



# Approximation of Sets Based on Partial Covering

Ph.D. Thesis

**Zoltán Csajbók**

SUPERVISOR: DR. TAMÁS MIHÁLYDEÁK CSC  
PROF. DR. ATTILA PETHŐ

UNIVERSITY OF DEBRECEN  
DOCTORAL COMMITTEE OF NATURAL SCIENCES  
DOCTORAL SCHOOL OF COMPUTER SCIENCES

Debrecen, 2011



Ezen értekezést a Debreceni Egyetem Természettudományi Tudományterületi Doktori Tanács Informatikai Tudományok Doktori Iskola Elméleti számítástudomány, adatvédelem és kriptográfia programja keretében készítettem a Debreceni Egyetem doktori (PhD) fokozatának elnyerése céljából.

Debrecen, 2011. október 14.

.....  
Csajbók Zoltán  
jelölt

Tanúsítom, hogy Csajbók Zoltán doktorjelölt 1999 - 2007 között az Diofantikus és Konstruktív Számelmélet PhD programjának keretében irányításommal végezte munkáját. Az értekezésben foglalt eredményekhez a jelölt önálló alkotó tevékenységével meghatározóan hozzájárult. Az értekezés elfogadását javaslom.

Debrecen, 2011. október 14.

.....  
Prof Dr. Pethő Attila  
témavezető

Tanúsítom, hogy Csajbók Zoltán doktorjelölt 2008 - 2011 között az Informatikai Tudományok Doktori Iskola Elméleti számítástudomány, adatvédelem és kriptográfia programjának keretében irányításommal végezte munkáját. Az értekezésben foglalt eredményekhez a jelölt önálló alkotó tevékenységével meghatározóan hozzájárult. Az értekezés elfogadását javaslom.

Debrecen, 2011. október 14.

.....  
Dr. Mihálydeák Tamás  
témavezető



# Approximation of Sets Based on Partial Covering

Értekezés a doktori (Ph.D.) fokozat megszerzése érdekében  
az informatika tudományágban.

Írta: Csajbók Zoltán okleveles matematikus, matematika szakos középiskolai tanár.

Készült a Debreceni Egyetem Informatikai Tudományok Doktori Iskolája (Elméleti  
számítástudomány, adatvédelem és kriptográfia programja) keretében.

Témavezető: Dr. Mihálydeák Tamás  
Prof. Dr. Pethő Attila

A doktori szigorlati bizottság:

elnök: Dr. ....

tagok: Dr. ....

Dr. ....

A doktori szigorlat időpontja: 2011. ....

Az értekezés bírálói:

Dr. ....

Dr. ....

Dr. ....

A bírálóbizottság:

elnök: Dr. ....

tagok: Dr. ....

Dr. ....

Dr. ....

Dr. ....

Az értekezés védésének időpontja: 2011. ....



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	A Historical Outline . . . . .	1
1.2	Basic Philosophical Background . . . . .	3
1.3	Our Approach . . . . .	6
1.4	Thesis Overview and our Results . . . . .	8
<b>2</b>	<b>Basic Concepts</b>	<b>13</b>
2.1	Basic Notations . . . . .	13
2.2	Galois Connections . . . . .	14
<b>3</b>	<b>General Approximation Frameworks</b>	<b>17</b>
3.1	An Initial Approximation Framework . . . . .	17
3.2	A General Set Theoretic Approximation Framework . . . . .	19
<b>4</b>	<b>Fundamentals of Rough Set Theory</b>	<b>29</b>
4.1	Basic Notions . . . . .	29
4.2	Granularity Aspects of Rough Set Theory . . . . .	34
4.3	Galois Connection of Upper and Lower Approximations . . . . .	35
<b>5</b>	<b>Approximation of Sets Based on Partial Covering</b>	<b>41</b>
5.1	Introduction . . . . .	41
5.2	Base Systems . . . . .	42
5.3	Single-Layered Base Systems . . . . .	42
5.4	Lower and Upper $\mathfrak{B}$ -Approximations . . . . .	45
5.5	Representation of Sets . . . . .	48
5.6	Exactness in $\mathfrak{B}$ -approximation spaces . . . . .	50
5.7	A Possible Interpretation of Our Approach . . . . .	52

<b>6</b>	<b>Galois Connections</b>	<b>55</b>
6.1	Regular Galois Connection . . . . .	55
6.2	Partial Galois Connection . . . . .	58
6.2.1	On Partial Lower $\mathfrak{B}$ -approximations . . . . .	58
6.2.2	Partial Upper $\mathfrak{B}$ -approximations . . . . .	60
<b>7</b>	<b>Applications</b>	<b>65</b>
7.1	Natural Computing—A Biological Example . . . . .	66
7.1.1	A Brief Outline of the MÉTA program . . . . .	66
7.1.2	Model of Behavioral Features of Natural Vegetation . . . . .	69
7.2	A General Tool-Based Approximation Framework . . . . .	71
7.3	Simultaneous Anomaly and Misuse Intrusion Detections . . . . .	82
7.3.1	Introduction . . . . .	82
7.3.2	The Intrusion Detection Model . . . . .	84
	<b>Summary</b>	<b>93</b>
	<b>Összefoglaló (Hungarian summary)</b>	<b>97</b>
	<b>Bibliography</b>	<b>102</b>
<b>A</b>	<b>List of Papers of the Author</b>	<b>113</b>
<b>B</b>	<b>Additional Publications</b>	<b>117</b>
<b>C</b>	<b>List of talks of the author</b>	<b>123</b>
<b>D</b>	<b>Acknowledgements</b>	<b>125</b>
<b>E</b>	<b>Köszönetnyilvánítások</b>	<b>127</b>



# Chapter 1

## Introduction

### 1.1 A Historical Outline

The rough set theory (RST), among others, is a mathematical tool to manage inexact, uncertain, incomplete and imperfect data. It was invented by Zdzisław Pawlak in the early 1980s [53, 54].

The starting point is a nonempty *finite* set  $U$  of distinguishable objects, called the *universe of discourse*, and an *equivalence relation*  $\varepsilon$  on  $U$  [55]. The partition of  $U$  generated by  $\varepsilon$  is denoted by  $U/\varepsilon$ , and its elements are called  $\varepsilon$ -*elementary* sets (Fig. 1.1). An  $\varepsilon$ -elementary set can be viewed as a set of indiscernible objects characterized by the same available information about them [58, 68]. In addition, any union of  $\varepsilon$ -elementary sets is referred to as *definable* set (Fig. 1.2).

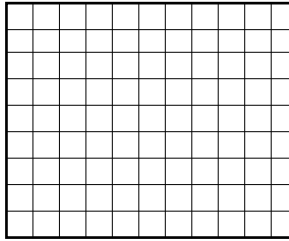


Figure 1.1:  $\varepsilon$ -elementary sets

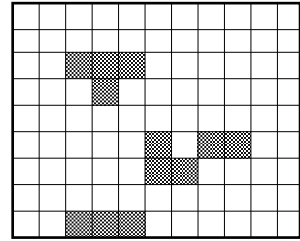


Figure 1.2. Definable sets

Any subset  $X \subseteq U$  can be naturally approximated by two sets called the lower and upper  $\varepsilon$ -approximations of  $X$ . The *lower  $\varepsilon$ -approximation* of  $X$  is the union of all the  $\varepsilon$ -elementary sets which are the subsets of  $X$  (Fig. 1.3), whereas the *upper  $\varepsilon$ -approximation* of  $X$  is the union of all the  $\varepsilon$ -elementary sets that have a nonempty intersection with  $X$  (Fig. 1.4).

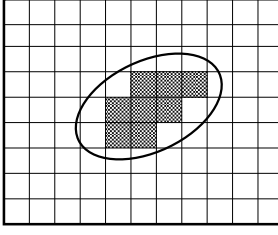


Figure 1.3. Lower approximation

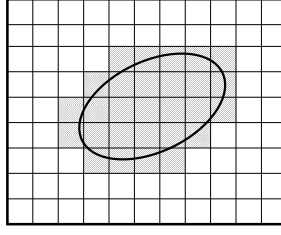


Figure 1.4. Upper approximation

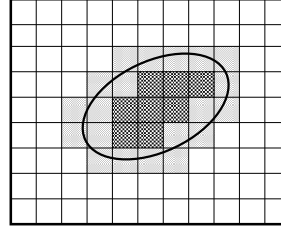


Figure 1.5. Lower-upper approximation

The difference between upper and lower  $\varepsilon$ -approximations is called the  $\varepsilon$ -boundary of  $X$  (Fig. 1.5). The subset  $X$  is  $\varepsilon$ -crisp (exact), if its  $\varepsilon$ -boundary is the empty set,  $\varepsilon$ -rough (inexact) otherwise.

Let  $\sigma(U/\varepsilon)$  denote the extension of  $U/\varepsilon$  with all the unions of some  $\varepsilon$ -elementary sets and the empty set. It is easy to see that  $\sigma(U/\varepsilon) \subseteq 2^U$  is a  $\sigma$ -algebra generated by  $U/\varepsilon$ , i.e. it is nonempty, closed under complementations and countable unions. In other words,  $(U, \sigma(U/\varepsilon))$  is an Alexandrov topological space with the basis  $U/\varepsilon$ .  $\sigma(U/\varepsilon)$  is the family of all open and closed sets [36, 67].

