\{x_1\}^{\uparrow\downarrow} \cap X_2 = \{x_1\} \cap X_2 = \emptyset \subseteq A_2$, $\{x_4\}^{\uparrow\downarrow} \cap X_2 = X_1 \cap X_2 = X_2 \not\subseteq A_2$,
   hence $A_1 = A_2 \cup \{x_1\} = \{x_1\}$, and
   $\{y_2\}^{\downarrow\uparrow} \cap Y_2 = \{y_2\} \cap Y_2 = \emptyset \subseteq B_2$,
   hence $B_1 = B_2 \cup \{y_2\} = Y_1$. So, $\langle A_1, B_1\rangle = \langle \{x_1\}, Y_1\rangle$.

I_1	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×	×		
x_3		×	×	×
x_4		×		
x_5		×	×	

I_2	y_1	y_3	y_4
x_2	×		
x_3		×	×
x_5		×	

2. for $\langle A_2, B_2\rangle = \langle \{x_2\}, \{y_1\}\rangle$ we have
   $\{x_1\}^{\uparrow\downarrow} \cap X_2 = \emptyset \subseteq A_2$, $\{x_4\}^{\uparrow\downarrow} \cap X_2 = X_2 \not\subseteq A_2$,
   hence $A_1 = A_2 \cup \{x_1\} = \{x_1, x_2\}$, and
   $\{y_2\}^{\downarrow\uparrow} \cap Y_2 = \{y_2\} \cap Y_2 = \emptyset \subseteq B_2$,
   hence $B_1 = B_2 \cup \{y_2\} = \{y_1, y_2\}$. So, $\langle A_1, B_1\rangle = \langle \{x_1, x_2\}, \{y_1, y_2\}\rangle$.
3. for $\langle A_2, B_2\rangle = \langle \{x_3\}, \{y_3, y_4\}\rangle$ we have
   $\{x_1\}^{\uparrow\downarrow} \cap X_2 = \emptyset \subseteq A_2$, $\{x_4\}^{\uparrow\downarrow} \cap X_2 = X_2 \not\subseteq A_2$,
   hence $A_1 = A_2 \cup \{x_1\} = \{x_1, x_3\}$, and
   $\{y_2\}^{\downarrow\uparrow} \cap Y_2 = \{y_2\} \cap Y_2 = \emptyset \subseteq B_2$,
   hence $B_1 = B_2 \cup \{y_2\} = \{y_2, y_3, y_4\}$. So, $\langle A_1, B_1\rangle = \langle \{x_1, x_3\}, \{y_2, y_3, y_4\}\rangle$.

I_1	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×	×		
x_3		×	×	×
x_4		×		
x_5		×	×	

I_2	y_1	y_3	y_4
x_2	×		
x_3		×	×
x_5		×	

4. for $\langle A_2, B_2\rangle = \langle\{x_3, x_5\}, \{y_3\}\rangle$ we have
$\{x_1\}^{\uparrow\downarrow} \cap X_2 = \emptyset \subseteq A_2$, $\{x_4\}^{\uparrow\downarrow} \cap X_2 = X_2 \not\subseteq A_2$,
hence $A_1 = A_2 \cup \{x_1\} = \{x_1, x_3, x_5\}$, and
$\{y_2\}^{\downarrow\uparrow} \cap Y_2 = \{y_2\} \cap Y_2 = \emptyset \subseteq B_2$,
hence $B_1 = B_2 \cup \{y_2\} = \{y_2, y_3\}$. So, $\langle A_1, B_1\rangle = \langle\{x_1, x_3, x_5\}, \{y_2, y_3\}\rangle$.
5. for $\langle A_2, B_2\rangle = \langle X_2, \emptyset\rangle$ we have
$\{x_1\}^{\uparrow\downarrow} \cap X_2 = \emptyset \subseteq A_2$, $\{x_4\}^{\uparrow\downarrow} \cap X_2 = X_2 \subseteq A_2$,
hence $A_1 = A_2 \cup \{x_1, x_4\} = X_1$, and
$\{y_2\}^{\downarrow\uparrow} \cap Y_2 = \{y_2\} \cap Y_2 = \emptyset \subseteq B_2$,
hence $B_1 = B_2 \cup \{y_2\} = \{y_2\}$. So, $\langle A_1, B_1\rangle = \langle X_1, \{y_2\}\rangle$.

Example 2.52. Determine a reduced context from the following formal context. Use the reduced context to compute $\mathcal{B}(X, Y, I)$.

I	y_1	y_2	y_3	y_4	y_5
x_1					
x_2		×		×	
x_3		×	×	×	
x_4		×		×	×
x_5		×	×		
x_6		×	×	×	
x_7	×	×	×		

Hint: First clarify, then compute arrow relations.

2.8 Basic Algorithm For Computing Concept Lattices

We now consider the problem of computing concept lattices, i.e. following problem:
INPUT: formal context $\langle X, Y, I\rangle$,
OUTPUT: concept lattice $\mathcal{B}(X, Y, I)$ (possibly plus $\leq$)

- Sometimes one needs to compute the set $\mathcal{B}(X, Y, I)$ of formal concepts only.
- Sometimes one needs to compute both the set $\mathcal{B}(X, Y, I)$ and the conceptual hierarchy $\leq$. $\leq$ can be computed from $\mathcal{B}(X, Y, I)$ by definition of $\leq$. But this is not efficient. Algorithms exist which can compute $\mathcal{B}(X, Y, I)$ and $\leq$ simultaneously, which is more efficient (faster) than first computing $\mathcal{B}(X, Y, I)$ and then computing $\leq$.

A good survey on algorithms for computing Kuznetsov S. O., Obiedkov S. A.: Comparing performance of algorithms for generating concept lattices. *J. Experimental & Theoretical Artificial Intelligence* **14**(2003), 189–216.

We will describe NextClosure which can be considered a basic algorithm for computing $\mathcal{B}(X, Y, I)$. The following are the basic characteristics of this algorithm:

- author: Bernhard Ganter (1987)
- input: formal context $\langle X, Y, I\rangle$,
- output: $\mathrm{Int}(X, Y, I)$ ... all intents (dually, $\mathrm{Ext}(X, Y, I)$ ... all extents),
- list all intents (or extents) in lexicographic order,
- note that $\mathcal{B}(X, Y, I)$ can be reconstructed from $\mathrm{Int}(X, Y, I)$ due to

$$\mathcal{B}(X, Y, I) = \{\langle B^{\downarrow}, B\rangle \mid B \in \mathrm{Int}(X, Y, I)\},$$

– one of most popular algorithms, easy to implement,

– we present NextClosure for intents.

Suppose $Y = \{1, \dots, n\}$ (that is, we denote attributes by positive integers, this way, we fix an ordering of attributes).

Definition 2.53 (lexicographic ordering of sets of attributes). For $A, B \subseteq Y$, $i \in \{1, \dots, n\}$ put

$$\begin{aligned} A <_i B \quad &\text{iff} \quad i \in B - A \text{ a } A \cap \{1, \dots, i-1\} = B \cap \{1, \dots, i-1\}, \\ A < B \quad &\text{iff} \quad A <_i B \text{ for some } i. \end{aligned}$$

Note that $<$ is a lexicographic ordering (thus, every two distinct sets $A, B \subseteq$ are comparable w.r.t. $<$). For $i = 1$, we put $\{1, \dots, i-1\} = \emptyset$. Note also that one may think of $B \subseteq Y$ in terms of its characteristic vector. For $Y = \{1,2,3,4,5,6,7\}$ and $B = \{1,3,4,6\}$, the characteristic vector of B is 1011010.

Example 2.54. Let $Y = \{1,2,3,4,5,6\}$, consider sets $\{1\}$, $\{2\}$, $\{2,3\}$, $\{3,4,5\}$, $\{3,6\}$, $\{1,4,5\}$. We have

- $\{2\} <_1 \{1\}$ because $1 \in \{1\} - \{2\} = \{1\}$ and $A \cap \emptyset = B \cap \emptyset$. Characteristic vectors: $010000 <_1 100000$.
- $\{3,6\} <_4 \{3,4,5\}$ because $4 \in \{3,4,5\} - \{3,6\} = \{4,5\}$ and $A \cap \{1,2,3\} = B \cap \{1,2,3\}$. Characteristic vectors: $001001 <_4 001110$.
- All sets ordered lexicographically:
  $\{3,6\} <_4 \{3,4,5\} <_2 \{2\} <_3 \{2,3\} <_1 \{1\} <_4 \{1,4,5\}$.
  Characteristic vectors:
  $001001 <_4 001110 <_2 010000 <_3 011000 <_1 100000 <_4 100110$.

Note that if $B_1 \subset B_2$ then $B_1 < B_2$.

Definition 2.55. For $A \subseteq Y$, $i \in \{1, \dots, n\}$, put

$$A \oplus i := ((A \cap \{1, \dots, i-1\}) \cup \{i\})^{\downarrow\uparrow}.$$

Example 2.56.

I	1	2	3	4
x_1	×	×		×
x_2	×	×	×	×
x_3		×		

- $A = \{1,3\}$, $i = 2$.
  $A \oplus i = ((\{1,3\} \cap \{1,2\}) \cup \{2\})^{\downarrow\uparrow} = (\{1\} \cup \{2\})^{\downarrow\uparrow} = \{1,2\}^{\downarrow\uparrow} = \{1,2,4\}$.
- $A = \{2\}$, $i = 1$.
  $A \oplus i = ((\{2\} \cap \emptyset) \cup \{1\})^{\downarrow\uparrow} = \{1\}^{\downarrow\uparrow} = \{1,2,4\}$.

Lemma 2.57. *For any $B, D, D_1, D_2 \subseteq Y$:*

(1) If $B <_i D_1$, $B <_j D_2$, and $i < j$ then $D_2 <_i D_1$;

(2) if $i \notin B$ then $B < B \oplus i$;

(3) if $B <_i D$ and $D = D^{\downarrow\uparrow}$ then $B \oplus i \subseteq D$;

(4) if $B <_i D$ and $D = D^{\downarrow\uparrow}$ then $B <_i B \oplus i$.

Proof. (1) by easy inspection.

(2) is true because $B \cap \{1, \ldots, i-1\} \subseteq B \oplus i \cap \{1, \ldots, i-1\}$ and $i \in (B \oplus i) - B$.

(3) Putting $C_1 = B \cap \{1, \ldots, i-1\}$ and $C_2 = \{i\}$ we have $C_1 \cup C_2 \subseteq D$, and so $B \oplus i = (C_1 \cup C_2)^{\downarrow\uparrow} \subseteq D^{\downarrow\uparrow} = D$.

(4) By assumption, $B \cap \{1, \ldots, i-1\} = D \cap \{1, \ldots, i-1\}$. Furthermore, (3) yields $B \oplus i \subseteq D$ and so $B \cap \{1, \ldots, i-1\} \supseteq B \oplus i \cap \{1, \ldots, i-1\}$. On the other hand, $B \oplus i \cap \{1, \ldots, i-1\} \supseteq (B \cap \{1, \ldots, i-1\})^{\downarrow\uparrow} \cap \{1, \ldots, i-1\} \supseteq B \cap \{1, \ldots, i-1\}$. Therefore, $B \cap \{1, \ldots, i-1\} = B \oplus i \cap \{1, \ldots, i-1\}$. Finally, $i \in B \oplus i$. $\square$

The following is a main theorem we need for the NextClosure algorithm.

Theorem 2.58 (lexicographic successor)**.** *The least intent B^+ greater (w.r.t. $<$) than $B \subseteq Y$ is given by*

$$B^+ = B \oplus i$$

where i is the greatest one with $B <_i B \oplus i$.

Proof. Let B^+ be the least intent greater than B (w.r.t. $<$). We have $B < B^+$ and thus $B <_i B^+$ for some i such that $i \in B^+$. By Lemma (4), $B <_i B \oplus i$, i.e. $B < B \oplus i$. Lemma (3) yields $B \oplus i \leq B^+$ which gives $B^+ = B \oplus i$ since B^+ is the least intent with $B < B^+$. It remains to show that i is the greatest one satisfying $B <_i B \oplus i$. Suppose $B <_k B \oplus k$ for $k > i$. By Lemma (1), $B \oplus k <_i B \oplus i$ which is a contradiction to $B \oplus i = B^+ < B \oplus k$ (B^+ is the least intent greater than B and so $B^+ < B \oplus k$). Therefore we have $k = i$. $\square$

pseudo-code of NextClosure algorithm:

```
A:=∅↓↑; (leastIntent)
store(A);
while not(A=Y) do
    A:=A+;
    store(A);
endwhile.
```

Note that the time complexity of computing A^+ is $O(|X| \cdot |Y|^2)$:
complexity of computing $C^{\uparrow}$ is $O(|X| \cdot |Y|)$, for $D^{\downarrow}$ it is $O(|X| \cdot |Y|)$, thus for $D^{\downarrow\uparrow}$ it is $O(|X| \cdot |Y|)$; complexity of computing $A \oplus i$ is thus $O(|X| \cdot |Y|)$; to get A^+ we need to compute $A \oplus i$ $|Y|$-times in the worst case. As a result, complexity of computing A^+ is $O(|X| \cdot |Y|^2)$.

Therefore, the time complexity of NextClosure is $O(|X| \cdot |Y|^2 \cdot |\mathcal{B}(X, Y, I)|)$.

Note also that NextClosure has almost no space requirements. However, NextClosure does not directly give information about $\leq$.

Example 2.59 (NextClosure Algorithm – simulation)**.** Simulate NextClosure algorithm on the following example.

I	1	2	3
x_1	×	×	×
x_2	×		×
x_3		×	×
x_4	×		

1. $A = \emptyset^{\downarrow\uparrow} = \emptyset$.
2. Next, we are looking for A^+, i.e. $\emptyset^+$, which is $A \oplus i$ s.t. i is the largest one with $A <_i A \oplus i$. We proceed for $i = 3, 2, 1$ and test whether $A <_i A \oplus i$:

- $i = 3$: $A \oplus i = \{3\}^{\downarrow\uparrow} = \{3\}$ and $\emptyset <_3 \{3\} = A \oplus i$, therefore $A^+ = \{3\}$.

3. Next, $\{3\}^+$:
   - $i = 3$: $A \oplus i = \{3\}^{\downarrow\uparrow} = \{3\}$ and $\{3\} \not<_3 \{3\} = A \oplus i$, therefore we proceed for $i = 2$.
   - $i = 2$: $A \oplus i = \{2\}^{\downarrow\uparrow} = \{2, 3\}$ and $\{3\} <_2 \{2, 3\} = A \oplus i$, therefore $A^+ = \{2, 3\}$.

4. Next, $\{2, 3\}^+$:
   - $i = 3$: $A \oplus i = \{2, 3\}^{\downarrow\uparrow} = \{2, 3\}$ and $\{2, 3\} \not<_3 \{2, 3\} = A \oplus i$, therefore we proceed for $i = 2$.
   - $i = 2$: $A \oplus i = \{2\}^{\downarrow\uparrow} = \{2, 3\}$ and $\{2, 3\} \not<_2 \{2, 3\} = A \oplus i$, therefore we proceed for $i = 1$.
   - $i = 1$: $A \oplus i = \{1\}^{\downarrow\uparrow} = \{1\}$ and $\{2, 3\} <_1 \{1\} = A \oplus i$, therefore we $A^+ = \{1\}$.

5. Next, $\{1\}^+$:
   - $i = 3$: $A \oplus i = \{1, 3\}^{\downarrow\uparrow} = \{1, 3\}$ and $\{1\} <_3 \{1, 3\} = A \oplus i$, therefore $A^+ = \{1, 3\}$.

6. Next, $\{1, 3\}^+$:
   - $i = 3$: $A \oplus i = \{1, 3\}^{\downarrow\uparrow} = \{1, 3\}$ and $\{1, 3\} \not<_3 \{1, 3\} = A \oplus i$, therefore we proceed for $i = 2$.
   - $i = 2$: $A \oplus i = \{1, 2\}^{\downarrow\uparrow} = \{1, 2, 3\}$ and $\{1, 3\} <_2 \{1, 2, 3\} = A \oplus i$, therefore $A^+ = \{1, 2, 3\} = Y$.

Therefore, the intents from $\mathrm{Int}(X, Y, I)$, ordered lexicographically, are:
$\emptyset < \{3\} < \{2, 3\} < \{1\} < \{1, 3\} < \{1, 2, 3\}$.

I	1	2	3
x_1	×	×	×
x_2	×		×
x_3		×	×
x_4	×		

$\mathrm{Int}(X, Y, I) = \{\emptyset, \{3\}, \{2, 3\}, \{1\}, \{1, 3\}, \{1, 2, 3\}\}$.

From this list, we can get the corresponding extents:
$X = \emptyset^{\downarrow}$, $\{x_1, x_2, x_3\} = \{3\}^{\downarrow}$, $\{x_1, x_3\} = \{2, 3\}^{\downarrow}$, $\{x_1, x_3, x_4\} = \{1\}^{\downarrow}$, $\{x_1, x_2\} = \{1, 3\}^{\downarrow}$, $\{x_1\} = \{1, 2, 3\}^{\downarrow}$.

Therefore, $\mathcal{B}(X, Y, I)$ consists of: $\langle\{x_1\}, \{1, 2, 3\}\rangle$, $\langle\{x_1, x_2\}, \{1, 3\}\rangle$, $\langle\{x_1, x_3\}, \{2, 3\}\rangle$, $\langle\{x_1, x_2, x_3\}, \{3\}\rangle$, $\langle\{x_1, x_2, x_4\}, \{1\}\rangle$, $\langle\{x_1, x_2, x_3, x_4\}, \emptyset\rangle$.

Note the following:

- If ${}^{\downarrow\uparrow}$ is replaced by an arbitrary closure operator C, NextClosure computes all fixpoints of C. This is easy to see: all that matters in the proofs of Theorem and Lemma justifying correctness of NextClosure, is that ${}^{\downarrow\uparrow}$ is a closure operator.

- Therefore, NextClosure is essentially an algorithm for computing all fixpoints of a given closure operator C.

- Computational complexity of NextClosure depends on computational complexity of computing $C(A)$ (computing closure of arbitrary set A).

3 Attribute Implications

Goals: This chapter provides basic information regarding particular attribute dependencies in cross-tables. These dependencies are called attribute implications.

Keywords: attribute implication, attribute dependency, entailment, non-redundant basis, Armstrong axioms.

3.1 Basic Notions Regarding Attribute Implications

Attribute implications represent data dependencies such as

- every number divisible by 2 and 3 is divisible by 6,
- every patient with symptom s_2 and symptom s_5 has also symptom s_1 and symptom s_3.

Definition 3.1 (attribute implication)**.** Let Y be a non-empty set (of attributes). An *attribute implication* over Y is an expression

$$A \Rightarrow B$$

where $A \subseteq Y$ and $B \subseteq Y$ (A and B are sets of attributes).

Example 3.2.
- Let $Y = \{y_1, y_2, y_3, y_4\}$. Then $\{y_2, y_3\} \Rightarrow \{y_1, y_4\}$, $\{y_2, y_3\} \Rightarrow \{y_1, y_2, y_3\}$, $\emptyset \Rightarrow \{y_1, y_2\}$, $\{y_2, y_4\} \Rightarrow \emptyset$ are AIs over Y.
- Let $Y = \{$watches-TV, eats-unhealthy-food, runs-regularly, normal-blood-pressure, high-blood-pressure$\}$. Then
  $\{$watches-TV, eats-unhealthy-food$\} \Rightarrow \{$high-blood-pressure$\}$,
  $\{$runs-regularly$\} \Rightarrow \{$normal-blood-pressure$\}$
  are attribute implications over Y.

Basic semantic structures in which we evaluate attribute implications are rows of tables (of formal contexts). Table rows can be regarded as sets of attributes. In table

	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×			×
x_3				

,

rows corresponding to x_1, x_2, and x_3 can be regarded as sets $M_1 = \{y_1, y_2, y_3, y_4\}$, $M_2 = \{y_1, y_4\}$, and $M_3 = \emptyset$.

Therefore, we need to define a notion of a validity of an AI in a set M of attributes.

Definition 3.3 (validity of attribute implication)**.** An attribute implication $A \Rightarrow B$ over Y is *true* (valid) in a set $M \subseteq Y$ iff

$$A \subseteq M \text{ implies } B \subseteq M.$$

- We write
$$||A \Rightarrow B||_M = \begin{cases} 1 & \text{if } A \Rightarrow B \text{ is true in } M, \\ 0 & \text{if } A \Rightarrow B \text{ is not true in } M. \end{cases}$$
- Let M be a set of attributes of some object x. $||A \Rightarrow B||_M = 1$ says "if x has all attributes from A then x has all attributes from B", because "if x has all attributes from C" is equivalent to $C \subseteq M$.

Example 3.4. Let $Y = \{y_1, y_2, y_3, y_4\}$. Then

$A \Rightarrow B$	M	$\|\|A \Rightarrow B\|\|_M$	why
$\{y_2, y_3\} \Rightarrow \{y_1\}$	$\{y_2\}$	1	$A \not\subseteq M$
$\{y_2, y_3\} \Rightarrow \{y_1\}$	$\{y_1, y_2\}$	1	$A \not\subseteq M$
$\{y_2, y_3\} \Rightarrow \{y_1\}$	$\{y_1, y_2, y_3\}$	1	$A \subseteq M$ and $B \subseteq M$
$\{y_2, y_3\} \Rightarrow \{y_1\}$	$\{y_2, y_3, y_4\}$	0	$A \subseteq M$ but $B \not\subseteq M$
$\{y_2, y_3\} \Rightarrow \{y_1\}$	$\emptyset$	1	$A \not\subseteq \emptyset$
$\emptyset \Rightarrow \{y_1\}$	$\{y_1, y_4\}$	1	$\emptyset \subseteq M$ and $B \subseteq M$.
$\emptyset \Rightarrow \{y_1\}$	$\{y_3, y_4\}$	0	$\emptyset \subseteq M$ but $B \not\subseteq M$.
$\{y_2, y_3\} \Rightarrow \emptyset$	any M	1	$\emptyset \subseteq M$

We now extend the validity of $A \Rightarrow B$ to collections $\mathcal{M}$ of M's (collections of subsets of attributes), i.e. define validity of $A \Rightarrow B$ in $\mathcal{M} \subseteq 2^Y$.

Definition 3.5. Let $\mathcal{M} \subseteq 2^Y$ (elements of $\mathcal{M}$ are subsets of attributes). An attribute implication $A \Rightarrow B$ over Y is true (valid) in $\mathcal{M}$ if $A \Rightarrow B$ is true in each $M \in \mathcal{M}$.

– Again,

$$||A \Rightarrow B||_{\mathcal{M}} = \begin{cases} 1 & \text{if } A \Rightarrow B \text{ is true in } \mathcal{M}, \\ 0 & \text{if } A \Rightarrow B \text{ is not true in } \mathcal{M}. \end{cases}$$

Therefore, $||A \Rightarrow B||_{\mathcal{M}} = \min_{M \in \mathcal{M}} ||A \Rightarrow B||_M$.

Definition 3.6 (validity of attribute implications in formal contexts)**.** An attribute implication $A \Rightarrow B$ over Y is true in a table (formal context) $\langle X, Y, I\rangle$ iff $A \Rightarrow B$ is true in

$$\mathcal{M} = \{\{x\}^\uparrow \mid x \in X\}.$$

– We write $||A \Rightarrow B||_{\langle X,Y,I\rangle} = 1$ if $A \Rightarrow B$ is true in $\langle X, Y, I\rangle$.

– Note that, $\{x\}^\uparrow$ is the set of attributes of x (row corresponding to x). Hence, $\mathcal{M} = \{\{x\}^\uparrow \mid x \in X\}$ is the collection whose members are just sets of attributes of objects (i.e., rows) of $\langle X, Y, I\rangle$. Therefore, $||A \Rightarrow B||_{\langle X,Y,I\rangle} = 1$ iff $A \Rightarrow B$ is true in each row of $\langle X, Y, I\rangle$ iff
for each $x \in X$:
 if x has all attributes from A then x has all attributes from B.

Example 3.7. Consider attributes normal blood pressure (nbp), high blood pressure (hbp), watches TV (TV), eats unhealthy food (uf), runs regularly (r), and table

I	nbp	hbp	TV	uf	r
a	×				×
b	×			×	×
c		×	×	×	
d		×		×	
e	×				

Then

$A \Rightarrow B$	$\|\|A \Rightarrow B\|\|_{\langle X,Y,I\rangle}$	why
{r} ⇒ {nbp}	1	
{TV,uf} ⇒ {hbp}	1	
{TV} ⇒ {hbp}	1	
{uf} ⇒ {hbp}	0	b counterexample
{nbp} ⇒ {r}	0	e counterexample
{nbp,hbp} ⇒ {r,TV}	1	A never satisfied
{uf,r} ⇒ {r}	1	

– In the previous example: {TV,uf} ⇒ {hbp} intuitively follows from {TV} ⇒ {hbp}. Therefore, provided we establish validity of {TV} ⇒ {hbp}, AI {TV,uf} ⇒ {hbp} is redundant.
Another example: $A \Rightarrow C$ follows from $A \Rightarrow B$ and $B \Rightarrow C$ (for any A, B, C).

– We need to capture intuitive notion of entailment of attribute implications. We use standard notions of a theory and model.

– Eventually, we want to have a small set T of AIs which are valid in $\langle X, Y, I\rangle$ such that all other AIs which are true in $\langle X, Y, I\rangle$ follow from T.

Definition 3.8 (theory, model)**.** A *theory* (over Y) is any set T of attribute implications (over Y).
A *model* of a theory T is any $M \subseteq Y$ such that every $A \Rightarrow B$ from T is true in M.

- $\mathrm{Mod}(T)$ denotes all models of a theory T, i.e. $$\mathrm{Mod}(T) = \{M \subseteq Y \mid \text{for each } A \Rightarrow B \in T : A \Rightarrow B \text{ is true in } M\}.$$
- Intuitively, a theory is some "important" set of attribute implications. For instance, T may contain AIs established to be true in data (extracted from data).
- Intuitively, a model of T is (a set of attributes of some) object which satisfies every AI from T.
- Notions of theory and model do not depend on some particular $\langle X, Y, I\rangle$.

Example 3.9 (theories over $\{y_1, y_2, y_3\}$).
- $T_1 = \{\{y_3\} \Rightarrow \{y_1, y_2\}, \{y_1, y_3\} \Rightarrow \{y_2\}\}$.
- $T_2 = \{\{y_3\} \Rightarrow \{y_1, y_2\}\}$.
- $T_3 = \{\{y_1, y_3\} \Rightarrow \{y_2\}\}$.
- $T_4 = \{\{y_1\} \Rightarrow \{y_3\}, \{y_3\} \Rightarrow \{y_1\}, \{y_2\} \Rightarrow \{y_2\}\}$.
- $T_5 = \emptyset$.
- $T_6 = \{\emptyset \Rightarrow \{y_1\}, \emptyset \Rightarrow \{y_3\}\}$.
- $T_7 = \{\{y_1\} \Rightarrow \emptyset, \{y_2\} \Rightarrow \emptyset, \{y_3\} \Rightarrow \emptyset\}$.
- $T_8 = \{\{y_1\} \Rightarrow \{y_2\}, \{y_2\} \Rightarrow \{y_3\}, \{y_3\} \Rightarrow \{y_1\}\}$.

Example 3.10 (models of theories over $\{y_1, y_2, y_3\}$). Determine $\mathrm{Mod}(T)$ of the following theories over $\{y_1, y_2, y_3\}$.