In Pawlak's theory, the lower and upper  $\varepsilon$ -approximations can be defined by three equivalent forms. These three forms are based on elements,  $\varepsilon$ -elementary sets and the  $\sigma$ -algebra  $\sigma(U/\varepsilon)$  [76, 77, 80]. In any case, both lower and upper  $\varepsilon$ -approximations of any subset  $X \subseteq U$  belong to the  $\sigma$ -algebra  $\sigma(U/\varepsilon)$ .

The three equivalent definitions offer different interpretations of Pawlak's approximations. According to the element based formulation, the lower and upper approximation operators can be interpreted as the necessity and possibility operators of *modal logic* [35, 81]. The  $\sigma$ -algebra based formulation relates them to interior and closure operators in *topological spaces* [86]. The formulation based on  $\varepsilon$ -elementary sets has been served as the "pattern" of *granular computing* developments [44, 69, 79, 83]. Nowadays, the granular computing is a fast developing of the one branch of information technology.

The generalization of Pawlak's approximations can go along one of the three equivalent definitions mentioned above. A natural generalization of Pawlak's idea via the element based definition is that the equivalence relation is replaced by any other type of binary relation on  $U$  [32, 36]. Another generalization can be obtained by using any covering of the universe and the imitation of the  $\varepsilon$ -elementary set based definition [86, 87].

The case of  $\sigma$ -algebra based definition is a little more complicated. In the language of Alexandrov topological spaces, the  $\sigma$ -algebra  $\sigma(U/\varepsilon)$  is the family of clopen sets, i.e. the family of open sets coincides with the family of closed sets. The family of open sets is related to the lower approximation or interior operator, whereas the family of closed sets is related to the upper approximation or closure operator. As a possible generalization, one may use two different subsystems of the powerset of  $U$  [78]. A subsystem for the lower approximation which must be closed under unions and another subsystem for the upper approximation which, in turn, must be closed under intersections. Moreover, in order to keep the duality of lower and upper approximation operators, the elements of two subsystems must be related to each other through the complementation. In addition, this latter restriction can also be removed [76].

A list of some *research directions* on the rough set foundations and the rough set based methods can be found in [57].

Rough set theory can be applied among others in the areas of artificial intelligence, cognitive sciences, medicine and economics. It provides a powerful foundation to reveal and discover important structures and patterns in data and to classify complex objects. One of the main advantages of rough set theory is that it does not need any preliminary or additional information about data [62, 64]. This attractive property of rough set theory is of especial importance for instance to data mining, machine learning, decision analysis, knowledge management, expert systems, pattern recognition, medicine, engineering, banking, financial and market analysis [62, 64, 85].

## 1.2 Basic Philosophical Background

There is a philosophical interpretation of the rough set theory too. It may also be seen as a relatively new possible mathematical approach to *vagueness* [38, 52, 57, 63]. According to the entry for ‘vagueness’ in the *Stanford Encyclopedia of Philosophy*:

There is wide agreement that a term is vague to the extent that it has borderline cases. This makes the notion of a borderline case crucial in accounts of vagueness. ([70], the two introductory sentences.)

Vagueness is standardly defined as the possession of borderline cases. For example, ‘tall’ is vague because a man who is 1.8 meters in height is neither clearly tall nor clearly non-tall. No amount of *conceptual analysis* or *empirical investigation* can settle whether a 1.8 meter man is tall. ([70], Chapter 1. *The italics are mine.*)

Borderline cases are inquiry resistant. Indeed, the inquiry resistance typically recurses. For in addition to the unclarity of the borderline case, there is normally unclarity as to where the unclarity begins. In other words ‘borderline case’ has borderline cases. This higher order vagueness shows that ‘vague’ is vague. ([70], Chapter 1.)

That is vague terms lack well-defined extensions—there is no sharp boundary between tall people and the rest [38]. In other words a set of objects is *vague* if objects exist that cannot be classified as belonging to either the set or its complement [58]. It should immediately be noted that in this context the notion ‘set’ is used in a pre-theoretic sense.

The ‘vagueness’ is a more than two thousand-year-old problem. Its origins back to the so-called *Sorites paradox* [38, 39, 61, 73] attributed to Aristotle’s contemporary Eubulides of Miletus (4th c. BC), the Megarian logician. The word ‘sorites’ in Greek it means ‘heap’. (To be more precise, the paradox derives its name from the Greek word *soros*.) Note that the far known Liar paradox in its purest form is also attributed to Eubulides.

One of the form of the Sorites paradox is the following. Of course, one stone does not make a heap. Adding only one stone to what is not yet a heap surely cannot make a heap. Repeating this step adding stones one by one we arrive at the conclusion the heaps do not exist not even if they consist of more than, say, 100,000 stones. Then where do we draw the line between what is a heap of stones and what is not?

I agree with Priest [61]: the Sorites is a very hard paradox, possibly harder than the Liar. For the Liar can be isolated, whereas the Sorites is everywhere and can take us anywhere. And I agree that the paradox is so hard because it systematically imposes upon us the existence of unbelievable or otherwise unacceptable *cut-off points*. No solution can avoid explaining why this happens. ([73], p. 24. *The italics are mine.*)

The counter-intuitiveness of Sorites phenomena lies in the fact that there must be a cut-off, *regardless* of where exactly it is located in the soritical sequence. ([73], p. 34. *The italics are the author’s.*)

The Sorites paradox is not mere a curiosity such as R. Keefe and P. Smith remarked in [40]. To confirm this statement, let us look at the following example coming from medicine.

**1.1 Example.** Let us consider the fasting blood glucose test [3] which is used to screen for and diagnose diabetes. It is measured on a fast basis, i.e. collected after an 8 and 10 hours fast. The test measures the amount of

glucose in the blood right at the time of sample collection. On the clinical practice recommendations of the American Diabetes Association, the fasting glucose level is normal if the test result is between 3.9 mmol/L and 5.5 mmol/L, and indicates diabetes over 11.1 mmol/L on more than one testing occasion.

Now, e.g., the fasting glucose level 4.5 mmol/L is normal. Plausibly, increasing the normal fasting glucose level by 0.001 mmol/L (or 0.00001 mmol/L, if necessary) cannot make a difference. So, if the fasting glucose level 4.0 mmol/L is normal then 4.0 mmol/L plus 0.001 mmol/L is also normal. Now, since the fasting glucose level 4.001 mmol/L is normal, 4.001 mmol/L plus 0.001 mmol/L is also normal; and so on. Consequently, any fasting glucose level is normal, even if it is greater than, say, 25.0 mmol/L.

The sorites paradox was not an attractive problem until the late 19th century. Next, numerous logicians and philosophers have dealt with it. The anthology [40] collects for the first time the most important classical papers in the field.

The vagueness associated with the boundary region approach was first formulated in 1893 by G. Frege [27], next Peirce in 1902 [59].

Pawlak's fundamental view of vagueness can be characterized as "unable to classify" [56, 57, 58]. As Pawlak and Skowron have written in [58]:

In contrast to odd numbers, the notion of a beautiful painting is vague, because we are *unable to classify* uniquely all paintings into two classes: beautiful and not beautiful. Some paintings cannot be decided whether they are beautiful or not and thus they remain in the doubtful area. Thus, beauty is not a precise but a vague concept. ([58], p. 5. *The italics are mine.*)

However, in spite of the fact that vagueness is very interesting phenomenon in philosophy, it is not allowed within standard mathematics. Pawlak's information-based solution concerning vagueness is the following:<sup>1</sup>

[...] in the proposed approach, we assume that any vague concept is replaced by a *pair of precise concepts*—called the lower and the upper

---

<sup>1</sup>There is another contemporary information-based solution proposal concerning vagueness, namely, Zadeh's fuzzy set theory [82]. "Zadeh's introduction of fuzzy sets was not meant to be a contribution to the philosophy of vagueness. It was motivated by the need for a computational representation for linguistic terms appearing in statements, which are often intended to provide synthetic information about complex situations." ([24], p. 893). Fuzzy set theory is complementary to rough set theory. In this thesis, this aspect is only mentioned here.

approximation of the vague concept. The lower approximation consists of all objects which surely belong to the concept and the upper approximation contains all objects which possibly belong to the concept. The difference between the upper and the lower approximation constitutes the boundary region of the vague concept. Approximations are two basic operations in rough set theory.

Hence, rough set theory expresses vagueness *not by means of membership, but by employing a boundary region of a set*. If the boundary region of a set is empty it means that the set is crisp, otherwise the set is rough (inexact). A non-empty boundary region of a set means that our knowledge about the set is not sufficient to define the set precisely. ([58], p. 6. *The italics are mine.*)

In sum, Pawlak's approach can be viewed as a specific implementation of Frege's idea of vagueness [27], i.e. imprecision is expressed by a boundary region of a set.