- $T_1 = \{\{y_3\} \Rightarrow \{y_1, y_2\}, \{y_1, y_3\} \Rightarrow \{y_2\}\}$.
  $\mathrm{Mod}(T_1) = \{\emptyset, \{y_1\}, \{y_2\}, \{y_1, y_2\}, \{y_1, y_2, y_3\}\}$,
- $T_2 = \{\{y_3\} \Rightarrow \{y_1, y_2\}\}$.
  $\mathrm{Mod}(T_2) = \{\emptyset, \{y_1\}, \{y_2\}, \{y_1, y_2\}, \{y_1, y_2, y_3\}\}$ (note: $T_2 \subset T_1$ but $\mathrm{Mod}(T_1) = \mathrm{Mod}(T_2)$),
- $T_3 = \{\{y_1, y_3\} \Rightarrow \{y_2\}\}$.
  $\mathrm{Mod}(T_3) = \{\emptyset, \{y_1\}, \{y_2\}, \{y_3\}, \{y_1, y_2\}, \{y_2, y_3\}, \{y_1, y_2, y_3\}\}$ (note: $T_3 \subset T_1$, $\mathrm{Mod}(T_1) \subset \mathrm{Mod}(T_2)$),
- $T_4 = \{\{y_1\} \Rightarrow \{y_3\}, \{y_3\} \Rightarrow \{y_1\}, \{y_2\} \Rightarrow \{y_2\}\}$.
  $\mathrm{Mod}(T_4) = \{\emptyset, \{y_2\}, \{y_1, y_3\}, \{y_1, y_2, y_3\}\}$
- $T_5 = \emptyset$. $\mathrm{Mod}(T_5) = 2^{\{y_1, y_2, y_3\}}$. Why: $M \in \mathrm{Mod}(T)$ iff for each $A \Rightarrow B$: if $A \Rightarrow B \in T$ then $||A \Rightarrow B||_M = 1$.
- $T_6 = \{\emptyset \Rightarrow \{y_1\}, \emptyset \Rightarrow \{y_3\}\}$. $\mathrm{Mod}(T_6) = \{\{y_1, y_3\}, \{y_1, y_2, y_3\}\}$.
- $T_7 = \{\{y_1\} \Rightarrow \emptyset, \{y_2\} \Rightarrow \emptyset, \{y_3\} \Rightarrow \emptyset\}$. $\mathrm{Mod}(T_7) = 2^{\{y_1, y_2, y_3\}}$.
- $T_8 = \{\{y_1\} \Rightarrow \{y_2\}, \{y_2\} \Rightarrow \{y_3\}, \{y_3\} \Rightarrow \{y_1\}\}$. $\mathrm{Mod}(T_8) = \{\emptyset, \{y_1, y_2, y_3\}\}$.

The next notion we deal with is that of a semantic consequence (entailment).

Definition 3.11 (semantic consequence)**.** An attribute implication $A \Rightarrow B$ *follows semantically* from a theory T, which is denoted by $$T \models A \Rightarrow B,$$ iff $A \Rightarrow B$ is true in every model M of T,

- Therefore, $T \models A \Rightarrow B$ iff for each $M \subseteq Y$: if $M \in \mathrm{Mod}(T)$ then $||A \Rightarrow B||_M = 1$.
- Intuitively, $T \models A \Rightarrow B$ iff $A \Rightarrow B$ is true in every situation where every AI from T is true (replace "situation" by "model").
- Later on, we will see how to efficiently check whether $T \models A \Rightarrow B$.
- Terminology: $T \models A \Rightarrow B \ldots A \Rightarrow B$ follows semantically from $T \ldots A \Rightarrow B$ is semantically entailed by $T \ldots A \Rightarrow B$ is a semantic consequence of T.

How to decide by definition whether $T \models A \Rightarrow B$?

1. Determine $\mathrm{Mod}(T)$.
2. Check whether $A \Rightarrow B$ is true in every $M \in \mathrm{Mod}(T)$; if yes then $T \models A \Rightarrow B$; if not then $T \not\models A \Rightarrow B$.

Example 3.12 (semantic entailment)**.** Let $Y = \{y_1, y_2, y_3\}$. Determine whether $T \models A \Rightarrow B$.

- $T = \{\{y_3\} \Rightarrow \{y_1, y_2\}, \{y_1, y_3\} \Rightarrow \{y_2\}\}$, $A \Rightarrow B$ is $\{y_2, y_3\} \Rightarrow \{y_1\}$.
  1. $\mathrm{Mod}(T) = \{\emptyset, \{y_1\}, \{y_2\}, \{y_1, y_2\}, \{y_1, y_2, y_3\}\}$.
  2. $||\{y_2, y_3\} \Rightarrow \{y_1\}||_\emptyset = 1$, $||\{y_2, y_3\} \Rightarrow \{y_1\}||_{\{y_1\}} = 1$, $||\{y_2, y_3\} \Rightarrow \{y_1\}||_{\{y_2\}} = 1$, $||\{y_2, y_3\} \Rightarrow \{y_1\}||_{\{y_1,y_2\}} = 1$, $||\{y_2, y_3\} \Rightarrow \{y_1\}||_{\{y_1,y_2,y_3\}} = 1$.
  Therefore, $T \models A \Rightarrow B$.
- $T = \{\{y_3\} \Rightarrow \{y_1, y_2\}, \{y_1, y_3\} \Rightarrow \{y_2\}\}$, $A \Rightarrow B$ is $\{y_2\} \Rightarrow \{y_1\}$.
  1. $\mathrm{Mod}(T) = \{\emptyset, \{y_1\}, \{y_2\}, \{y_1, y_2\}, \{y_1, y_2, y_3\}\}$.
  2. $||\{y_2\} \Rightarrow \{y_1\}||_\emptyset = 1$, $||\{y_2\} \Rightarrow \{y_1\}||_{\{y_1\}} = 1$, $||\{y_2\} \Rightarrow \{y_1\}||_{\{y_2\}} = 0$, we can stop.
  Therefore, $T \not\models A \Rightarrow B$.

Example 3.13. Let $Y = \{y_1, y_2, y_3\}$. Determine whether $T \models A \Rightarrow B$.

- $T_1 = \{\{y_3\} \Rightarrow \{y_1, y_2\}, \{y_1, y_3\} \Rightarrow \{y_2\}\}$.
  $A \Rightarrow B$: $\{y_1, y_2\} \Rightarrow \{y_3\}$, $\emptyset \Rightarrow \{y_1\}$.
- $T_2 = \{\{y_3\} \Rightarrow \{y_1, y_2\}\}$.
  $A \Rightarrow B$: $\{y_3\} \Rightarrow \{y_2\}$, $\{y_3, y_2\} \Rightarrow \emptyset$.
- $T_3 = \{\{y_1, y_3\} \Rightarrow \{y_2\}\}$.
  $A \Rightarrow B$: $\{y_3\} \Rightarrow \{y_1, y_2\}$, $\Rightarrow \emptyset$.
- $T_4 = \{\{y_1\} \Rightarrow \{y_3\}, \{y_3\} \Rightarrow \{y_2\}, \}$.
  $A \Rightarrow B$: $\{y_1\} \Rightarrow \{y_2\}$, $\{y_1\} \Rightarrow \{y_1, y_2, y_3\}$.
- $T_5 = \emptyset$.
  $A \Rightarrow B$: $\{y_1\} \Rightarrow \{y_2\}$, $\{y_1\} \Rightarrow \{y_1, y_2, y_3\}$.
- $T_6 = \{\emptyset \Rightarrow \{y_1\}, \emptyset \Rightarrow \{y_3\}\}$.
  $A \Rightarrow B$: $\{y_1\} \Rightarrow \{y_3\}$, $\emptyset \Rightarrow \{y_1, y_3\}$ $\{y_1\} \Rightarrow \{y_2\}$.
- $T_7 = \{\{y_1\} \Rightarrow \emptyset, \{y_2\} \Rightarrow \emptyset, \{y_3\} \Rightarrow \emptyset\}$.
  $A \Rightarrow B$: $\{y_1, y_2\} \Rightarrow \{y_3\}$, $\{y_1, y_2\} \Rightarrow \emptyset$.
- $T_8 = \{\{y_1\} \Rightarrow \{y_2\}, \{y_2\} \Rightarrow \{y_3\}, \{y_3\} \Rightarrow \{y_1\}\}$.
  $A \Rightarrow B$: $\{y_1\} \Rightarrow \{y_3\}$, $\{y_1, y_3\} \Rightarrow \{y_2\}$.

3.2 Armstrong Rules and Reasoning With Attribute Implications

– Some attribute implications semantically follow from others.

– Example: $A \Rightarrow C$ follows from $A \Rightarrow B$ and $B \Rightarrow C$ (for every $A, B, C \subseteq Y$), i.e. $\{A \Rightarrow B, B \Rightarrow C\} \models A \Rightarrow C$.

– Therefore, we can introduce a deduction rule
  (Tra) from $A \Rightarrow B$ and $B \Rightarrow C$ infer $A \Rightarrow C$.

– We can use such rule to derive new AI such as

  – start from $T = \{\{y_1\} \Rightarrow \{y_2, y_5\}, \{y_2, y_5\} \Rightarrow \{y_3\}, \{y_3\} \Rightarrow \{y_2, y_4\}\}$,
  – apply (Tra) to the first and the second AI in T to infer $\{y_1\} \Rightarrow \{y_3\}$,
  – apply (Tra) to $\{y_1\} \Rightarrow \{y_3\}$ and the second AI in T to infer $\{y_1\} \Rightarrow \{y_2, y_4\}$.

Therefore, the following question arises:

- Is there a collection of simple deduction rules which allow us to determine whether $T \models A \Rightarrow B$?, i.e., rules such that
- 1. if $A \Rightarrow B$ semantically follows from T then one can derive $A \Rightarrow B$ from T using those rules (like above) and
- 2. if one can derive $A \Rightarrow B$ from T then $A \Rightarrow B$ semantically follows from T.

Our system for reasoning about attribute implications consists of the following (schemes of) deduction rules:

(Ax) infer $A \cup B \Rightarrow A$,
(Cut) from $A \Rightarrow B$ and $B \cup C \Rightarrow D$ infer $A \cup C \Rightarrow D$,

for every $A, B, C, D \subseteq Y$.

- (Ax) is a rule without the input part "from …", i.e. $A \cup B \Rightarrow A$ can be inferred from any AIs.
- (Cut) has both the input and the output part.
- Rules for reasoning about AIs go back to Armstrong's research on reasoning about functional dependencies in databases:
  Armstrong W. W.: Dependency structures in data base relationships. IFIP Congress, Geneva, Switzerland, 1974, pp. 580–583.
- There are several systems of deduction rules which are equivalent to (Ax), (Cut), see later.

Example 3.14 (how to use deduction rules)**.** (Cut)
If we have two rules which are of the form $A \Rightarrow B$ and $B \cup C \Rightarrow D$, we can derive (in a single step, using deduction rule (Cut)) a new AI of the form $A \cup C \Rightarrow D$.
Consider AIs $\{r, s\} \Rightarrow \{t, u\}$ and $\{t, u, v\} \Rightarrow \{w\}$.
Putting $A = \{r, s\}$, $B = \{t, u\}$, $C = \{v\}$, $D = \{w\}$, $\{r, s\} \Rightarrow \{t, u\}$ is of the form $A \Rightarrow B$, $\{t, u, v\} \Rightarrow \{w\}$ is of the form $A \cup C \Rightarrow D$,
and we can infer $A \cup C \Rightarrow D$ which is $\{r, s, v\} \Rightarrow \{w\}$.

(Ax)
We can derive (in a single step, using deduction rule (Ax), with no assumptions) a new AI of the form $A \cup B \Rightarrow A$.
For instance, we can infer $\{y_1, y_3, y_4, y_5\} \Rightarrow \{y_3, y_5\}$. Namely, putting $A = \{y_3, y_5\}$ and $B = \{y_1, y_4\}$, $A \cup B \Rightarrow A$ becomes $\{y_1, y_3, y_4, y_5\} \Rightarrow \{y_3, y_5\}$.

How to formalize the concept of a derivation of new AIs using our rules?

Definition 3.15 (proof)**.** A proof of $A \Rightarrow B$ from a set T of AIs is a sequence

$$A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n$$

of AIs satisfying:

1. $A_n \Rightarrow B_n$ is just $A \Rightarrow B$,
2. for every $i = 1, 2, \ldots, n$:
   - either $A_i \Rightarrow B_i$ is from T ("assumption"),
   - or $A_i \Rightarrow B_i$ results by application of (Ax) or (Cut) to some of preceding AIs $A_j \Rightarrow B_j$'s ("deduction").

In such case, we write $T \vdash A \Rightarrow B$ and say that $A \Rightarrow B$ is provable (derivable) from T using (Ax) and (Cut).

- proof as a sequence?: makes sense: informally, we understand a proof to be a sequence of our arguments which we take from 1. assumptions (from T) of 2. infer pro previous arguments by deduction steps.

Example 3.16 (simple proof)**.** Proof of $P \Rightarrow R$ from $T = \{P \Rightarrow Q, Q \Rightarrow R\}$ is a sequence:

$$P \Rightarrow Q, Q \Rightarrow R, P \Rightarrow R$$

because: $P \Rightarrow Q \in T$; $Q \Rightarrow R \in T$; $P \Rightarrow R$ can be inferred from $P \Rightarrow Q$ and $Q \Rightarrow R$ using (Cut). Namely, put $A = P$, $B = Q$, $C = Q$, $D = R$; then $A \Rightarrow B$ becomes $P \Rightarrow Q$, $B \cup C \Rightarrow D$ becomes $Q \Rightarrow R$, and $A \cup C \Rightarrow D$ becomes $P \Rightarrow R$.

Note that this works for any particular sets P, Q, R. For instance for $P = \{y_1, y_3\}$, $Q = \{y_3, y_4, y_5\}$, $R = \{y_2, y_4\}$, or
$P = \{$watches-TV,unhealthy-food$\}$, $Q = \{$high-blood-pressure$\}$, $R = \{$often-visits-doctor$\}$.

In the latter case, we inferred:
$\{$watches-TV,unhealthy-food$\} \Rightarrow \{$often-visits-doctor$\}$ from
$\{$watches-TV,unhealthy-food$\} \Rightarrow \{$high-blood-pressure$\}$ and $\{$high-blood-pressure$\} \Rightarrow \{$often-visits-doctor$\}$.

The notions of a deduction rule and proof are syntactic notions. Proof results by "manipulation of symbols" according to deduction rules. We do not refer to any data table when deriving new AIs using deduction rules.
A typical scenario: (1) We extract a set T of AIs from data table and then (2) infer further AIs from T using deduction rules. In (2), we do not use the data table. Next, we turn to the following two notions:

- Soundness: Is our inference using (Ax) and (Cut) sound? That is, is it the case that IF $T \vdash A \Rightarrow B$ ($A \Rightarrow B$ can be inferred from T) THEN $T \models A \Rightarrow B$ ($A \Rightarrow B$ semantically follows from T, i.e., $A \Rightarrow B$ is true in every table in which all AIs from T are true)?

- Completeness: Is our inference using (Ax) and (Cut) complete? That is, is it the case that IF $T \models A \Rightarrow B$ THEN $T \vdash A \Rightarrow B$?

Definition 3.17 (derivable rule)**.** Deduction rule

from $A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n$ infer $A \Rightarrow B$

is *derivable* from (Ax) and (Cut) if $\{A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n\} \vdash A \Rightarrow B$.

- Derivable rule = new deduction rule = shorthand for a derivation using the basic rules (Ax) and (Cut).

- Why derivable rules: They are natural rules which can speed up proofs.

- Derivable rules can be used in proofs (in addition to the basic rules (Ax) and (Cut)). Why: By definition, a single deduction step using a derivable rule can be replaced by a sequence of deduction steps using the original deduction rules (Ax) and (Cut) only.

Theorem 3.18 (derivable rules)**.** *The following rules are derivable from (Ax) and (Cut):*

(Ref) *infer* $A \Rightarrow A$,

(Wea) *from* $A \Rightarrow B$ *infer* $A \cup C \Rightarrow B$,

(Add) *from* $A \Rightarrow B$ *and* $A \Rightarrow C$ *infer* $A \Rightarrow B \cup C$,

(Pro) *from* $A \Rightarrow B \cup C$ *infer* $A \Rightarrow B$,

(Tra) *from* $A \Rightarrow B$ *and* $B \Rightarrow C$ *infer* $A \Rightarrow C$,

for every $A, B, C, D \subseteq Y$.

Proof. In order to avoid confusion with symbols A, B, C, D used in (Ax) and (Cut), we use P, Q, R, S instead of A, B, C, D in (Ref)–(Tra).

(Ref): We need to show $\{\} \vdash P \Rightarrow P$, i.e. that $P \Rightarrow P$ is derivable using (Ax) and (Cut) from the empty set of assumptions.
Easy, just put $A = P$ and $B = P$ in (Ax). Then $A \cup B \Rightarrow A$ becomes $P \Rightarrow P$. Therefore, $P \Rightarrow P$ can be inferred (in a single step) using (Ax), i.e., a one-element sequence $P \Rightarrow P$ is a proof of $P \Rightarrow P$. This shows $\{\} \vdash P \Rightarrow P$.

(Wea): We need to show $\{P \Rightarrow Q\} \vdash P \cup R \Rightarrow Q$.
A proof (there may be several proofs, this is one of them) is:

$$P \cup R \Rightarrow P, P \Rightarrow Q, P \cup R \Rightarrow Q.$$

Namely, 1. $P \cup R \Rightarrow P$ is derived using (Ax), 2. $P \Rightarrow Q$ is an assumption, $P \cup R \Rightarrow Q$ is derived from $P \cup R \Rightarrow P$ and $P \Rightarrow Q$ using (Cut) (put $A = P \cup R$, $B = P$, $C = P$, $D = Q$).

(Add): EXERCISE.

(Pro): We need to show $\{P \Rightarrow Q \cup R\} \vdash P \Rightarrow Q$.
A proof is:

$$P \Rightarrow Q \cup R, Q \cup R \Rightarrow Q, P \Rightarrow Q.$$

Namely, 1. $P \Rightarrow Q \cup R$ is an assumption, 2. $Q \cup R \Rightarrow Q$ by application of (Ax), 3. $P \Rightarrow Q$ by application of (Cut) to $P \Rightarrow Q \cup R, Q \cup R \Rightarrow Q$ (put $A = P$, $B = C = Q \cup R$, $D = Q$).

(Tra): We need to show $\{P \Rightarrow Q, Q \Rightarrow R\} \vdash P \Rightarrow R$. This was checked earlier. □

- (Ax) …"axiom", and (Cut) …"rule of cut",
- (Ref) …"rule of reflexivity", (Wea) …"rule of weakening", (Add) …"rule of additivity", (Pro) …"rule of projectivity", (Ref) …"rule of transitivity".

Alternative notation for deduction rules: rule "from $A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n$ infer $A \Rightarrow B$" displayed as

$$\frac{A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n}{A \Rightarrow B}.$$

So, (Ax) and (Cut) displayed as

$$\frac{}{A \cup B \Rightarrow A} \text{ and } \frac{A \Rightarrow B, B \cup C \Rightarrow D}{A \cup C \Rightarrow D}.$$

Definition 3.19 (sound deduction rules)**.** Deduction rule "from $A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n$ infer $A \Rightarrow B$" is sound if

$$\{A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n\} \models A \Rightarrow B.$$

- Soundness of a rule: if $A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n$ are true in a data table, then $A \Rightarrow B$ needs to be true in that data table, too.
- Meaning: Sound deduction rules do not allow us to infer "untrue" AIs from true AIs.

Theorem 3.20. *(Ax) and (Cut) are sound.*

Proof. (Ax): We need to check $\{\} \models A \cup B \Rightarrow A$, i.e. that $A \cup B \Rightarrow A$ semantically follows from an empty set T of assumptions. That is, we need to check that $A \cup B \Rightarrow A$ is true in any $M \subseteq Y$ (notice: any $M \subseteq Y$ is a model of the empty set of AIs). This amounts to verifying

$$A \cup B \subseteq M \text{ implies } A \subseteq M,$$

which is evidently true.

(Cut): We need to check $\{A \Rightarrow B, B \cup C \Rightarrow D\} \models A \cup C \Rightarrow D$. Let M be a model of $\{A \Rightarrow B, B \cup C \Rightarrow D\}$. We need to show that M is a model of $A \cup C \Rightarrow D$, i.e. that

$$A \cup D \subseteq M \text{ implies } D \subseteq M.$$

Let thus $A \cup C \subseteq M$. Then $A \subseteq M$, and since we assume M is a model of $A \Rightarrow B$, we need to have $B \subseteq M$. Furthermore, $A \cup C \subseteq M$ yields $C \subseteq M$. That is, we have $B \subseteq M$ and $C \subseteq M$, i.e. $B \cup C \subseteq M$. Now, taking $B \cup C \subseteq M$ and invoking the assumption that M is a model of $B \cup C \Rightarrow D$ gives $D \subseteq M$. □

Corollary 3.21 (soundness of inference using (Ax) and (Cut))**.** *If $T \vdash A \Rightarrow B$ then $T \models A \Rightarrow B$.*

Proof. Direct consequence of previous theorem: Let

$$A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n$$

be a proof from T. It suffices to check that every model M of T is a model of $A_i \Rightarrow B_i$ for $i = 1, \ldots, n$. We check this by induction over i, i.e., we assume that M is a model of $A_j \Rightarrow B_j$'s for $j < i$ and check that M is a model of $A_i \Rightarrow B_i$. There are two options: 1. Either $A_i \Rightarrow B_i$ if from T. Then, trivially, M is a model of $A_i \Rightarrow B_i$ (our assumption). 2. Or, $A_i \Rightarrow B_i$ results by (Ax) or (Cut) to some $A_j \Rightarrow B_j$'s for $j < i$. Then, since we assume that M is a model of $A_j \Rightarrow B_j$'s, we get that M is a model of $A_i \Rightarrow B_i$ by soundness of (Ax) and (Cut). □

Corollary 3.22 (soundness of derived rules)**.** *(Ref), (Wea), (Add), (Pro), (Tra) are sound.*

Proof. As an example, take (Wea). Note that (Wea) is a derived rule. This means that $\{A \Rightarrow B\} \vdash A \cup C \Rightarrow B$. Applying previous corollary yields $\{A \Rightarrow B\} \models A \cup C \Rightarrow B$ which means, by definition, that (Wea) is sound. □

- We have two notions of consequence, semantic and syntactic.
- Semantic: $T \models A \Rightarrow B \ldots A \Rightarrow B$ semantically follows from T.
- Syntactic: $T \vdash A \Rightarrow B \ldots A \Rightarrow B$ syntactically follows from T (is provable from T).
- We know (previous corollary on soundness) that $T \vdash A \Rightarrow B$ implies $T \models A \Rightarrow B$.
- Next, we are going to check completeness, i.e. $T \models A \Rightarrow B$ implies $T \vdash A \Rightarrow B$.

Definition 3.23 (semantic closure, syntactic closure)**.**
- *Semantic closure* of T is the set $$sem(T) = \{A \Rightarrow B \mid T \models A \Rightarrow B\}$$ of all AIs which semantically follow from T.
- *Syntactic closure* of T is the set $$syn(T) = \{A \Rightarrow B \mid T \vdash A \Rightarrow B\}$$ of all AIs which syntactically follow from T (i.e., are provable from T using (Ax) and (Cut)).
- T is *semantically closed* if $T = sem(T)$.
- T is *syntactically closed* if $T = syn(T)$.
- It can be checked that $sem(T)$ is the least set of AIs which is semantically closed and which contains T.
- It can be checked that $syn(T)$ is the least set of AIs which is syntactically closed and which contains T.

Lemma 3.24. *T is syntactically closed iff for any $A, B, C, D \subseteq Y$*

1. *$A \cup B \Rightarrow B \in T$,*
2. *if $A \Rightarrow B \in T$ and $B \cup C \Rightarrow D \in T$ implies $A \cup C \Rightarrow D \in T$.*

Proof. "$\Rightarrow$": If T is syntactically closed then any AI which is provable from T needs to be in T. In particular, $A \cup B \Rightarrow B$ is provable from T, therefore $A \cup B \Rightarrow B \in T$; if $A \Rightarrow B \in T$ and $B \cup C \Rightarrow D \in T$ then, obviously, $A \cup C \Rightarrow D$ is provable from T (by using (Cut)), therefore $A \cup C \Rightarrow D \in T$.

"$\Leftarrow$": If 1. and 2. are satisfied then, obviously, any AI which is provable from T needs to belong to T, i.e. T is syntactically closed. $\square$

This says that T is syntactically closed iff T is closed under deduction rules (Ax) and (Cut).

Lemma 3.25. *If T is semantically closed then T is syntactically closed.*

Proof. Let T be semantically closed. In order to see that T is syntactically closed, it suffices to verify 1. and 2. of previous Lemma.
1.: We have $T \models A \cup B \Rightarrow B$ (we even have $\{\} \models A \cup B \Rightarrow B$). Since T is semantically closed, we get $A \cup B \Rightarrow B \in T$.
2.: Let $A \Rightarrow B \in T$ and $B \cup C \Rightarrow D \in T$. Since $\{A \Rightarrow B, B \cup C \Rightarrow D\} \models A \cup C \Rightarrow D$ (cf. soundness of (Cut)), we have $T \models A \cup C \Rightarrow D$. Now, since T is semantically closed, we get $A \cup C \Rightarrow D \in T$, verifying 2. $\square$

Lemma 3.26. *If T is syntactically closed then T is semantically closed.*

Proof. Let T be syntactically closed. In order to show that T is semantically closed, it suffices to show $sem(T) \subseteq T$. We prove this by showing that if $A \Rightarrow B \notin T$ then $A \Rightarrow B \notin sem(T)$. Recall that since T is syntactically closed, T is closed under all (Ref)–(Tra).

Let thus $A \Rightarrow B \notin T$. To see $A \Rightarrow B \notin sem(T)$, we show that there is $M \in \mathrm{Mod}(T)$ which is not a model of $A \Rightarrow B$. For this purpose, consider $M = A^+$ where A^+ is the largest one such that $A \Rightarrow A^+ \in T$. A^+ exists. Namely, consider all AIs $A \Rightarrow C_1, \ldots, A \Rightarrow C_n \in T$. Note that at least one such AI exists. Namely, $A \Rightarrow A \in T$ by (Ref). Now, repeated application of (Add) yields $A \Rightarrow \bigcup_{i=1}^{n} C_i \in T$ and we have $A^+ = \bigcup_{i=1}^{n} C_i$.

Now, we need to check that (a) $||A \Rightarrow B||_{A^+} = 0$ (i.e., A^+ is not a model of $A \Rightarrow B$) and (b) for every $C \Rightarrow D \in T$ we have $||C \Rightarrow D||_{A^+} = 1$ (i.e., A^+ is a model of T).

(a): We need to show $||A \Rightarrow B||_{A^+} = 0$. By contradiction, suppose $||A \Rightarrow B||_{A^+} = 1$. Since $A \subseteq A^+$, $||A \Rightarrow B||_{A^+} = 1$ yields $B \subseteq A^+$. Since $A \Rightarrow A^+ \in T$, (Pro) would give $A \Rightarrow B \in T$, a contradiction to $A \Rightarrow B \notin T$.

(b): Let $C \Rightarrow D \in T$. We need to show $||C \Rightarrow D||_{A^+} = 1$, i.e. if $C \subseteq A^+$ then $D \subseteq A^+$.
To see this, it is sufficient to verify that if $C \subseteq A^+$ then $A \Rightarrow D \in T$.
Namely, since A^+ is the largest one for which $A \Rightarrow A^+ \in T$, $A \Rightarrow D \in T$ implies $D \subseteq A^+$. So let $C \subseteq A^+$. We have
(b1) $A \Rightarrow A^+ \in T$ (by definition of A^+),
(b2) $A^+ \Rightarrow C \in T$ (this follows by (Pro) from $C \subseteq A^+$),
(b3) $C \Rightarrow D \in T$ (our assumption).
Therefore, applying (Tra) to (b1), (b2), (b3) twice gives $A \Rightarrow D \in T$. $\square$

Theorem 3.27 (soundness and completeness)**.** *$T \vdash A \Rightarrow B$ iff $T \models A \Rightarrow B$.*

Proof. Clearly, it suffices to check $syn(T) = sem(T)$. Recall: $A \Rightarrow B \in syn(T)$ means $T \vdash A \Rightarrow B$, $A \Rightarrow B \in sem(T)$ means $T \models A \Rightarrow B$.