### 1.3 Our Approach

There are many possibilities to generalize the rough set theory. To sum up, our approach has three main foundation-stones:

- (1) “*unable to classify*” as the base of vagueness,
- (2) its presentation in a *point-free* manner, and
- (3) *partiality* of our knowledge about the universe.

*Ad 1.* Rough set theory has been served as a “pattern” of granular computing (GrC). However, there are fundamental differences between them. Granular computing and also rough set theory have three semantic views, in particular, uncertainty theory, knowledge engineering and how-to-solve/compute-it [42, 43]. The most important difference between the two theories is best illustrated in connection with the uncertainty theory. Pawlak uses “unable to classify” as the *base of uncertainty*, while the granular computing regards a granule as a *unit of uncertainty* [43].

*Ad 2.* The philosophy of the rough set theory relies on the assumption that some information (data, knowledge) are associated with every object of the universe. Objects characterized by the same information are *indiscernible* or *similar* in view of the available information about them. A set of all indiscernible or similar objects form a unit of the basic knowledge. Such

a unit can be seen in a *point-wise* manner, i.e. the content of the unit is visible, and in a *point-free* manner, i.e. the content of the unit is hidden. We abstract each unit into a point. Such collection of points is called the quotient structure. We will work on quotient structures, in other words we manage units in the point-free manner.

For more details concerning points (1) and (2) see [42, 43, 44].

*Ad 3.* In real life, information being at our disposal is generally insufficient. Consequently, it is natural to assume that there may be objects which we are unable to characterize at all. Moreover, there are features with which we can form a set of objects effectively, but we cannot form its complement effectively at the same time. For instance, the complements of a recursively enumerable set is not necessarily a recursively enumerable set as well [50].

In the rough set theory, the sets used for to approximation are the equivalence classes which are pairwise disjoint and cover the base set. If we give up the requirement of the pairwise disjoint, we get a kind of generalization of the theory (Fig. 1.6). Its detailed elaboration can be found in the literature.

The main question of the thesis is what would happen if we gave up not only the pairwise disjoint but also the covering of the base set (Fig 1.7.) The resulting system is called the *approximation of sets based on partial covering*.

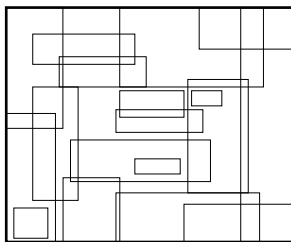


Figure 1.6. Giving up the pairwise disjoint

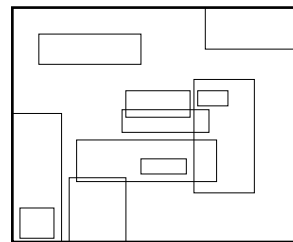


Figure 1.7 Giving up the covering: partial base system

In our thesis we examine the properties of the approximation of sets under these unusual conditions. At this most general abstraction level, we make the only essential condition that *the lower approximation of any set must be included in its upper approximation*.

In the theory of the approximation of sets based on partial covering, let our starting point be an arbitrary nonempty family  $\mathfrak{B}$  of subsets of an arbitrary nonempty universe of discourse  $U$  [14, 15, 60, 84]. Its elements are called  $\mathfrak{B}$ -sets. On the analogy of the definition of the  $\sigma$ -algebra  $\sigma(U/\varepsilon)$ , let  $\mathfrak{D}_{\mathfrak{B}}$  denote the extension of  $\mathfrak{B}$  with the empty set and all the unions of some  $\mathfrak{B}$ -sets. In other words,  $\mathfrak{D}_{\mathfrak{B}}$  is closed under arbitrary unions and contains

every set in  $\mathfrak{B}$  plus the empty set. However,  $\mathfrak{D}_{\mathfrak{B}}$  *neither covers the universe* (i.e. it does not contain  $U$ ) *nor forms  $\sigma$ -algebra* in general. Similarly to the rough set theory, any union of  $\mathfrak{B}$ -sets is referred to as  $\mathfrak{B}$ -definable set.

Our notion of lower and upper approximations are straightforward point-free generalizations of Pawlak's same approximation operators imitating the  $\varepsilon$ -elementary set based formulae and both of them belong to  $\mathfrak{D}_{\mathfrak{B}}$ . So, our lower and upper approximation operators are of the form  $2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}$ . This approach corresponds to the "a priori" attitude in the sense of [11]. Here,  $\mathfrak{D}_{\mathfrak{B}}$  is the family of the fundamental sets of our framework which can be seen as the *tools* which we use to approximate any subset of  $U$ . However, we have to emphasize that  $\mathfrak{D}_{\mathfrak{B}}$  is just the set of definable sets, not the set of exact sets (in the sense of Chapter 5).

Our discussion will be within an overall approximation framework the scope of which ranges from the weak approximation pair of maps on  $U$  [25] to the notion of Galois connection on  $2^U$  [21, 22, 32]. Along this framework, the common features of both the rough set theory and our approach can be treated uniformly. In addition, most notions of Pawlak's rough set theory constitute compound ones and they are split into two or more parts in our approach. This framework helps us to understand the state of their compound nature and to specify their constituents in a more general context.

Last but not least, it has been proved that the partial approximation of sets can be applied to solving practical problems [15, 12, 18, 19].

## 1.4 Thesis Overview and our Results

The present thesis can be divided into three main parts. (1) Chapter 1–2 are two introductory chapters; (2) Chapter 3–6 contain our theoretical results; (3) Chapter 7 presents different real-life applications of our approach.

More precisely, this dissertation consists of the following parts.

Chapter 1 is an introduction. It contains a historical outline, a philosophical background, the brief summary of our approach and, finally, the thesis overview and our main results.

Chapter 2 summarizes the basic concepts and notations used throughout the thesis.

Chapter 3 defines two general approximation frameworks, a large-scaled initial one, called the Initial Approximation Framework, and a finer-scaled



one, called the General Set Theoretic Approximation Framework. They allow us to treat the common features of the classic rough set theory and its generalizations uniformly.

The results of Chapter 3 are based on

Z. Csajbók: *Partial Approximative Set Theory: A Generalization of the Rough Set Theory*, Proceedings of SoCPaR 2010, IEEE (see, [14])

Z. Csajbók: *On the General Set Theoretical Framework of Set Approximation*, Proceedings of RST 2011 (see, [20])

Chapter 4 is devoted to the basic concepts and properties of the classic rough set theory relying on the General Set Theoretic Approximation Framework. We partly restate some well-known facts in the language of our approximation framework and provide new point-free proofs for a few of them.

The results of Chapter 4 are based on

Z. Csajbók: *Approximation of sets based on partial covering*, Theoretical Computer Science: Theory of Natural Computing Thematic Special Issue, 2011 (see, [12])

Z. Csajbók: *On the Partial Approximation of Sets*, Acta Mathematicae et Sociologica (see, [16])

Z. Csajbók: *Partial Approximative Set Theory: A Generalization of the Rough Set Theory*, Proceedings of SoCPaR 2010, IEEE (see, [14])

Z. Csajbók: *Rudiments of Partial Approximative Set Theory* (in Hungarian), Proceedings of the 3rd International Doctoral (PHD/DLA) Conference, 2009 (see, [13])

Chapter 5 presents a special approximation framework based on the partial covering of the universe. It is fully integrated into the General Set Theoretic Approximation Framework.

After some introductory remarks in Section 5.1, Section 5.2 defines the most fundamental concepts of our approach, the base system  $\mathfrak{B}$  and the family of  $\mathfrak{B}$ -definable subsets.

Section 5.3 introduce a constrained version of  $\mathfrak{B}$ , called the single-layered base system. This allows us to prove some properties of our approximation framework which in a sense are similar to the properties of classic rough set theory.

Section 5.4 defines the lower and upper approximations based on partial covering of the universe. First, we prove that they fit into the General Set Theoretic Approximation Framework. Lower  $\mathfrak{B}$ -approximation is always contractive, but upper  $\mathfrak{B}$ -approximation is extensive if and only if the base system  $\mathfrak{B}$  covers the universe. We also show that the  $\mathfrak{B}$ -definable property is generally not equivalent to the equality of lower and upper  $\mathfrak{B}$ -approximations unlike Pawlak's rough set theory. The universe, the family of  $\mathfrak{B}$ -definable sets, the lower and upper  $\mathfrak{B}$ -approximations form together a so-called  $\mathfrak{B}$ -approximation space.

Section 5.5 discusses the  $\mathfrak{B}$ -representations of the  $\mathfrak{B}$ -definable sets. A subset  $D$  is  $\mathfrak{B}$ -representable, if there exists exactly one family of  $\mathfrak{B}$ -sets such that its union equals to  $D$ . We prove that all  $\mathfrak{B}$ -definable subset of the universe are  $\mathfrak{B}$ -representable if and only if the base system  $\mathfrak{B}$  is single-layered. We also give the explicit  $\mathfrak{B}$ -representations of  $\mathfrak{B}$ -definable subsets, among others, the lower and upper  $\mathfrak{B}$ -approximations, when the base system  $\mathfrak{B}$  is single-layered.