"$sem(T) \subseteq syn(T)$": Since $syn(T)$ is syntactically closed, it is also semantically closed (previous lemma). Therefore, $sem(syn(T)) = syn(T)$ (semantic closure of

$syn(T)$ is just $syn(T)$ because $syn(T)$ is semantically closed). Furthermore, since $T \subseteq syn(T)$, we have $sem(T) \subseteq sem(syn(T))$. Putting this together gives
$$sem(T) \subseteq sem(syn(T)) = syn(T).$$

"$syn(T) \subseteq sem(T)$": Since $sem(T)$ is semantically closed, it is also syntactically closed (previous lemma). Therefore, $syn(sem(T)) = sem(T)$. Furthermore, since $T \subseteq sem(T)$, we have $syn(T) \subseteq syn(sem(T))$. Putting this together gives
$$syn(T) \subseteq syn(sem(T)) = sem(T).$$

□

Summary:

- (Ax) and (Cut) are elementary deduction rules.
- Proof …formalizes derivation process of new AIs from other AIs.
- We have two notions of consequence:
  - $T \models A \Rightarrow B$ …semantic consequence ($A \Rightarrow B$ is true in every model of T).
  - $T \vdash A \Rightarrow B$ …syntactic consequence ($A \Rightarrow B$ is provable T, i.e. can be derived from T using deduction rules).
- Note: proof = syntactic manipulation, no reference to semantic notions; in order to know what $T \vdash A \Rightarrow B$ means, we do not have to know what it means that an AI $A \Rightarrow B$ is true in M.
- Derivable rules (Ref)–(Tra) …derivable rule = shorthand, inference of new AIs using derivable rules can be replaced by inference using original rules (Ax) and (Cut).
- Sound rule …derives true conclusions from true premises; (Ax) and (Cut) are sound; in detail, for (Cut): soundness of (Cut) means that for every M in which both $A \Rightarrow B$ and $B \cup C \Rightarrow D$ are true, $A \cup C \Rightarrow D$ needs to be true, too.
- Soundness of inference using sound rules: if $T \vdash A \Rightarrow B$ ($A \Rightarrow B$ is provable from T) then $T \models A \Rightarrow B$ ($A \Rightarrow B$ semantically follows from T), i.e. if $A \Rightarrow B$ is provable from T then $A \Rightarrow B$ is true in every M in which every AI from T is true. Therefore, soundness of inference means that if we take an arbitrary M and take a set T of AIs which are true in M, then evey AI $A \Rightarrow B$ which we can infer (prove) from T using our inference rules needs to be true in M.
- Consequence: rules, such as (Ref)–(Tra), which can be derived from sound rules are sound.
- $sem(T)$ …set of all AIs which are semantic consequences of T,
  $syn(T)$ …set of all AIs which are syntactic consequences of T (provable from T).
- T is syntactically closed iff T is closed under (Ax) and (Cut).
- (Syntactico-semantical) completeness of rules (Ax) and (Cut): $T \vdash A \Rightarrow B$ iff $T \models A \Rightarrow B$.

Example 3.28.

- Explain why $\{\} \models A \Rightarrow B$ means that (1) $A \Rightarrow B$ is true in every $M \subseteq Y$, (2) $A \Rightarrow B$ is true in every formal context $\langle X, Y, I \rangle$.
- Explain why soundness of inference implies that if we take an arbitrary formal context $\langle X, Y, I \rangle$ and take a set T of AIs which are true in $\langle X, Y, I \rangle$, then evey AI $A \Rightarrow B$ which we can infer (prove) from T using our inference rules needs to be true in $\langle X, Y, I \rangle$.
- Let $\mathcal{R}_1$ and $\mathcal{R}_2$ be two sets of deduction rules, e.g. $\mathcal{R}_1 = \{$(Ax), (Cut)$\}$. Call $\mathcal{R}_1$ and $\mathcal{R}_2$ equivalent if every rule from $\mathcal{R}_2$ is a derived rule in terms of rules from $\mathcal{R}_1$ and, vice versa, every rule from $\mathcal{R}_1$ is a derived rule in terms of rules from $\mathcal{R}_2$.
  For instance, we know that taking $\mathcal{R}_1 = \{$(Ax), (Cut)$\}$, every rule from $\mathcal{R}_2 = \{$(Ref),…, (Tra)$\}$ is a derived rule in terms of rules of $\mathcal{R}_1$.
  Verify that $\mathcal{R}_1 = \{$(Ax), (Cut)$\}$ and $\mathcal{R}_2 = \{$(Ref),(Wea), (Cut)$\}$ are equivalent.

Example 3.29. – Explain: If $\mathcal{R}_1$ and $\mathcal{R}_2$ are equivalent sets of inference rules then $A \Rightarrow B$ is provable from T using rules from $\mathcal{R}_1$ iff $A \Rightarrow B$ is provable from T using rules from $\mathcal{R}_2$.

- Explain: Let $\mathcal{R}_2$ be a set of inference rules equivalent to $\mathcal{R}_1 = \{(\text{Ax}), (\text{Cut})\}$. Then $A \Rightarrow B$ is provable from T using rules from $\mathcal{R}_2$ iff $T \models A \Rightarrow B$.
- Verify that $sem(\cdots)$ is a closure operator, i.e. that $T \subseteq sem(T)$, $T_1 \subseteq T_2$ implies $sem(T_1) \subseteq sem(T_2)$, and $sem(T) = sem(sem(T))$.
- Verify that $syn(\cdots)$ is a closure operator, i.e. that $T \subseteq syn(T)$, $T_1 \subseteq T_2$ implies $syn(T_1) \subseteq syn(T_2)$, and $syn(T) = syn(syn(T))$.
- Verify that for any T, $sem(T)$ is the least semantically closed set which contains T.
- Verify that for any T, $syn(T)$ is the least syntactically closed set which contains T.

3.3 Models of Attribute Implications

For a set T of attribute implications, denote

$$\mathrm{Mod}(T) = \{M \subseteq Y \mid ||A \Rightarrow B||_M = 1 \text{ for every } A \Rightarrow B \in T\}$$

That is, $\mathrm{Mod}(T)$ is the set of all models of T.

Definition 3.30 (closure system). A closure system in a set Y is any system $\mathcal{S}$ of subsets of Y which contains Y and is closed under arbitrary intersections.

That is, $Y \in \mathcal{S}$ and $\bigcap \mathcal{R} \in \mathcal{S}$ for every $\mathcal{R} \subseteq \mathcal{S}$ (intersection of every subsystem $\mathcal{R}$ of $\mathcal{S}$ belongs to $\mathcal{S}$).

$\{\{a\}, \{a, b\}, \{a, d\}, \{a, b, c, d\}\}$ is a closure system in $\{a, b, c, d\}$ while $\{\{a, b\}, \{c, d\}, \{a, b, c, d\}\}$ is not.

There is a one-to-one relationship between closure systems in Y and closure operators in Y. Given a closure operator C in Y, $\mathcal{S}_C = \{A \in 2^X \mid A = C(A)\} = \mathrm{fix}(C)$ is a closure system in Y.

Given a closure system in Y, putting

$$C_{\mathcal{S}}(A) = \bigcap \{B \in \mathcal{S} \mid A \subseteq B\}$$

for any $A \subseteq X$, $C_{\mathcal{S}}$ is a closure operator on Y. This is a one-to-one relationship, i.e. $C = C_{\mathcal{S}_C}$ and $\mathcal{S} = \mathcal{S}_{C_{\mathcal{S}}}$ (we omit proofs).

Lemma 3.31. *For a set T of attribute implications, $\mathrm{Mod}(T)$ is a closure system in Y.*

Proof. First, $Y \in \mathrm{Mod}(T)$ because Y is a model of any attribute implication.

Second, let $M_j \in \mathrm{Mod}(T)$ $(j \in J)$. For any $A \Rightarrow B \in T$, if $A \subseteq \bigcap_j M_j$ then for each $j \in J$: $A \subseteq M_j$, and so $B \subseteq M_j$ (since $M_j \in \mathrm{Mod}(T)$, thus in particular $M_j \models A \Rightarrow B$), from which we have $B \subseteq \bigcap_j M_j$.

We showed that $\mathrm{Mod}(T)$ contains Y and is closed under intersections, i.e. $\mathrm{Mod}(T)$ is a closure system. □

Remark 3.32. (1) If T is the set of all attribute implications valid in a formal context $\langle X, Y, I\rangle$, then $\mathrm{Mod}(T) = \mathrm{Int}(X, Y, I)$, i.e. models of T are just all the intents of the concept lattice $\mathcal{B}(X, Y, I)$ (see later).

(2) Another connection to concept lattices is: $A \Rightarrow B$ is valid in $\langle X, Y, I\rangle$ iff $A^{\downarrow} \subseteq B^{\downarrow}$ iff $B \subseteq A^{\downarrow\uparrow}$ (see later).

Since $\mathrm{Mod}(T)$ is a closure system, we can consider the corresponing closure operator $C_{\mathrm{Mod}(T)}$ (i.e., the fixed points of $C_{\mathrm{Mod}(T)}$ are just models of T). Therefore, for every $A \subseteq Y$ there exist the least model of $\mathrm{Mod}(T)$ which contains A, namely, such least model is just $C_{\mathrm{Mod}(T)}(A)$.

Theorem 3.33 (testing entailment via least model)**.** *For any $A \Rightarrow B$ and any T, we have*

$$T \models A \Rightarrow B \quad \textit{iff} \quad ||A \Rightarrow B||_{C_{\mathrm{Mod}(T)}(A)} = 1,$$

i.e., $A \Rightarrow B$ semantically follows from T iff $A \Rightarrow B$ is true in the least model $C_{\mathrm{Mod}(T)}(A)$ of T which contains A.

Proof. "$\Rightarrow$": If $T \models A \Rightarrow B$ then, by definition, $A \Rightarrow B$ is true in every model of T. Therefore, in particular, $A \Rightarrow B$ is true in $C_{\mathrm{Mod}(T)}(A)$.

"$\Leftarrow$": Let $A \Rightarrow B$ be true in $C_{\mathrm{Mod}(T)}(A)$. Since $A \subseteq C_{\mathrm{Mod}(T)}(A)$, we have $B \subseteq C_{\mathrm{Mod}(T)}(A)$. We need to check that $A \Rightarrow B$ is true in every model of T. Let thus $M \in \mathrm{Mod}(T)$. If $A \nsubseteq M$ then, clearly, $A \Rightarrow B$ is true in M. If $A \subseteq M$ then, since M is a model of T containing A, we have $C_{\mathrm{Mod}(T)}(A) \subseteq M$. Putting together with $B \subseteq C_{\mathrm{Mod}(T)}(A)$, we get $B \subseteq M$, i.e. $A \Rightarrow B$ is true in M. ☐

– Previous theorem $\Rightarrow$ testing $T \models A \Rightarrow B$ by checking whether $A \Rightarrow B$ is true in a single particular model of T. This is much better than going by definition $\models$ (definition says: $T \models A \Rightarrow B$ iff $A \Rightarrow B$ is true in every model of T).
– How can we obtain $C_{\mathrm{Mod}(T)}(A)$?

Definition 3.34. For $Z \subseteq Y$, T a set of implications, put

1. $Z^T = Z \cup \bigcup\{B \mid A \Rightarrow B \in T, A \subseteq Z\}$,
2. $Z^{T_0} = Z$,
3. $Z^{T_n} = (Z^{T_{n-1}})^T$ (for $n \geq 1$).

Define define operator $C : 2^Y \to 2^Y$ by

$C(Z) = \bigcup_{n=0}^{\infty} Z^{T_n}$

Theorem 3.35. *Given T, C (defined on previous slide) is a closure operator in Y such that* $C(Z) = C_{\mathrm{Mod}(T)}(Z)$.

Proof. First, check that C is a closure operator.
$Z = Z^{T_0}$ yields $Z \subseteq C(Z)$.
Evidently, $Z_1 \subseteq Z_2$ implies $Z_1^T \subseteq Z_2^T$ which implies $Z_1^{T_1} \subseteq Z_2^{T_1}$ which implies $Z_1^{T_2} \subseteq Z_2^{T_2}$ which implies $\ldots Z_1^{T_n} \subseteq Z_2^{T_n}$ for any n. That is, $Z_1 \subseteq Z_2$ implies $C(Z_1) = \bigcup_{n=0}^{\infty} Z_1^{T_n} \subseteq \bigcup_{n=0}^{\infty} Z_2^{T_n} = C(Z_2)$.
$C(Z) = C(C(Z))$: Clearly, $Z^{T_0} \subseteq Z^{T_1} \subseteq \cdots Z^{T_n} \subseteq \cdots$.
Since Y is finite, the above sequence terminates after a finite number n_0 of steps, i.e. there is n_0 such that $C(Z) = \bigcup_{n=0}^{\infty} Z^{T_n} = Z^{T_{n_0}}$.
This means $(Z^{T_{n_0}})^T = Z^{T_{n_0}} = C(Z)$ which gives $C(Z) = C(C(Z))$.

Next, we check $C(Z) = C_{\mathrm{Mod}(T)}(Z)$.
1. $C(Z)$ is a model of T containing Z:
Above, we checked that$C(Z)$ contains Z. Take any $A \Rightarrow B \in T$ and verify that $A \Rightarrow B$ is valid in $C(Z)$ (i.e., $C(Z)$ is a model of $A \Rightarrow B$). Let $A \subseteq C(Z)$. We need to check $B \subseteq C(Z)$. $A \subseteq C(Z)$ means that for some n, $A \subseteq Z^{T_n}$. But then, by definition, $B \subseteq (Z^{T_n})^T$ which gives $B \subseteq Z^{T_{n+1}} \subseteq C(Z)$.

2. $C(Z)$ is the least model of T containing Z:
Let M be a model of T containing Z, i.e. $Z^{T_0} = Z \subseteq M$. Then $Z^T \subseteq M^T$ (just check

definition of $(\cdots)^T$). Evidently, $M = M^T$. Therefore, $Z^{T_1} = Z^T \subseteq M$. Applying this inductively gives $Z^{T_2} \subseteq M$, $Z^{T_3} \subseteq M$, ... Putting together yields $C(Z) = \bigcup_{n=0}^{\infty} Z^{T_n} \subseteq M$. That is, $C(Z)$ is contained in every model M of T and is thus the least one. $\square$

- Therefore, C is the closure operator which computes, given $Z \subseteq Y$, the least model of T containing Z.
- As argued in the proof, since Y is finite, $\bigcup_{n=0}^{\infty} Z^{T_n}$ "stops" after a finite number of steps. Namely, there is n_0 such that $Z^{T_n} = Z^{T_{n_0}}$ for $n > n_0$.
- The least such n_0 is the smallest n with $Z^{T_n} = Z^{T_{n+1}}$.
- Given T, $C(Z)$ can be computed: Use definition and stop whenever $Z^{T_n} = Z^{T_{n+1}}$. That is, put
$$C(Z) = Z \cup Z^{T_1} \cup Z^{T_2} \cup \cdots \cup Z^{T_n}.$$
- There is a more efficient algorithm (called LinClosure) for computing $C(Z)$. See Maier D.: The Theory of Relational Databases. CS Press, 1983.