Section 5.6 is about an especial important notion of approximation spaces, namely, the exactness. In Pawlak's approximation spaces the notions of 'crisp' (i.e. the exactness) and 'definable' are synonymous to each other. However, a  $\mathfrak{B}$ -definable subset is not necessarily  $\mathfrak{B}$ -crisp. Consequently, the notions of 'definable' and 'crisp' (exactness) are not synonymous to each other in  $\mathfrak{B}$ -approximation spaces.

In Section 5.7, we give a possible interpretation of our approach.

The results of Chapter 5 are based on

Z. Csajbók: *Approximation of sets based on partial covering*,  
Theoretical Computer Science: Theory of Natural Computing  
Thematic Special Issue, 2011 (see, [12])

Z. Csajbók: *On the Partial Approximation of Sets*, Acta Medic-  
inae et Sociologica (see, [16])

In Chapter 6, we investigate what conditions have to be satisfied by the upper and lower  $\mathfrak{B}$ -approximations so that they form a Galois connection on  $(2^U, \subseteq)$ .

In Section 6.1 we prove that the upper and lower  $\mathfrak{B}$ -approximations form a Galois connection on  $(2^U, \subseteq)$  if and only if the base system  $\mathfrak{B}$  is a partition of  $U$ .

In Section 6.2, we deal with partial lower and upper  $\mathfrak{B}$ -approximations.

The empty set may be the lower  $\mathfrak{B}$ -approximation of certain nonempty subsets provided that all singletons are not  $\mathfrak{B}$ -definable. Excluding to allow that the empty set to be the lower  $\mathfrak{B}$ -approximation of a nonempty subset, we obtain the partial variant of the lower  $\mathfrak{B}$ -approximation. We show that under well-defined conditions there exists a unique total extension of the partial lower  $\mathfrak{B}$ -approximation which is exactly the lower  $\mathfrak{B}$ -approximation.

The empty set may be the upper  $\mathfrak{B}$ -approximation of certain nonempty subsets provided that the base system does not cover the universe. Excluding these uncommon cases we obtain the partial variant of the upper  $\mathfrak{B}$ -approximation. We prove that the partial upper  $\mathfrak{B}$ -approximation and the lower  $\mathfrak{B}$ -approximation form a partial Galois connection in the sense of Miné if and only if the  $\mathfrak{B}$ -sets are pairwise disjoint.

The results of Chapter 6 are based on

Z. Csajbók: *Approximation of sets based on partial covering*, Theoretical Computer Science: Theory of Natural Computing Thematic Special Issue, 2011 (see, [12])

Z. Csajbók: *Partial Approximative Set Theory: A View from Galois Connections*, Proceedings of ICAI 2010, Eger, Hungary, UNIDEB Faculty of Informatics – Eszterházy Károly College, 2011 [17]

Z. Csajbók: *Partial Approximative Set Theory: A Generalization of the Rough Set Theory*, Proceedings of SoCPaR 2010, IEEE (see, [14])

Chapter 7, to demonstrate the effectiveness of our approach, presents three real-life applications.

Section 7.1. The first application shows the relationship of our approach with natural computing via a biological application. In particular, we show how our approach helps us to understand some behavioral features of the natural vegetation heritage of Hungary.

The results of Section 7.1 are based on

Z. Csajbók: *Approximation of sets based on partial covering*, Theoretical Computer Science: Theory of Natural Computing Thematic Special Issue, 2011 (see, [12])

Section 7.2. The second application presents a general tool-based approximation framework. In practice, two relevant groups of observed objects can be separated. A group whose elements really possess some features in question and another group whose elements do not substantially possess the same features. To model this situation, two separated approximation spaces are defined over the universe. Then, any collections of the observed objects can *simultaneously* be approximated in the two approximation spaces.

The results of Section 7.2 are based on

Z. Csajbók, T. Mihálydeák: *A General Tool-Based Approximation Framework Based on Partial Approximation of Sets*, Proceedings of RSFDGrC 2011, Moscow, Russia, Springer-Verlag, LNAI 6743, 2011 (see, [19])

Section 7.3. The third example applies the tool-based approximation framework to model Intrusion Detection Systems (IDS) in computer security. In accordance with this framework, anomalies and misuses can be detected at the same time due to its simultaneous nature.

The results of Section 7.3 are based on

Z. Csajbók, T. Mihálydeák: *A General Tool-Based Approximation Framework Based on Partial Approximation of Sets*, Proceedings of RSFDGrC 2011, Moscow, Russia, Springer-Verlag, LNAI 6743, 2011 (see, [19])

Z. Csajbók: *Simultaneous Anomaly and Misuse Intrusion Detections Based on Partial Approximative Set Theory*, Proceedings PDP 2011, Ayia Napa, Cyprus, 2011, IEEE (see, [18])

Z. Csajbók: *A Security Model for Personal Information Security Management Based on Partial Approximative Set Theory*, Proceedings of IMCSIT 2010, Wisła, Poland, 2010, IEEE-PTI (see, [15])

## Chapter 2

# Basic Concepts

### 2.1 Basic Notations

Let  $U$  be any nonempty set. Let  $\mathfrak{A} \subseteq 2^U$  be a family of sets whose elements are subsets of  $U$ .

The union and intersection of  $\mathfrak{A}$  are  $\bigcup \mathfrak{A} = \{x \mid \exists A \in \mathfrak{A} (x \in A)\}$  and  $\bigcap \mathfrak{A} = \{x \mid \forall A \in \mathfrak{A} (x \in A)\}$ , respectively.

If  $\mathfrak{A}$  is empty we define  $\bigcup \emptyset = \emptyset$  and  $\bigcap \emptyset = U$ .

If  $\epsilon \subseteq U \times U$  is an arbitrary binary relation on  $U$ , let  $[x]_\epsilon$  denote the  $\epsilon$ -related elements to  $x$ , i.e.,  $[x]_\epsilon = \{y \in U \mid (x, y) \in \epsilon\}$ . They are called  $\epsilon$ -elementary sets, and the family of  $[x]_\epsilon$  is denoted by  $U/\epsilon$ .

Let  $|A|$  denote the *cardinality* of any set  $A$ .

Let  $X$  and  $Y$  be nonempty sets and  $f : X \rightarrow Y$  be a map. If  $\text{dom} f = X$ ,  $f$  is *total*, if  $\text{dom} f \subsetneq X$ ,  $f$  is *partial*. If  $f$  is a partial map, then  $\text{dom} f = \emptyset$  is allowed. For the purpose of simplicity we will talk about partial maps without direct references to their partiality. However, statements with respect to partial maps always concern their restrictions to their domains.

A nonempty set  $P$  together with a partial order  $\leq$  on  $P$  is called a *partial ordered set* or a *poset*, in symbol  $(P, \leq)$ . Any subset of a poset is in itself a poset which is partially ordered by the same (relative or induced) partial ordering relation.

A self-map  $f : P \rightarrow P$  on  $(P, \leq)$  is

- *extensive* if  $x \leq f(x)$ ;
- *contractive* if  $f(x) \leq x$ ;
- *idempotent* if  $f(f(x)) = f(x)$ ;

- *normalized* if  $f(m) = m$ , when the minimal element  $m \in P$  exists;
- *co-normalized* if  $f(M) = M$ , when the maximal element  $M \in P$  exists.

If  $(P, \leq_P)$  and  $(Q, \leq_Q)$  are two posets, a map  $f : P \rightarrow Q$  is *monotone* or *order-preserving* when  $x \leq_P y \Rightarrow f(x) \leq_Q f(y)$ , and *antitone* or *order-reversing* when  $x \leq_P y \Rightarrow f(y) \leq_Q f(x)$ .

A map  $f : P \rightarrow Q$  is the *order isomorphism* between  $(P, \leq_P)$  and  $(Q, \leq_Q)$  if  $f$  is a bijection and both  $f$  and  $f^{-1}$  are monotone. In this case, it is said that  $P$  and  $Q$  are *isomorph*.

## 2.2 Galois Connections

Let  $(P, \leq_P)$  and  $(Q, \leq_Q)$  be two posets. Let the quadruple  $(P, f, g, Q)$  denote the pair of maps  $f : P \rightarrow Q$  and  $g : Q \rightarrow P$ .

**2.1 Definition.** The pair of maps  $(P, f, g, Q)$  is a (*regular*) *Galois connection* between  $P$  and  $Q$ , in notation  $\mathbb{G}(P, f, g, Q)$ , if

$$\forall p \in P \forall q \in Q (f(p) \leq_Q q \Leftrightarrow p \leq_P g(q)).$$

The map  $f$  is called the *lower adjoint* and  $g$  is called the *upper adjoint* of the Galois connection.

If  $P = Q$ ,  $\mathbb{G}(P, f, g, P)$  is said a *Galois connection on  $P$* .

The following theorem gives a useful characterization of Galois connections.

**2.2 Proposition.** ([32], Lemma 79) *The pair of maps  $(P, f, g, Q)$  is a Galois connection if and only if*

(1)  $p \leq_P g(f(p))$  for all  $p \in P$  and  $f(g(q)) \leq_Q q$  for all  $q \in Q$ ;

(2) the maps  $f$  and  $g$  are monotone.