Example 3.36. Back to one of our previous examples: Let $Y = \{y_1, y_2, y_3\}$. Determine whether $T \models A \Rightarrow B$.

- $T = \{\{y_3\} \Rightarrow \{y_1, y_2\}, \{y_1, y_3\} \Rightarrow \{y_2\}\}$, $A \Rightarrow B$ is $\{y_2, y_3\} \Rightarrow \{y_1\}$.
  1. $\mathrm{Mod}(T) = \{\emptyset, \{y_1\}, \{y_2\}, \{y_1, y_2\}, \{y_1, y_2, y_3\}\}$.
  2. By definition: $||\{y_2, y_3\} \Rightarrow \{y_1\}||_{\emptyset} = 1$, $||\{y_2, y_3\} \Rightarrow \{y_1\}||_{\{y_1\}} = 1$, $||\{y_2, y_3\} \Rightarrow \{y_1\}||_{\{y_2\}} = 1$, $||\{y_2, y_3\} \Rightarrow \{y_1\}||_{\{y_1,y_2\}} = 1$, $||\{y_2, y_3\} \Rightarrow \{y_1\}||_{\{y_1,y_2,y_3\}} = 1$.
  Therefore, $T \models A \Rightarrow B$.
  3. Now, using our theorem: The least model of T containing $A = \{y_2, y_3\}$ is $C_{\mathrm{Mod}(T)}(A) = \{y_1, y_2, y_3\}$. Therefore, to verify $T \models A \Rightarrow B$, we just need to check whether $A \Rightarrow B$ is true in $\{y_1, y_2, y_3\}$. Since $||\{y_2, y_3\} \Rightarrow \{y_1\}||_{\{y_1,y_2,y_3\}} = 1$, we conclude $T \models A \Rightarrow B$.

- $T = \{\{y_3\} \Rightarrow \{y_1, y_2\}, \{y_1, y_3\} \Rightarrow \{y_2\}\}$, $A \Rightarrow B$ is $\{y_2\} \Rightarrow \{y_1\}$.
  1. $\mathrm{Mod}(T) = \{\emptyset, \{y_1\}, \{y_2\}, \{y_1, y_2\}, \{y_1, y_2, y_3\}\}$.
  2. By definition: $||\{y_2\} \Rightarrow \{y_1\}||_{\emptyset} = 1$, $||\{y_2\} \Rightarrow \{y_1\}||_{\{y_1\}} = 1$, $||\{y_2\} \Rightarrow \{y_1\}||_{\{y_2\}} = 0$, we can stop.
  Therefore, $T \not\models A \Rightarrow B$.
  3. Now, using our theorem: The least model of T containing $A = \{y_2\}$ is $C_{\mathrm{Mod}(T)}(A) = \{y_2\}$. Therefore, to verify $T \models A \Rightarrow B$, we need to check whether $A \Rightarrow B$ is true in $\{y_2\}$. Since $||\{y_2\} \Rightarrow \{y_1\}||_{\{y_2\}} = 0$, we conclude $T \not\models A \Rightarrow B$.

Example 3.37. Let $Y = \{y_1, \ldots, y_{10}\}$, $T = \{\{y_1, y_4\} \Rightarrow \{y_3\}, \{y_2, y_4\} \Rightarrow \{y_1\}, \{y_1, y_2\} \Rightarrow \{y_4, y_7\}, \{y_2, y_7\} \Rightarrow \{y_3\}, \{y_6\} \Rightarrow \{y_4\}, \{y_2, y_8\} \Rightarrow \{y_3\}, \{y_9\} \Rightarrow \{y_1, y_2, y_7\}\}$

1. Decide whether $T \models A \Rightarrow B$ for $A \Rightarrow B$ being $\{y_2, y_5, y_6\} \Rightarrow \{y_3, y_7\}$.

We need to check whether $||A \Rightarrow B||_{C_{\mathrm{Mod}(T)}(A)} = 1$. First, we compute $C_{\mathrm{Mod}(T)}(A) = \bigcup_{n=0}^{\infty} A^{T_n}$. Recall:
$$A^{T_n} = A^{T_{n-1}} \cup \bigcup\{D \mid C \Rightarrow D \in T, C \subseteq A^{T_n}\}.$$

- $A^{T_0} = A = \{y_2, y_5, y_6\}$.
- $A^{T_1} = A^{T_0} \cup \bigcup\{\{y_4\}\} = \{y_2, y_4, y_5, y_6\}$.
  Note: $\{y_4\}$ added because for $C \Rightarrow D$ being $\{y_6\} \Rightarrow \{y_4\}$ we have $\{y_6\} \subseteq A^{T_0}$.
- $A^{T_2} = A^{T_1} \cup \bigcup\{\{y_1\}, \{y_4\}\} = \{y_1, y_2, y_4, y_5, y_6\}$.
- $A^{T_3} = A^{T_2} \cup \bigcup\{\{y_3\}, \{y_1\}, \{y_4\}\} = \{y_1, y_2, y_3, y_4, y_5, y_6\}$.
- $A^{T_4} = A^{T_3} \cup \bigcup\{\{y_3\}, \{y_1\}, \{y_4, y_7\}, \{y_4\}\} = \{y_1, y_2, y_3, y_4, y_5, y_6, y_7\}$.

– $A^{T_5} = A^{T_4} \cup \bigcup\{\{y_3\}, \{y_1\}, \{y_4, y_7\}, \{y_4\}\} = \{y_1, y_2, y_3, y_4, y_5, y_6, y_7\} = A^{T_4}$, STOP.

Therefore, $C_{\mathrm{Mod}(T)}(A) = \{y_1, y_2, y_3, y_4, y_5, y_6, y_7\}$. Now, we need to check if $||A \Rightarrow B||_{C_{\mathrm{Mod}(T)}(A)} = 1$, i.e. if

$$||\{y_2, y_5, y_6\} \Rightarrow \{y_3, y_7\}||_{\{y_1,y_2,y_3,y_4,y_5,y_6,y_7\}} = 1.$$

Since this is true, we conclude $T \models A \Rightarrow B$.

2. Decide whether $T \models A \Rightarrow B$ for $A \Rightarrow B$ being $\{y_1, y_2, y_8\} \Rightarrow \{y_4, y_7\}$.

We need to check whether $||A \Rightarrow B||_{C_{\mathrm{Mod}(T)}(A)} = 1$. First, we compute $C_{\mathrm{Mod}(T)}(A) = \bigcup_{n=0}^{\infty} A^{T_n}$.

- $A^{T_0} = A = \{y_1, y_2, y_8\}$.
- $A^{T_1} = A^{T_0} \cup \bigcup\{\{y_3\}\} = \{y_1, y_2, y_3, y_8\}$.
- $A^{T_2} = A^{T_1} \cup \bigcup\{\{y_7\}, \{y_3\}\} = \{y_1, y_2, y_3, y_7, y_8\}$.
- $A^{T_3} = A^{T_2} \cup \bigcup\{\{y_7\}, \{y_3\}\} = \{y_1, y_2, y_3, y_7, y_8\} = A^{T_2}$, STOP.

Thus, $C_{\mathrm{Mod}(T)}(A) = \{y_1, y_2, y_3, y_7, y_8\}$. Now, we need to check if $||A \Rightarrow B||_{C_{\mathrm{Mod}(T)}(A)} = 1$, i.e. if $||\{y_1, y_2, y_8\} \Rightarrow \{y_4, y_7\}||_{\{y_1,y_2,y_3,y_7,y_8\}} = 1$. Since this is not true, we conclude $T \not\models A \Rightarrow B$.

3.4 Non-Redundant Bases of Attribute Implications

Definition 3.38 (non-redundant set of AIs). A set T of attribute implications is *called non-redundant* if for any $A \Rightarrow B \in T$ we have

$$T - \{A \Rightarrow B\} \not\models A \Rightarrow B.$$

That is, if T' results from T be removing an arbitrary $A \Rightarrow B$ from T, then $A \Rightarrow B$ does not semantically follow from T', i.e. T' is weaker than T.

How to check if T is redundant or not? Pseudo-code:

```
for A ⇒ B ∈ T do
    T' := T − {A ⇒ B};
    if T' |= A ⇒ B then
        output(``REDUNDANT'');
        stop;
    endif;
endfor;
output(``NONREDUNDANT'').
```

- Checking $T' \models A \Rightarrow B$: as described above, i.e. test whether $||A \Rightarrow B||_{C_{\mathrm{Mod}(T')}(A)} = 1$.
- Modification of this algorithm gives an algorithm which, given T, returns a non-redundant subset nrT of T which is equally strong as T, i.e. for any $C \Rightarrow D$,
$$T \models C \Rightarrow D \quad \text{iff} \quad nrT \models C \Rightarrow D.$$
Pseudo-code:

```
nrT := T;
for A ⇒ B ∈ nrT do
    T' := nrT − {A ⇒ B};
    if T' |= A ⇒ B then
```

```
            nrT := T';
        endif;
    endfor;
    output(nrT).
```

Definition 3.39 (complete set of AIs)**.** Let $\langle X, Y, I\rangle$ be a formal context, T be a set of attribute implications over Y. T is called *complete* in $\langle X, Y, I\rangle$ if for any attribute implication $C \Rightarrow D$ we have

$$C \Rightarrow D \text{ is true in } \langle X, Y, I\rangle \quad \text{IFF} \quad T \models C \Rightarrow D.$$

– This is a different notion of completeness (different from syntactico-semantical completeness of system (Ax) and (Cut) of Armstrong rules).
– Meaning: T is complete iff validity of any AI $C \Rightarrow D$ in data $\langle X, Y, I\rangle$ is encoded in T via entailment: $C \Rightarrow D$ is true in $\langle X, Y, I\rangle$ iff $C \Rightarrow D$ follows from T. That is, T gives complete information about which AIs are true in data.
– Definition directly yields: If T is complete in $\langle X, Y, I\rangle$ then every $A \Rightarrow B$ from T is true in $\langle X, Y, I\rangle$. Why: because $T \models A \Rightarrow B$ for every $A \Rightarrow B$ from T.

Theorem 3.40 (criterion for T being complete in $\langle X, Y, I\rangle$)**.** *T is complete in $\langle X, Y, I\rangle$ iff* $\mathrm{Mod}(T) = \mathrm{Int}(X, Y, I)$, *i.e. models of T are just intents of formal concepts from $\mathcal{B}(X, Y, I)$.*

Proof. Omitted. □

Definition 3.41 (non-redundant basis of $\langle X, Y, I\rangle$)**.** Let $\langle X, Y, I\rangle$ be a formal context. A set T of attribute implications over Y is called a *non-redundant basis* of $\langle X, Y, I\rangle$ iff

1. T is complete in $\langle X, Y, I\rangle$,
2. T is non-redundant.

– Another way to say that T is a non-redundant basis of $\langle X, Y, I\rangle$:
(a) every AI from T is true in $\langle X, Y, I\rangle$;
(b) for any other AI $C \Rightarrow D$: $C \Rightarrow D$ is true in $\langle X, Y, I\rangle$ iff $C \Rightarrow D$ follows from T;
(c) no proper subset $T' \subseteq T$ satisfies (a) and (b).

Example 3.42 (testing non-redundancy of T)**.** Let $Y = \{ab2, ab6, abs, ac, cru, ebd\}$ with the following meaning of attributes: $ab2$ …has 2 or more airbags, $ab6$ …has 6 or more airbags, abs …has ABS, ac …has air-conditioning, ebd …has EBD.
Let T consist of the following attribute implications: $\{ab6\} \Rightarrow \{abs, ac\}$, $\{\} \Rightarrow \{ab2\}$, $\{ebd\} \Rightarrow \{ab6, cru\}$, $\{ab6\} \Rightarrow \{ab2\}$.
Determine whether T is non-redundant.
We can use the above algorithm, and proceed as follows: We go over all $A \Rightarrow B$ from T and test whether $T' \models A \Rightarrow B$ where $T' = T - \{A \Rightarrow B\}$.

– $A \Rightarrow B = \{ab6\} \Rightarrow \{abs, ac\}$. Then, $T' = \{\{\} \Rightarrow \{ab2\}, \{ebd\} \Rightarrow \{ab6, cru\}, \{ab6\} \Rightarrow \{ab2\}\}$. In order to decide whether $T' \models \{ab6\} \Rightarrow \{abs, ac\}$, we need to compute $C_{\mathrm{Mod}(T')}(\{ab6\})$ and check $||\{ab6\} \Rightarrow \{abs, ac\}||_{C_{\mathrm{Mod}(T')}(\{ab6\})}$. Putting $Z = \{ab6\}$, and denoting $Z^{T'_i}$ by Z^i,

we get $Z^0 = \{ab6\}$, $Z^1 = \{ab2, ab6\}$, $Z^2 = \{ab2, ab6\}$, we can stop and we have $C_{\mathrm{Mod}(T')}(\{ab6\}) = \bigcup_{i=0^1} Z^i = \{ab2, ab6\}$. Now, $||\{ab6\} \Rightarrow \{abs, ac\}||_{C_{\mathrm{Mod}(T')}(\{ab6\})} = ||\{ab6\} \Rightarrow \{abs, ac\}||_{\{ab2, ab6\}} = 0$, i.e. $T' \not\models \{ab6\} \Rightarrow \{abs, ac\}$. That is, we need to go further.
– $A \Rightarrow B = \{\} \Rightarrow \{ab2\}$. Then, $T' = \{\{ab6\} \Rightarrow \{abs, ac\}, \{ebd\} \Rightarrow \{ab6, cru\}, \{ab6\} \Rightarrow \{ab2\}\}$. In order to decide whether $T' \models \{\} \Rightarrow \{ab2\}$, we need to compute $C_{\mathrm{Mod}(T')}(\{\})$ and check $||\{\} \Rightarrow \{ab2\}||_{C_{\mathrm{Mod}(T')}(\{\})}$. Putting

$Z = \{\}$, and denoting $Z^{T'_i}$ by Z^i, we get $Z^0 = \{\}$, $Z^1 = \{\}$ (because there is no $A \Rightarrow B \in T'$ such that $A \subseteq \{\}$), we can stop and we have $C_{\mathrm{Mod}(T')}(\{\}) = Z^0 = \{\}$. Now, $||\{\} \Rightarrow \{ab2\}||_{C_{\mathrm{Mod}(T')}(\{\})} = ||\{\} \Rightarrow \{ab2\}||_{\{\}} = 0$, i.e. $T' \not\models \{\} \Rightarrow \{ab2\}$. That is, we need to go further.

- $A \Rightarrow B = \{ebd\} \Rightarrow \{ab6, cru\}$. Then, $T' = \{\{ab6\} \Rightarrow \{abs, ac\}, \{\} \Rightarrow \{ab2\}, \{ab6\} \Rightarrow \{ab2\}\}$. In order to decide whether $T' \models \{ebd\} \Rightarrow \{ab6, cru\}$, we need to compute

  $C_{\mathrm{Mod}(T')}(\{ebd\})$ and check $||\{ebd\} \Rightarrow \{ab6, cru\}||_{C_{\mathrm{Mod}(T')}(\{ebd\})}$. Putting $Z = \{ebd\}$, and denoting $Z^{T'_i}$ by Z^i, we get $Z^0 = \{ebd\}$, $Z^1 = \{ab2, ebd\}$, $Z^2 = \{ab2, ebd\}$, we can stop and we have $C_{\mathrm{Mod}(T')}(\{ebd\}) = Z^0 = \{ab2, ebd\}$. Now, $||\{ebd\} \Rightarrow \{ab6, cru\}||_{C_{\mathrm{Mod}(T')}(\{ab2,ebd\})} = ||\{ebd\} \Rightarrow \{ab6, cru\}||_{\{ab2,ebd\}} = 0$, i.e. $T' \not\models \{ebd\} \Rightarrow \{ab6, cru\}$. That is, we need to go further.
- $A \Rightarrow B = \{ab6\} \Rightarrow \{ab2\}$. Then, $T' = \{\{ab6\} \Rightarrow \{abs, ac\}, \{\} \Rightarrow \{ab2\}, \{ebd\} \Rightarrow \{ab6, cru\}\}$. In order to decide whether $T' \models \{ab6\} \Rightarrow \{ab2\}$, we need to compute $C_{\mathrm{Mod}(T')}(\{ab6\})$ and check $||\{ab6\} \Rightarrow \{ab2\}||_{C_{\mathrm{Mod}(T')}(\{ab6\})}$. Putting $Z = \{ab6\}$, and denoting $Z^{T'_i}$ by Z^i, we get $Z^0 = \{ab6\}$, $Z^1 = \{ab2, ab6, abs, ac\}$, $Z^2 = \{ab2, ab6, abs, ac\}$, we can stop and we have $C_{\mathrm{Mod}(T')}(\{ab6\}) = \bigcup_{i=0^1} Z^i = \{ab2, ab6, abs, ac\}$. Now, $||\{ab6\} \Rightarrow \{ab2\}||_{C_{\mathrm{Mod}(T')}(\{ab6\})} = ||\{ab6\} \Rightarrow \{ab2\}||_{\{ab2,ab6,abs,ac\}} = 1$, i.e. $T' \models \{ab6\} \Rightarrow \{ab2\}$. Therefore, T is redundant (we can remove $\{ab6\} \Rightarrow \{ab2\}$).

We can see that T is redundant by observing that $T' \vdash \{ab6\} \Rightarrow \{ab2\}$ where $T' = T - \{\{ab6\} \Rightarrow \{ab2\}\}$. Namely, we can infer $\{ab6\} \Rightarrow \{ab2\}$ from $\{\} \Rightarrow \{ab2\}$ by (Wea). Syntactico-semantical completeness yields $T' \models \{ab6\} \Rightarrow \{ab2\}$, hence T is redundant.

Example 3.43 (deciding whether T is complete w.r.t $\langle X, Y, I\rangle$). Consider attributes normal blood pressure (nbp), high blood pressure (hbp), watches TV (TV), eats unhealthy food (uf), runs regularly (r), persons $a, \ldots, e$, and formal context (table) $\langle X, Y, I\rangle$

I	nbp	hbp	TV	uf	r
a	×				×
b	×			×	×
c		×	×	×	
d		×		×	
e	×				

Decide whether T is complete w.r.t. $\langle X, Y, I\rangle$ for sets T described below.

Due to the above theorem, we need to check $\mathrm{Mod}(T) = \mathrm{Int}(X, Y, I)$. That is, we need to compute $\mathrm{Int}(X, Y, I)$ and $\mathrm{Mod}(T)$ and compare.

We have $\mathrm{Int}(X, Y, I) = \{\{\}, \{nbp\}, \{uf\}, \{uf, hbp\}, \{nbp, r\}, \{uf, hbp, TV\}, \{nbp, r, uf\}, \{hbp, nbp, r, TV, uf\}\}$

1. T consists of $\{r\} \Rightarrow \{nbp\}$, $\{TV, uf\} \Rightarrow \{hbp\}$, $\{r, uf\} \Rightarrow \{TV\}$.
   T is not complete w.r.t. $\langle X, Y, I\rangle$ because $\{r, uf\} \Rightarrow \{TV\}$ is not true in $\langle X, Y, I\rangle$ (person b is a counterexample). Recall that if T is complete, every AI from T is true in $\langle X, Y, I\rangle$.
2. T consists of $\{r\} \Rightarrow \{nbp\}$, $\{TV, uf\} \Rightarrow \{hbp\}$, $\{TV\} \Rightarrow \{hbp\}$.
   In this case, every AI from T is true in $\langle X, Y, I\rangle$. But still, T is not complete. Namely, $\mathrm{Mod}(T) \not\subseteq \mathrm{Int}(X, Y, I)$. For instance, $\{hbp, TV\} \in \mathrm{Mod}(T)$ but $\{hbp, TV\} \notin \mathrm{Int}(X, Y, I)$.

   In this case, T is too weak. T does not entail all attribute implications which are true in $\langle X, Y, I\rangle$. For instance $\{hbp, TV\} \Rightarrow \{uf\}$ is true in $\langle X, Y, I\rangle$ but $T \not\models \{hbp, TV\} \Rightarrow \{uf\}$. Indeed, $\{hbp, TV\}$ is a model of T but $||\{hbp, TV\} \Rightarrow \{uf\}||_{\{hbp,TV\}} = 0$.

3. T consists of $\{r\} \Rightarrow \{nbp\}$, $\{TV, uf\} \Rightarrow \{hbp\}$, $\{TV\} \Rightarrow \{uf\}$, $\{TV\} \Rightarrow \{hbp\}$, $\{hbp, TV\} \Rightarrow \{uf\}$, $\{nbp, uf\} \Rightarrow \{r\}$, $\{hbp\} \Rightarrow \{uf\}$, $\{uf, r\} \Rightarrow \{nbp\}$, $\{nbp, TV\} \Rightarrow \{r\}$, $\{hbp, nbp\} \Rightarrow \{r, TV\}$.
   One can check that $\mathrm{Mod}(T)$ consists of $\{\}$, $\{nbp\}$, $\{uf\}$, $\{uf, hbp\}$, $\{nbp, r\}$, $\{uf, hbp, TV\}$, $\{nbp, r, uf\}$, $\{hbp, nbp, r, TV, uf\}\}$. Therefore, $\mathrm{Mod}(T) = \mathrm{Int}(X, Y, I)$. This implies that T is complete in $\langle X, Y, I\rangle$.

   (An easy way to check it is to check that every intent from $\mathrm{Int}(X, Y, I)$ is a model of T (there are 8 intents in our case), and that no other subset of Y is a model of T (there are $2^5 - 8 = 24$ such subsets in our case). As an example, take $\{hbp, uf, r\} \notin \mathrm{Int}(X, Y, I)$. $\{hbp, uf, r\}$ is not a model of T because $\{hbp, uf, r\}$ is not a model of $\{r\} \Rightarrow \{nbp\}$.)

Example 3.44 (reducing T to a non-redundant set**).** Continuing our previous example, consider again T consisting of $\{r\} \Rightarrow \{nbp\}$, $\{TV, uf\} \Rightarrow \{hbp\}$, $\{TV\} \Rightarrow \{uf\}$, $\{TV\} \Rightarrow \{hbp\}$, $\{hbp, TV\} \Rightarrow \{uf\}$, $\{nbp, uf\} \Rightarrow \{r\}$, $\{hbp\} \Rightarrow \{uf\}$, $\{uf, r\} \Rightarrow \{nbp\}$, $\{nbp, TV\} \Rightarrow \{r\}$, $\{hbp, nbp\} \Rightarrow \{r, TV\}$.

From the previous example we know that T is complete in $\langle X, Y, I\rangle$. Check whether T is non-redundant. If not, transform T into a non-redundant set nrT. (Note: nrT is then a non-redundant basis of $\langle X, Y, I\rangle$.)

Using the above algorithm, we put $nrT := T$ and go through all $A \Rightarrow B \in nrT$ and perform: If for $T' := nrT - \{A \Rightarrow B\}$ we find out that $T' \models A \Rightarrow B$, we remove $A \Rightarrow B$ from nrT, i.e. we put $nrT := T'$. Checking $T' \models A \Rightarrow B$ is done by verifying whether $||A \Rightarrow B||_{C_{\mathrm{Mod}(T')}(A)}$.

- For $A \Rightarrow B = \{r\} \Rightarrow \{nbp\}$: $T' := nrT - \{\{r\} \Rightarrow \{nbp\}\}$, $C_{\mathrm{Mod}(T')}(A) = \{r\}$ and $||A \Rightarrow B||_{\{r\}} = 0$, thus $T' \not\models A \Rightarrow B$, and nrT does not change.

- For $A \Rightarrow B = \{TV, uf\} \Rightarrow \{hbp\}$: $T' := nrT - \{\{TV, uf\} \Rightarrow \{hbp\}\}$, $C_{\mathrm{Mod}(T')}(A) = \{TV, uf, hbp\}$ and $||A \Rightarrow B||_{\{TV,uf,hbp\}} = 1$, thus $T' \models A \Rightarrow B$, and we remove $\{TV, uf\} \Rightarrow \{hbp\}$ from nrT. That is, $nrT = T - \{\{TV, uf\} \Rightarrow \{hbp\}\}$.
- For $A \Rightarrow B = \{TV\} \Rightarrow \{uf\}$: $T' := nrT - \{\{TV\} \Rightarrow \{uf\}\}$, $C_{\mathrm{Mod}(T')}(A) = \{TV, hbp, uf\}$ and $||A \Rightarrow B||_{\{TV,hbp,uf\}} = 1$, thus $T' \models A \Rightarrow B$, and we remove $\{TV\} \Rightarrow \{uf\}$ from nrT. That is, $nrT = T - \{\{TV, uf\} \Rightarrow \{hbp\}, \{TV\} \Rightarrow \{uf\}\}$.
- For $A \Rightarrow B = \{TV\} \Rightarrow \{hbp\}$: $T' := nrT - \{\{TV\} \Rightarrow \{hbp\}\}$, $C_{\mathrm{Mod}(T')}(A) = \{TV\}$ and $||A \Rightarrow B||_{\{TV\}} = 0$, thus $T' \not\models A \Rightarrow B$, nrT does not change. That is, $nrT = T - \{\{TV, uf\} \Rightarrow \{hbp\}, \{TV\} \Rightarrow \{uf\}\}$.
- For $A \Rightarrow B = \{hbp, TV\} \Rightarrow \{uf\}$: $T' := nrT - \{\{hbp, TV\} \Rightarrow \{uhf\}\}$, $C_{\mathrm{Mod}(T')}(A) = \{hbp, TV, uf\}$ and $||A \Rightarrow B||_{\{hbp,TV,uf\}} = 1$, thus $T' \models A \Rightarrow B$, we remove $\{hbp, TV\} \Rightarrow \{uf\}$ from nrT. That is, $nrT = T - \{\{TV, uf\} \Rightarrow \{hbp\}$, $\{TV\} \Rightarrow \{uf\}, \{hbp, TV\} \Rightarrow \{uf\}\}$.

- For $A \Rightarrow B = \{nbp, uf\} \Rightarrow \{r\}$: $T' := nrT - \{\{nbp, uf\} \Rightarrow \{r\}\}$, $C_{\mathrm{Mod}(T')}(A) = \{nbp, uf\}$ and $||A \Rightarrow B||_{\{nbp,uf\}} = 0$, thus $T' \not\models A \Rightarrow B$ and nrT does not change. That is, $nrT = T - \{\{TV, uf\} \Rightarrow \{hbp\}, \{TV\} \Rightarrow \{uf\}, \{hbp, TV\} \Rightarrow \{uf\}\}$.
- For $A \Rightarrow B = \{hbp\} \Rightarrow \{uf\}$: $T' := nrT - \{\{hbp\} \Rightarrow \{uf\}\}$, $C_{\mathrm{Mod}(T')}(A) = \{hbp\}$ and $||A \Rightarrow B||_{\{hbp\}} = 0$, thus $T' \not\models A \Rightarrow B$ and nrT does not change. That is, $nrT = T - \{\{TV, uf\} \Rightarrow \{hbp\}, \{TV\} \Rightarrow \{uf\}, \{hbp, TV\} \Rightarrow \{uf\}\}$.
- For $A \Rightarrow B = \{uf, r\} \Rightarrow \{nbp\}$: $T' := nrT - \{\{uf, r\} \Rightarrow \{nbp\}\}$, $C_{\mathrm{Mod}(T')}(A) = \{uf, r, nbp\}$ and $||A \Rightarrow B||_{\{uf,r,nbp\}} = 1$, thus $T' \models A \Rightarrow B$ and we remove $\{uf, r\} \Rightarrow \{nbp\}$ from nrT. That is, $nrT = T - \{\{TV, uf\} \Rightarrow \{hbp\}$, $\{TV\} \Rightarrow \{uf\}, \{hbp, TV\} \Rightarrow \{uf\}, \{uf, r\} \Rightarrow \{nbp\}\}$.