**2.3 Remark.** Here we adopted the definition of Galois connection in which the maps are monotone. It is also called monotone or covariant form. For more details on Galois connections, see, e.g. [21, 22, 28].

Finally, we will need the following notion.

**2.4 Definition.** ([46], **Definition 2.2.2**) A pair of maps  $(P, f, g, Q)$  is the *partial Galois connection* between  $P$  and  $Q$ , denoted by  $\partial\mathbb{G}(P, f, g, Q)$ , if

- (1)  $f : P \rightarrow Q$  is a monotone partial map,
- (2)  $g : Q \rightarrow P$  is a monotone total map,
- (3)  $f(g(q))$  exists for all  $q \in Q$ , and
- (4)  $\forall p \in P$  and  $\forall q \in Q$  such that  $f(p)$  is defined,  $f(p) \leq_Q q \Leftrightarrow p \leq_P g(q)$ .

**2.5 Remark.** In [46], A. Miné actually introduced the concept of  $\mathcal{F}$ -partial Galois connection  $\partial\mathbb{G}(P, f, g, Q)$  between the concrete domain  $P$  and the abstract domain  $Q$ , where  $\mathcal{F}$  is a set of concrete operators. We will apply this notion in the simplest form:  $P = Q = 2^U$  and  $\mathcal{F} = \emptyset$  which is allowed by Miné's definition.





## Chapter 3

# General Approximation Frameworks

In order to be able to discuss the common features of both the rough set theory and its possible generalizations uniformly, we define two general approximation frameworks, a large-scaled and a finer-scaled set theoretic one.

### 3.1 An Initial Approximation Framework

A large-scaled general framework of the set approximation first has been proposed by the author in [14].

Let  $U$  be a nonempty set and  $\langle l, u \rangle$  be an ordered pair of maps

$$l, u : 2^U \rightarrow 2^U$$

on  $(2^U, \subseteq)$ . Of course, the maps  $l$  and  $u$  are intended to be the *lower* and *upper approximations* of any subset  $X \subseteq U$ , respectively. Hence, the ordered pair  $\langle l, u \rangle$  is called the *approximation pair*.

The most essential features of an approximation pair  $\langle l, u \rangle$  can be summarized as follows.

- (0) (*Definability*) The subsets of a set are approximated by the beforehand given family of subsets of the set itself. The members of the beforehand given family of subsets are called *well defined*. In other words, the maps  $l$  and  $u$  are of the form

$$l, u : 2^U \rightarrow \mathfrak{D}(\subseteq 2^U),$$

where  $\mathfrak{D}$  is the family of well defined subsets of  $U$ .

Hereupon, the nature of an approximation pair  $\langle l, u \rangle$  depends on how the lower and upper approximations are related to each other and the subset itself to be approximated.

- (1) (*Monotonicity*) The maps  $l$  and  $u$  are monotone with respect to the inclusion relation  $\subseteq$  on  $2^U$ .
- (2) (*Weak approximation property*) An approximation pair  $\langle l, u \rangle$  is the *weak approximation pair* on  $U$  if

$$\forall X \in 2^U \ (l(X) \subseteq u(X)).$$

- (3) (*Strong approximation property*) An approximation pair  $\langle l, u \rangle$  is the *strong approximation pair* on  $U$ , if each subset  $X \in 2^U$  is bounded by  $l(X)$  and  $u(X)$ :

$$\forall X \in 2^U \ (l(X) \subseteq X \subseteq u(X)).$$

- (4) (*Approximation hypothesis*) The pair of maps  $(2^U, u, l, 2^U)$  forms a *Galois connection* on  $(2^U, \subseteq)$ , in notation  $\mathbb{G}(2^U, u, l, 2^U)$ , if

$$\forall X \in 2^U \ \forall Y \in 2^U \ (u(X) \subseteq Y \Leftrightarrow X \subseteq l(Y)).$$

**3.1 Remark.** *Ad (0).* It gives the most fundamental characterization of the approximation pair  $\langle l, u \rangle$ .

*Ad (1).* This property is a common and reasonable assumption.

*Ad (2).* The constraint  $l(X) \subseteq u(X)$  seems to be the weakest condition for a sensible concept of set approximation [11, 25].

*Ad (3).* This property is meaningful because the domain and codomain of  $l, u$  are the same [11].

*Ad (4).* In [51], a new hypothesis about approximation has been drawn up recently. According to this assumption, the notion of the “approximation” may be mathematically modelled by the notion of the Galois connection.

A finer-scaled characterization of the nature of the set approximation can be obtained with *further specifications concerning the family of well defined subsets*. These additional specifications will be performed in the next Section.

### 3.2 A General Set Theoretic Approximation Framework

Let  $U$  be an arbitrary nonempty set called the *universe of discourse*.

The first definition gives us the family of fundamental sets of the framework which can be considered as *primary tools*.

**3.2 Definition.** Let  $\mathfrak{B} = \{B_i \mid i \in I\} \subseteq 2^U$  be a nonempty family of nonempty subsets of  $U$ , where  $I$  denotes an index set.

$\mathfrak{B}$  is called the *base system*, its members are the  $\mathfrak{B}$ -sets.

Some extensions of the base system  $\mathfrak{B}$  can be defined.

**3.3 Definition.** Let  $\mathfrak{D}_{\mathfrak{B}} \subseteq 2^U$  be an extension of  $\mathfrak{B}$  such that

- (1)  $\mathfrak{B} \subseteq \mathfrak{D}_{\mathfrak{B}}$ ;
- (2)  $\emptyset \in \mathfrak{D}_{\mathfrak{B}}$ .

The members of  $\mathfrak{D}_{\mathfrak{B}}$  are called *definable*, while the members of  $2^U \setminus \mathfrak{D}_{\mathfrak{B}}$  are *undefinable*.

Any extension  $\mathfrak{D}_{\mathfrak{B}}$  of  $\mathfrak{B}$  can be seen as *derived tools*.

**3.4 Example.** The simplest extension of  $\mathfrak{B}$  is  $\mathfrak{D}_{\mathfrak{B}} = \mathfrak{B} \cup \{\emptyset\}$ .

**3.5 Example.** Let  $\mathfrak{D}_{\mathfrak{B}} \subseteq 2^U$  be an extension of  $\mathfrak{B}$  such that

- (1)  $\emptyset \in \mathfrak{D}_{\mathfrak{B}}$ ;
- (2) for any index set  $I' \subseteq I$ , if  $\mathfrak{B}' = \{B_i \mid i \in I'\} \subseteq \mathfrak{B}$ , then  $\bigcup \mathfrak{B}' \in \mathfrak{D}_{\mathfrak{B}}$ .

Notice that  $\mathfrak{B} \subseteq \mathfrak{D}_{\mathfrak{B}}$ , and  $\mathfrak{D}_{\mathfrak{B}}$  is closed under arbitrary unions.

If the universe  $U$  is finite, and  $\mathfrak{B} = U/\varepsilon$ , where  $U/\varepsilon$  is a partition of  $U$  generated by an equivalence relation  $\varepsilon$  on  $U$ , then  $\mathfrak{D}_{\mathfrak{B}} = \sigma(U/\varepsilon)$ . In this case, this extension procedure is just the scheme which is in Pawlak's rough set theory.

**3.6 Example.** Let  $\mathfrak{D}_{\mathfrak{B}} \subseteq 2^U$  be an extension of  $\mathfrak{B}$  such that

- (1)  $\emptyset \in \mathfrak{D}_{\mathfrak{B}}$ ;

(2) if  $B_1, B_2 \in \mathfrak{B}$  then

- (a)  $B_1 \cup B_2 \in \mathfrak{D}_{\mathfrak{B}}$ ;
- (b)  $B_1 \cap B_2 \in \mathfrak{D}_{\mathfrak{B}}$ .

Notice that  $\mathfrak{B} \subseteq \mathfrak{D}_{\mathfrak{B}}$ ,  $\mathfrak{D}_{\mathfrak{B}}$  is closed under finite unions and intersections, and  $\bigcup \mathfrak{B}, \bigcap \mathfrak{B} \in \mathfrak{D}_{\mathfrak{B}}$  do not hold necessarily when the cardinality of  $I$  is not finite.

**3.7 Example.** If  $\sigma(\mathfrak{B})$  is the  $\sigma$ -algebra generated by  $\mathfrak{B}$  then  $\sigma(\mathfrak{B})$  is an extension of  $\mathfrak{B}$  since  $\emptyset \in \mathfrak{B}$  and  $\mathfrak{B} \subseteq \sigma(\mathfrak{B})$ .

We want to approximate of any subset  $S \in 2^U$  from “lower side” and “upper side”—no matter what they mean at this time. We have the only requirement at the highest level of abstraction that is to let the lower and upper approximations of subsets  $S$  be *definable*. We look at definable sets as *tools* to approximate subsets of the universe  $U$ .

If we look at the sets belonging to  $\mathfrak{B}$  as primary tools, it is a highly reasonable requirement that they should *exactly* be approximated by themselves from “lower side”. This property is called the *(lower) granularity* of  $\mathfrak{B}$ . If we gave it up, the roles of the primary tools would be depreciated.