- For $A \Rightarrow B = \{nbp, TV\} \Rightarrow \{r\}$: $T' := nrT - \{\{nbp, TV\} \Rightarrow \{r\}\}$, $C_{\mathrm{Mod}(T')}(A) = \{nbp, TV, hbp, uf, r\}$ and $||A \Rightarrow B||_{\{nbp,TV,hbp,uf,r\}} = 1$, thus $T' \models A \Rightarrow B$ and we remove $\{nbp, TV\} \Rightarrow \{r\}$ from nrT. That is, $nrT = T - \{\{TV, uf\} \Rightarrow \{hbp\}$, $\{TV\} \Rightarrow \{uf\}$, $\{hbp, TV\} \Rightarrow \{uf\}$, $\{uf, r\} \Rightarrow \{nbp\}$, $\{nbp, TV\} \Rightarrow \{r\}\}$.
- For $A \Rightarrow B = \{hbp, hbp\} \Rightarrow \{r, TV\}$: $T' := nrT - \{\{hbp, nbp\} \Rightarrow \{r, TV\}\}$, $C_{\mathrm{Mod}(T')}(A) = \{hbp, nbp, uf, r\}$ and $||A \Rightarrow B||_{\{hbp,nbp,uf,r\}} = 0$, thus $T' \not\models A \Rightarrow B$ and nrT does not change. That is, $nrT = T - \{\{TV, uf\} \Rightarrow \{hbp\}$, $\{TV\} \Rightarrow \{uf\}$, $\{hbp, TV\} \Rightarrow \{uf\}$, $\{uf, r\} \Rightarrow \{nbp\}$, $\{nbp, TV\} \Rightarrow \{r\}\}$.

We obtained $nrT = \{\{r\} \Rightarrow \{nbp\}, \{TV\} \Rightarrow \{hbp\}, \{nbp, uf\} \Rightarrow \{r\}, \{hbp\} \Rightarrow \{uf\}, \{hbp, nbp\} \Rightarrow \{r, TV\}\}$. nrT is a non-redundant set of AIs.

Since T is complete in $\langle X, Y, I\rangle$, nrT is complete in $\langle X, Y, I\rangle$, too (why?). Therefore, nrT is a non-redundant basis of $\langle X, Y, I\rangle$.

In the last example, we obtained a non-redundant basis nrT of $\langle X, Y, I\rangle$, $nrT = \{\{r\} \Rightarrow \{nbp\}, \{TV\} \Rightarrow \{hbp\}, \{nbp, uf\} \Rightarrow \{r\}, \{hbp\} \Rightarrow \{uf\}, \{hbp, nbp\} \Rightarrow \{r, TV\}\}$.

How to compute non-redundant bases from data?

We are going to present an approach based on the notion of a pseudo-intent. This approach is due to Guigues and Duquenne. The resulting non-redundant basis is called a Guigues-Duquenne basis.

Two main features of Guigues-Duquenne basis are

- it is computationally tractable,
- it is optimal in terms of its size (no other non-redundant basis has is smaller in terms of the number of AIs it contains).

Definition 3.45 (pseudo-intents)**.** A *pseudo-intent* of $\langle X, Y, I\rangle$ is a subset $A \subseteq Y$ for which

1. $A \neq A^{\downarrow\uparrow}$,
2. $B^{\downarrow\uparrow} \subseteq A$ for each pseudo-intent $B \subset A$.

Theorem 3.46 (Guigues-Duquenne basis)**.** *The set* $T = \{A \Rightarrow A^{\downarrow\uparrow} \mid A \textit{ is a pseudointent of } (X, Y, I)\}$ *of implications is a non-redundant basis of* $\langle X, Y, I\rangle$.

Proof. We show that T is complete and non-redundant.

Complete: It suffices to show that $\mathrm{Mod}(T) \subseteq \mathrm{Int}(X, Y, I)$. Let $C \in \mathrm{Mod}(T)$. Assume $C \neq C^{\downarrow\uparrow}$. Then C is a pseudo-intent (indeed, if $P \subset C$ is a pseudo-intent then since $||P \Rightarrow P^{\downarrow\uparrow}||_C = 1$, we get $P^{\downarrow\uparrow} \subseteq C$). But then $C \Rightarrow C^{\downarrow\uparrow} \in T$ and so $||C \Rightarrow C^{\downarrow\uparrow}||_C = 1$. But the last fact means that if $C \subseteq C$ (which is true) then $C^{\downarrow\uparrow} \subseteq C$ which would give $C^{\downarrow\uparrow} = C$, a contradiction with the assumption $C^{\downarrow\uparrow} \neq C$. Therefore, $C^{\downarrow\uparrow} = C$, i.e. $C \in \mathrm{Int}(X, Y, I)$.

Non-redundant: Take any $P \Rightarrow P^{\downarrow\uparrow}$. We show that $T - \{P \Rightarrow P^{\downarrow\uparrow}\} \not\models P \Rightarrow P^{\downarrow\uparrow}$. Since $||P \Rightarrow P^{\downarrow\uparrow}||_P = 0$ (obvious, check), it suffices to show that $||T - \{P \Rightarrow P^{\downarrow\uparrow}\}||_P = 1$. That is, we need to show that for each $Q \Rightarrow Q^{\downarrow\uparrow} \in T - \{P \Rightarrow P^{\downarrow\uparrow}\}$ we have $||Q \Rightarrow Q^{\downarrow\uparrow}||_P = 1$, i.e. that if $Q \subseteq P$ then $Q^{\downarrow\uparrow} \subseteq P$. But this follows from the definition of a pseudo-intent (apply to P). $\square$

Lemma 3.47. *If* P, Q *are intents or pseudo-intents and* $P \not\subseteq Q$, $Q \not\subseteq P$, *then* $P \cap Q$ *is an intent.*

Proof. Let $T = \{R \Rightarrow R^{\downarrow\uparrow} \mid R$ a pseudo-intent$\}$ be the G.-D. basis. Since T is complete, it is sufficient to show that $P \cap Q \in \mathrm{Mod}(T)$ (since then, $P \cap Q$ is a model of any implication which is true in $\langle X, Y, I\rangle$, and so $P \cap Q$ is an intent).

Obviously, P, Q are models of $T - \{P \Rightarrow P^{\downarrow\uparrow}, Q \Rightarrow Q^{\downarrow\uparrow}\}$, whence $P \cap Q$ is a model of $T - \{P \Rightarrow P^{\downarrow\uparrow}, Q \Rightarrow Q^{\downarrow\uparrow}\}$ (since the set of models is a closure system, i.e. closed under intersections).

Therefore, to show that $P \cap Q$ is a model of T, it is sufficient to show that $P \cap Q$ is a model of $\{P \Rightarrow P^{\downarrow\uparrow}, Q \Rightarrow Q^{\downarrow\uparrow}\}$. Due to symmetry, we only verify that $P \cap Q$ is a model of $\{P \Rightarrow P^{\downarrow\uparrow}$: But this is trivial: since $P \not\subseteq Q$, the condition "if $P \subseteq P \cap Q$ implies $P^{\downarrow\uparrow} \subseteq P \cap Q$" is satisfied for free. The proof is complete. □

Lemma 3.48. *If T is complete, then for each pseudo-intent P, T contains $A \Rightarrow B$ with $A^{\downarrow\uparrow} = P^{\downarrow\uparrow}$*

Proof. For pseudointent P, $P \neq P^{\downarrow\uparrow}$, i.e. P is not an intent. Therefore, P cannot be a model of T (since models of a complete T are intents). Therefore, there is $A \Rightarrow B \in T$ such that $||A \Rightarrow B||_P = 0$, i.e. $A \subseteq P$ but $B \not\subseteq P$. As $||A \Rightarrow B||_{\langle X,Y,I\rangle} = 1$, we have $B \subseteq A^{\downarrow\uparrow}$ (Thm. on basic connections ...). Therefore, $A^{\downarrow\uparrow} \not\subseteq P$ (otherwise $B \subseteq P$, a contradiction). Therefore, $A^{\downarrow\uparrow} \cap P$ is not an intent (). By the foregoing Lemma, $P \subseteq A^{\downarrow\uparrow}$ which gives $P^{\downarrow\uparrow} \subseteq A^{\downarrow\uparrow}$. On the other hand, $A \subseteq P$ gives $A^{\downarrow\uparrow} \subseteq P^{\downarrow\uparrow}$. Altogether, $A^{\downarrow\uparrow} = P^{\downarrow\uparrow}$, proving the claim. □

Theorem 3.49 (Guigues-Duquenne basis is the smalest one)**.** *If T is the Guigues-Duquenne base and T' is complete then $|T| \leq |T'|$.*

Proof. Direct corollary of the above Lemma. □

$\mathcal{P}$ ... set of all pseudointents of $\langle X, Y, I\rangle$

THE base we need to compute: $\{A \Rightarrow A^{\downarrow\uparrow} \mid A \in \mathcal{P}\}$

Q: What do we need? A: Compute all pseudointents.

We will see that the set of all $P \subseteq Y$ which are intents or pseudo-intents is a closure system.

Q: How to compute the fixed points (closed sets)?

For $Z \subseteq Y$, T a set of implications, put

$Z^T = Z \cup \bigcup\{B \mid A \Rightarrow B \in T, A \subset Z\}$

$Z^{T_0} = Z$
$Z^{T_n} = (Z^{T_{n-1}})^T$ $(n \geq 1)$

define $C_T : 2^Y \to 2^Y$ by

$C_T(Z) = \bigcup_{n=0}^{\infty} Z^{T_n}$ (note: terminates, Y finite)

Note: this is different from the operator computing the least model $C_{\mathrm{Mod}(T)}(A)$ of T containing A (instead of $A \subseteq Z$, we have $A \subset Z$ here).

Theorem 3.50. *Let $T = \{A \Rightarrow A^{\downarrow\uparrow} \mid A \in \mathcal{P}\}$ (G.-D. base). Then*

1. *C_T is a closure operator,*
2. *P is a fixed point of C_T iff $P \in \mathcal{P}$ (P is a pseudo-intent) or $P \in \mathrm{Int}(X, Y, I)$ (P is an intent).*

Proof. 1. easy (analogous to the proof concerning the closure operator for $C_{\mathrm{Mod}(T)}(A)$).

2. $\mathcal{P} \cup \mathrm{Int}(X, Y, I) \subseteq \mathrm{fix}(C_T)$: easy. $\mathrm{fix}(C_T) \subseteq \mathcal{P} \cup \mathrm{Int}(X, Y, I)$: It suffices to show that if $P \in \mathrm{fix}(C_T)$ is not an intent ($P \neq P^{\downarrow\uparrow}$) then P is an pseudo-intent. So take $P \in \mathrm{fix}(C_T)$, i.e. $P = C_T(P)$, which is not an intent. Take any pseudointent $Q \subset P$. By definition (notice that $Q \Rightarrow Q^{\downarrow\uparrow} \in T$), $Q^{\downarrow\uparrow} \subseteq C_T(P) = P$ which means that P is a pseudo-intent. $\square$

So: $\mathrm{fix}(C_T) = \mathcal{P} \cup \mathrm{Int}(X, Y, I)$

Therefore, to compute $\mathcal{P}$, we can compute $\mathrm{fix}(C_T)$ and exclude $\mathrm{Int}(X, Y, I)$, i.e. $\mathcal{P} = \mathrm{fix}(C_T) - \mathrm{Int}(X, Y, I)$.

computing $\mathrm{fix}(C_T)$: by Ganter's NextClosure algorithm.

Caution! In order to compute C_T, we need T, i.e. we need $\mathcal{P}$, which we do not know in advance. Namely, recall what we are doing:

- Given input data $\langle X, Y, I\rangle$, we need to compute G.-D. basis $T = \{A \Rightarrow A^{\downarrow\uparrow} \mid A \in \mathcal{P}\}$.
- For this, we need to compute $\mathcal{P}$ (pseudo-intents of $\langle X, Y, I\rangle$).
- $\mathcal{P}$ can be obtained from $z\mathrm{fix}(C_T)$ (fixed points of C_T).
- But to compute C_T, we need T (actually, we need only a part of T).

But we are not in *circulus vitiosus*: The part of T (or $\mathcal{P}$) which is needed at a given point is already available (computed) at that point.

Computing G.-D. basis manually is tedious. Algorithms available, e.g. Peter Burmeister's ConImp software.

References

[1] Barbut M., Monjardet B.: *L'ordre et la classification, algèbre et combinatoire, tome II.* Paris, Hachette, 1970.

[2] Carpineto C., Romano G.: *Concept Data Analysis. Theory and Applications.* J. Wiley, 2004.

[3] Correia J. H., Stumme G., Wille R.: Conceptual Knowledge Discovery—A Human-Centered Approach. *Applied Artificial Intelligence* **17**,3 (2003), 281–302.

[4] Ganter B.: Attribute Exploration with Background Knowledge. *Theor. Comput. Sci.* **217**,2 (1999), 215–233.

[5] Ganter B., Wille R.: *Formal Concept Analysis. Mathematical Foundations.* Springer, 1999.

[6] Ganter B., Stumme G., Wille R.: *Formal Concept Analysis. Foundations and Applications.* Springer, 2005.

[7] Guigues J.-L., Duquenne V.: Familles minimales dimplications informatives resultant dun tableau de donnes binaires. *Math. Sci. Humaines* **95** (1986), 5–18.

[8] Kuznetsov S., Obiedkov S., Comparing performance of algorithms for generating concept lattices. *J. Experimental and Theoretical Articial Intelligence* **14**,2–3 (2002), 189–216.

[9] Ore O.: Galois connexions. *Trans. Amer. Math. Soc.* **55** (1944), 493–513.

[10] Wille R.: Restructuring lattice theory: an approach based on hierarchies of concepts. In: I. Rival (Ed.): *Ordered Sets*, 445–470, Reidel, Dordrecht-Boston, 1982.

[11] Wille R.: Methods of Conceptual Knowledge Processing. *ICFCA 2006*, LNAI **3874**, Springer, Heidelberg 2006, pp. 1–29.