In Pawlak’s rough set theory, however, not merely the granularity of  $U/\varepsilon$  but also the granularity of  $\sigma(U/\varepsilon)$  fulfills. It can be proved (Proposition 4.7, Corollary 4.9) that if  $D \in \sigma(U/\varepsilon)$ , then  $\underline{\varepsilon}(D) = D$  due to the particular construction of  $\mathfrak{D}_{U/\varepsilon}$  and definition of  $\underline{\varepsilon}$ .

A lower approximation is called *standard* if not only the primary tools in  $\mathfrak{B}$ , but also the derived tools in  $\mathfrak{D}_{\mathfrak{B}}$  are its fixpoints. In this thesis, we solely deal with standard lower approximations.

The following definition, at the next level of abstraction, is about the *minimum requirements* of standard lower and upper approximations.

**3.8 Definition.** Let  $\langle l, u \rangle$  be an approximation pair  $l, u : 2^U \rightarrow 2^U$  on  $(2^U, \subseteq)$ .

It is said that an approximation pair  $\langle l, u \rangle$  is the *weak (generalized) approximation pair on  $U$*  if

- (C0)  $l(2^U), u(2^U) \subseteq \mathfrak{D}_{\mathfrak{B}}$  (*definability* of  $l$  and  $u$ );
- (C1)  $l$  and  $u$  are monotone (*monotonicity* of  $l$  and  $u$ );

- (C2)  $u(\emptyset) = \emptyset$  (*normality of  $u$* );
- (C3) if  $D \in \mathfrak{D}_{\mathfrak{B}}$ , then  $l(D) = D$  (*granularity of  $\mathfrak{D}_{\mathfrak{B}}$* , i.e.  $l$  is standard);
- (C4) if  $S \in 2^U$ , then  $l(S) \subseteq u(S)$  (*approximation property*).

Informally, the intended meaning of the maps  $l$  and  $u$ , of course, is to express the lower and upper approximations of any subset of the universe  $U$  with the help of the beforehand given definable sets as *tools*.

Clearly, if  $\langle l, u \rangle$  is a weak approximation pair on  $U$ , the maps  $l, u$  are total and many-to-one in general.

**3.9 Proposition.** *Let  $\langle l, u \rangle$  be a weak approximation pair on  $U$ .*

- (1)  $l(\emptyset) = \emptyset$  (*normality of  $l$* );
- (2)  $\forall X \in 2^U (l(l(X)) = l(X))$  (*idempotency of  $l$* ).
- (3)  $S \in \mathfrak{D}_{\mathfrak{B}}$  if and only if  $l(S) = S$ .
- (4)  $u(2^U) \subseteq l(2^U) = \mathfrak{D}_{\mathfrak{B}}$ .

**Proof.** (1) By definition,  $\emptyset \in \mathfrak{D}_{\mathfrak{B}}$  and so  $l(\emptyset) = \emptyset$  by condition (C3).

(2)  $l(X) \in \mathfrak{D}_{\mathfrak{B}}$  and so  $l(l(X)) = l(X)$  by condition (C3).

(3)  $(\Rightarrow)$  It is just the same as the condition (C3).

$(\Leftarrow)$  Since  $l(S) \in \mathfrak{D}_{\mathfrak{B}}$ , and so  $l(S) = S \in \mathfrak{D}_{\mathfrak{B}}$  by Condition (C0).

(4)  $l(2^U) \subseteq \mathfrak{D}_{\mathfrak{B}}$  by condition (C0) and  $\mathfrak{D}_{\mathfrak{B}} \subseteq l(2^U)$  by condition (C3), thus  $l(2^U) = \mathfrak{D}_{\mathfrak{B}}$ .

Let  $S \in u(2^U) \subseteq \mathfrak{D}_{\mathfrak{B}}$ . By the condition (C3),  $S = l(S) \in \mathfrak{D}_{\mathfrak{B}} = l(2^U)$ , i.e.  $u(2^U) \subseteq l(2^U)$ .

To show that the inclusion  $u(2^U) \subseteq l(2^U)$  may be proper, let

- $U = \{a, b\}$ ,
- $\mathfrak{B} = \{\{a\}\}$ ,
- $\mathfrak{D}_{\mathfrak{B}} = \{\emptyset, \{a\}, \{a, b\}\}$ ,

- and  $l, u : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}$  be as follows:

$$X \mapsto l(X) = \begin{cases} \emptyset, & \text{if } X = \emptyset; \\ \{a\}, & \text{if } X = \{a\}; \\ \{a, b\}, & \text{otherwise.} \end{cases}$$

$$X \mapsto u(X) = \begin{cases} \emptyset, & \text{if } X = \emptyset; \\ \{a, b\}, & \text{otherwise.} \end{cases}$$

Conditions (C0)–(C4) can easily be checked, however,  $u(2^U) = \{\emptyset, \{a, b\}\} \subsetneq \{\emptyset, \{a\}, \{a, b\}\} = l(2^U) = \mathfrak{D}_{\mathfrak{B}}$ .  $\square$

The following example shows that for a weak approximation pair  $\langle l, u \rangle$  on  $U$  each condition (C0)–(C4) is independent of the other four.

**3.10 Example.** Let  $U$  be a nonempty set. Let us assume that there exist  $B_1, B_2 (\neq \emptyset) \in 2^U$  such that neither  $B_1 \subseteq B_2$  nor  $B_2 \subseteq B_1$  holds, and there exists a proper superset  $S$  of  $B_1$  (i.e.  $\emptyset \neq B_1 \subsetneq S \neq U$ ).

- (0) Let  $\mathfrak{B} = \{B_1\}$ ,  $\mathfrak{D}_{\mathfrak{B}} = \{\emptyset, B_1\}$  and  $l, u$  be the identity map, i.e.  $l, u : 2^U \rightarrow 2^U, X \mapsto X$ . These  $l$  and  $u$  trivially satisfy all the five conditions except (C0).
- (1) Let  $\mathfrak{B} = \{B_1, B_2\}$ ,  $\mathfrak{D}_{\mathfrak{B}} = \{\emptyset, B_1, B_2, B_1 \cup B_2\}$  and  $l, u : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}$  be as follows:

$$X \mapsto l(X) = \begin{cases} B_1, & \text{if } X = B_1; \\ B_2, & \text{if } X = B_2; \\ B_1 \cup B_2, & \text{if } X = B_1 \cup B_2, U; \\ \emptyset, & \text{otherwise.} \end{cases}$$

$$X \mapsto u(X) = \begin{cases} \emptyset, & \text{if } X = \emptyset; \\ B_1, & \text{if } X = B_1; \\ B_1 \cup B_2, & \text{if } X = B_1 \cup B_2, U; \\ B_2, & \text{otherwise.} \end{cases}$$

Conditions (C0), (C2), (C3) trivially hold. Let us check the condition

(C4):

$$\begin{aligned}
l(\emptyset) = \emptyset &\subseteq \emptyset = u(\emptyset) \\
l(B_1) = B_1 &\subseteq B_1 = u(B_1) \\
l(B_2) = B_2 &\subseteq B_2 = u(B_2) \\
l(B_1 \cup B_2) = B_1 \cup B_2 &\subseteq B_1 \cup B_2 = u(B_1 \cup B_2) \\
l(U) = B_1 \cup B_2 &\subseteq B_1 \cup B_2 = u(U) \\
l(S) = \emptyset &\subseteq B_2 = u(S)
\end{aligned}$$

and if  $S'(\neq \emptyset, B_1, B_2, B_1 \cup B_2, S, U) \in 2^U$ , then

$$l(S') = \emptyset \subseteq B_2 = u(S').$$

That is the condition (C4) also holds. However, in the case  $B_1 \subsetneq S$

$$\begin{aligned}
l(B_1) = B_1 &\not\subseteq \emptyset = l(S) \\
u(B_1) = B_1 &\not\subseteq B_2 = u(S).
\end{aligned}$$

Therefore, these  $l$  and  $u$  satisfy all the five conditions except (C1).

- (2) Let  $\mathfrak{B} = \{B_1, B_2\}$ ,  $\mathfrak{D}_{\mathfrak{B}} = \{\emptyset, B_1, B_2, B_1 \cup B_2\}$  and  $l, u : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}$  be as follows:

$$X \mapsto l(X) = \begin{cases} \emptyset, & \text{if } X = \emptyset; \\ B_1, & \text{if } X = B_1; \\ B_2, & \text{if } X = B_2; \\ B_1 \cup B_2, & \text{otherwise.} \end{cases}$$

$$X \mapsto u(X) = B_1 \cup B_2.$$

Conditions (C0), (C1), (C3) and (C4) hold, but  $u(\emptyset) = B_1 \cup B_2 \neq \emptyset$ .

Therefore, these  $l$  and  $u$  satisfy all the five conditions except (C2).

- (3) Let  $\mathfrak{B} = \{B_2\}$ ,  $\mathfrak{D}_{\mathfrak{B}} = \{\emptyset, B_2, B_1 \cup B_2\}$  and  $l, u : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}$  be as follows:

$$X \mapsto l(X) = \begin{cases} \emptyset, & \text{if } X = \emptyset; \\ B_2, & \text{if } X = B_1; \\ B_1 \cup B_2, & \text{otherwise.} \end{cases}$$

$$X \mapsto u(X) = \begin{cases} \emptyset, & \text{if } X = \emptyset; \\ B_1 \cup B_2, & \text{otherwise.} \end{cases}$$

Conditions (C0), (C1), (C2) trivially hold. Let us check the condition (C4):

$$\begin{aligned} l(\emptyset) = \emptyset &\subseteq \emptyset = u(\emptyset) \\ l(B_1) = B_2 &\subseteq B_1 \cup B_2 = u(B_1) \end{aligned}$$

and if  $S'(\neq \emptyset, B_1) \in 2^U$ , then

$$l(S') = B_1 \cup B_2 \subseteq B_1 \cup B_2 = u(S').$$

That is the condition (C4) also holds. However,

$$l(B_1) = B_2 \neq B_1.$$

Therefore, these  $l$  and  $u$  satisfy all the five conditions except (C3).

- (4) Let  $\mathfrak{B} = \{B_1, B_2\}$ ,  $\mathfrak{D}_{\mathfrak{B}} = \{\emptyset, B_1, B_2, B_1 \cup B_2\}$  and  $l, u : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}$  be as follows:

$$X \mapsto l(X) = \begin{cases} \emptyset, & \text{if } X = \emptyset; \\ B_1, & \text{if } X = B_1; \\ B_2, & \text{if } X = B_2; \\ B_1 \cup B_2, & \text{otherwise.} \end{cases}$$

$$X \mapsto u(X) = \emptyset.$$

These  $l$  and  $u$  trivially satisfy all the five conditions except (C4).

The next definition classifies the approximation pairs as how the lower and upper approximations of a subset are related to the subset itself to be approximated.

**3.11 Definition.** Let  $\langle l, u \rangle$  be an approximation pair  $l, u : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}$ .

It is said that an approximation pair  $\langle l, u \rangle$  is

- (C5) the *l-semi-strong approximation pair on U* if it is weak and if  $S \in 2^U$ , then  $l(S) \subseteq S$  ( $l$  is *contractive*);
- (C6) the *u-semi-strong approximation pair on U* if it is weak and if  $S \in 2^U$ , then  $S \subseteq u(S)$  ( $u$  is *extensive*);
- (C7) the *strong approximation pair on U* if it is  $l$ -semi-strong and  $u$ -semi-strong at the same time, i.e. each subset  $S \in 2^U$  is bounded by  $l(S)$  and  $u(S)$ :  $\forall S \in 2^U$  ( $l(S) \subseteq S \subseteq u(S)$ ).



If  $U$  is a nonempty set, and  $\mathfrak{D}_{\mathfrak{B}} = 2^U$ , it is straightforward that the approximation pair  $l, u : 2^U \rightarrow 2^U$ ,  $X \mapsto X$  is a strong approximation pair.

The next example shows that there are weak approximation pairs which are neither  $l$ -semi-strong nor  $u$ -semi-strong, not  $l$ -semi-strong but  $u$ -semi-strong,  $l$ -semi-strong but not  $u$ -semi-strong.

**3.12 Example.** Let  $U = \{a, b\}$  and  $\mathfrak{B} = \{\{a\}\}$  be the base system.

(1) Let  $\mathfrak{D}_{\mathfrak{B}} = \{\emptyset, \{a\}\}$ , and the maps  $l, u : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}$  be as follows:

$$X \mapsto l(X), u(X) = \begin{cases} \emptyset, & \text{if } X = \emptyset; \\ \{a\}, & \text{otherwise.} \end{cases}$$

Conditions (C0)–(C4) can easily be checked:

$$(C0) \quad l(2^U), u(2^U) = \{\emptyset, \{a\}\} = \mathfrak{D}_{\mathfrak{B}}.$$

(C1)  $l$  is monotone:

$$\begin{aligned} \emptyset \subset \{a\}, \{b\}, \{a, b\} &\Rightarrow l(\emptyset) = \emptyset \subset \{a\} = l(\{a\}), l(\{b\}), l(\{a, b\}). \\ \{a\}, \{b\} \subset \{a, b\} &\Rightarrow l(\{a\}), l(\{b\}) = \{a\} \subseteq \{a\} = l(\{a, b\}). \end{aligned}$$

The monotonicity of  $u$  can be proved in the same way.

$$(C2) \quad u(\emptyset) = \emptyset.$$

$$(C3) \quad l(\emptyset) = \emptyset, l(\{a\}) = \{a\}.$$

$$(C4) \quad l \text{ and } u \text{ are the same maps.}$$

However, for  $X = \{b\}$

$$l(\{b\}) = \{a\} \not\subseteq \{b\}; \{b\} \not\subseteq u(\{b\}) = \{a\}. \quad (3.2.1)$$

Therefore, the approximation pair  $\langle l, u \rangle$  is neither  $l$ -semi-strong nor  $u$ -semi-strong.

(2) Let  $\mathfrak{D}_{\mathfrak{B}} = \{\emptyset, \{a\}, \{a, b\}\}$ , and the maps  $l, u : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}$  be as follows:

$$X \mapsto l(X), u(X) = \begin{cases} \emptyset, & \text{if } X = \emptyset; \\ \{a\}, & \text{if } X = \{a\}; \\ \{a, b\}, & \text{otherwise.} \end{cases}$$

Conditions (C0)–(C4) can easily be checked:

$$(C0) \quad l(2^U), u(2^U) = \{\emptyset, \{a\}, \{a, b\}\} = \mathfrak{D}_{\mathfrak{B}}.$$

(C1)  $l$  is monotone:

$$\emptyset \subset \{a\} \Rightarrow l(\emptyset) = \emptyset \subset \{a\} = l(\{a\}).$$

$$\emptyset \subset \{a\}, \{a, b\} \Rightarrow l(\emptyset) = \emptyset \subset \{a, b\} = l(\{b\}), l(\{a, b\}).$$

$$\{a\} \subset \{a, b\} \Rightarrow l(\{a\}) = \{a\} \subset \{a, b\} = l(\{a, b\}).$$

$$\{b\} \subset \{a, b\} \Rightarrow l(\{b\}) = \{a, b\} \subseteq \{a, b\} = l(\{a, b\}).$$

The monotonicity of  $u$  can be proved in the same way.

(C2)  $u(\emptyset) = \emptyset$ .

(C3)  $l(\emptyset) = \emptyset, l(\{a\}) = \{a\}, l(\{a, b\}) = \{a, b\}$ .

(C4)  $l$  and  $u$  are the same maps.

Let us check that  $u$  is extensive:

- $\emptyset \subseteq \emptyset = u(\emptyset)$ ;
- $\{a\} \subseteq \{a\} = u(\{a\})$ ;
- $\{b\} \subseteq \{a, b\} = u(\{b\})$ ;
- $\{a, b\} \subseteq \{a, b\} = u(\{a, b\})$ .

However, in the case  $X = \{b\}$ ,

$$l(\{b\}) = \{a, b\} \not\subseteq \{b\}. \quad (3.2.2)$$

Therefore, the approximation pair  $\langle l, u \rangle$  is not  $l$ -semi-strong, but  $u$ -semi-strong.

(3) Let  $\mathfrak{D}_{\mathfrak{B}} = \{\emptyset, \{a\}, \{a, b\}\}$ , and the maps  $l, u : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}$  be as follows:

$$X \mapsto l(X), u(X) = \begin{cases} \emptyset, & \text{if } X = \emptyset, \{b\}; \\ \{a\}, & \text{if } X = \{a\}; \\ \{a, b\}, & \text{otherwise.} \end{cases}$$

Conditions (C0)–(C4) can easily be checked:

(C0)  $l(2^U), u(2^U) = \{\emptyset, \{a\}, \{a, b\}\} = \mathfrak{D}_{\mathfrak{B}}$ .

(C1)  $l$  is monotone:

$$\emptyset \subset \{a\} \Rightarrow l(\emptyset) = \emptyset \subset \{a\} = l(\{a\}),$$

$$\emptyset \subset \{b\} \Rightarrow l(\emptyset) = \emptyset \subseteq \emptyset = l(\{b\}),$$

$$\emptyset \subset \{a, b\} \Rightarrow l(\emptyset) = \emptyset \subset \{a, b\} = l(\{a, b\}),$$

$$\{a\} \subset \{a, b\} \Rightarrow l(\{a\}) = \{a\} \subset \{a, b\} = l(\{a, b\}),$$

$$\{b\} \subset \{a, b\} \Rightarrow l(\{b\}) = \emptyset \subset \{a, b\} = l(\{a, b\}).$$

The monotonicity of  $u$  can be proved in the same way.

$$(C2) \quad u(\emptyset) = \emptyset.$$

$$(C3) \quad l(\emptyset) = \emptyset, l(\{a\}) = \{a\}, l(\{a, b\}) = \{a, b\}.$$

$$(C4) \quad l \text{ and } u \text{ are the same maps.}$$

Let us check that  $l$  is contractive:

- $l(\emptyset) = \emptyset \subseteq \emptyset$ ;
- $l(\{a\}) = \{a\} \subseteq \{a\}$ ;
- $l(\{b\}) = \emptyset \subset \{b\}$ ;
- $l(\{a, b\}) = \{a, b\} \subseteq \{a, b\}$ .

However, in the case  $X = \{b\}$ ,

$$\{b\} \not\subseteq \emptyset = u(\{b\}). \quad (3.2.3)$$

Therefore, the approximation pair  $\langle l, u \rangle$  is  $l$ -semi-strong, but not  $u$ -semi-strong.

Using the preliminary notations, the notion of the generalized approximation space can be defined.

**3.13 Definition.** An ordered quadruple  $\langle U, \mathfrak{D}_{\mathfrak{B}}, l, u \rangle$  is the *weak/l-semi-strong/u-semi-strong/strong* generalized approximation space, if the approximation pair  $\langle l, u \rangle$  is weak/ $l$ -semi-strong/ $u$ -semi-strong/strong, respectively.

**3.14 Proposition.** Let  $\langle U, \mathfrak{D}_{\mathfrak{B}}, l, u \rangle$  be a generalized approximation space.

(1) If  $\langle U, \mathfrak{D}_{\mathfrak{B}}, l, u \rangle$  is weak, then

- (a)  $l(U) \subseteq \bigcup \mathfrak{D}_{\mathfrak{B}}$ ;
- (b)  $l(U) = \bigcup \mathfrak{D}_{\mathfrak{B}}$  if and only if  $\bigcup \mathfrak{D}_{\mathfrak{B}} \in \mathfrak{D}_{\mathfrak{B}}$ .
- (c)  $u(U) \subseteq \bigcup \mathfrak{D}_{\mathfrak{B}}$ .

(2) If  $\langle U, \mathfrak{D}_{\mathfrak{B}}, l, u \rangle$  is  $u$ -semi-strong, then  $u(U) = \bigcup \mathfrak{D}_{\mathfrak{B}} = U$ .

**Proof.**

- (1) (a) By the definition of  $l$ ,  $l(U) \in \mathfrak{D}_{\mathfrak{B}}$  and so  $l(U) \subseteq \bigcup \mathfrak{D}_{\mathfrak{B}}$ .
- (b)  $(\Rightarrow)$  By the definition of  $l$ ,  $l(U) = \bigcup \mathfrak{D}_{\mathfrak{B}} \in \mathfrak{D}_{\mathfrak{B}}$ .  
 $(\Leftarrow)$  Let us assume that  $\bigcup \mathfrak{D}_{\mathfrak{B}} \in \mathfrak{D}_{\mathfrak{B}}$ . Since  $\bigcup \mathfrak{D}_{\mathfrak{B}} \subseteq U$ , then by the condition (C3) and the monotonicity of  $l$ ,  $l(\bigcup \mathfrak{D}_{\mathfrak{B}}) = \bigcup \mathfrak{D}_{\mathfrak{B}} \subseteq l(U)$ . Comparing it with (1) (a), we obtain  $l(U) = \bigcup \mathfrak{D}_{\mathfrak{B}}$ .

(c) By the definitions of  $u$ ,  $u(U) \in \mathfrak{D}_{\mathfrak{B}}$  and so  $u(U) \subseteq \bigcup \mathfrak{D}_{\mathfrak{B}}$ .

(2)  $\langle U, \mathfrak{D}_{\mathfrak{B}}, l, u \rangle$  is weak, thus by Proposition 3.14, point (1)/(c),  $u(U) \subseteq \bigcup \mathfrak{D}_{\mathfrak{B}}$ . On the other hand, since  $u$  is extensive and monotone,  $\bigcup \mathfrak{D}_{\mathfrak{B}} \subseteq U$  implies  $\bigcup \mathfrak{D}_{\mathfrak{B}} \subseteq u(\bigcup \mathfrak{D}_{\mathfrak{B}}) \subseteq u(U)$ . Consequently,  $u(U) = \bigcup \mathfrak{D}_{\mathfrak{B}}$ .

Clearly,  $u(U) \subseteq U$ . Since  $u$  is extensive, thus  $U \subseteq u(U)$ . Therefore,  $u(U) = U$ .  $\square$

In generalized approximation spaces the notion of *well approximated sets* can be introduced. These sets are called crisp sets.

**3.15 Definition.** Let  $\langle U, \mathfrak{D}_{\mathfrak{B}}, l, u \rangle$  be a generalized approximation space and  $S \in 2^U$ .

The subset  $S$  is *crisp*, if  $l(S) = u(S)$ .

**3.16 Proposition.** Let  $\langle U, \mathfrak{D}_{\mathfrak{B}}, l, u \rangle$  be a strong generalized approximation space.

If  $S \in 2^U$  is *crisp*, then  $S$  is *definable*.

**Proof.**  $\langle U, \mathfrak{D}_{\mathfrak{B}}, l, u \rangle$  is strong, thus  $l(S) \subseteq S \subseteq u(S)$ . Since  $S$  is crisp, therefore  $l(S) = S = u(S)$ , and so  $S \in \mathfrak{D}_{\mathfrak{B}}$  by Proposition 3.9 (3).  $\square$

In general, the crisp property of a set does not imply its definability in not strong generalized approximation spaces. One can check that in all three cases of Example 3.12, the set  $\{b\}$  is crisp (because of  $l$  and  $u$  are the same maps, and so  $l(\{b\}) = u(\{b\})$  trivially holds), but  $\{b\}$  is not definable (i.e.  $\{b\} \notin \mathfrak{D}_{\mathfrak{B}}$ ). Of course, its lower and upper approximations are definable (i.e.  $l(\{b\}), u(\{b\}) \in \mathfrak{D}_{\mathfrak{B}}$ ).

## Chapter 4

# Fundamentals of Rough Set Theory

The basic concepts and properties of rough set theory can be found, e.g. in [33, 54, 55]. Here we will cite only notions and statements which are required in our subsequent work. Moreover, we partly restate these well-known facts in the language of the set theoretic approximation framework. On the other hand, we provide new point-free proofs for a few of them (see, especially, Section 4.2).

### 4.1 Basic Notions

Let  $U$  be a nonempty set and  $\varepsilon$  be an equivalence relation on  $U$ . In Pawlak's rough set theory the base system is the partition  $U/\varepsilon$ . Its extension  $\mathfrak{D}_{U/\varepsilon}$  contains  $U/\varepsilon$ , the empty set and closed under arbitrary unions. The members of  $\mathfrak{D}_{U/\varepsilon}$  are called  $\varepsilon$ -definable, while the members of  $2^U \setminus \mathfrak{D}_{U/\varepsilon}$  are called  $\varepsilon$ -undefinable.

**4.1 Remark.** By the special structure of  $U/\varepsilon$ ,  $\mathfrak{D}_{U/\varepsilon}$  is nonempty, closed under arbitrary unions, intersections and complementations. In other words,  $(U, \mathfrak{D}_{U/\varepsilon})$  is an Alexandrov topological space with the basis  $U/\varepsilon$ .

Having given the definable sets, Pawlak's approximation pair  $\langle \underline{\varepsilon}, \bar{\varepsilon} \rangle$  can be defined in three equivalent forms [76, 77, 80] as follows.

**4.2 Definition.** Let  $\langle \underline{\varepsilon}, \bar{\varepsilon} \rangle$  be an approximation pair  $\underline{\varepsilon}, \bar{\varepsilon} : 2^U \rightarrow 2^U$  on  $(2^U, \subseteq)$ .  $\langle \underline{\varepsilon}, \bar{\varepsilon} \rangle$  is a *Pawlak's approximation pair on  $U$* , if

the *lower*  $\varepsilon$ -approximation of a subset  $X \in 2^U$  is

$$\underline{\varepsilon}(X) = \{x \in U \mid [x]_\varepsilon \subseteq X\}, \quad (4.1.1a)$$

$$= \bigcup \{Y \mid Y \in U/\varepsilon, Y \subseteq X\}, \quad (4.1.1b)$$

$$= \bigcup \{D \mid D \in \mathfrak{D}_{U/\varepsilon}, D \subseteq X\}, \quad (4.1.1c)$$

and the *upper*  $\varepsilon$ -approximation of a subset  $X \in 2^U$  is

$$\overline{\varepsilon}(X) = \{x \in U \mid [x]_\varepsilon \cap X \neq \emptyset\}, \quad (4.1.2a)$$

$$= \bigcup \{Y \mid Y \in U/\varepsilon, Y \cap X \neq \emptyset\}, \quad (4.1.2b)$$

$$= \bigcap \{D \mid D \in \mathfrak{D}_{U/\varepsilon}, X \subseteq D\}, \quad (4.1.2c)$$

**4.3 Remark.** The above equations respectively emphasize the *local* (4.1.1a, 4.1.2a), *global* (4.1.1b, 4.1.2b) and *topological* (4.1.1c, 4.1.2c) nature of Pawlak's approximations. From another point of view, the local approach is *point-wise* and the two latter ones are *point-free* in nature.

Our approach is relying on the generalization of formulae 4.1.1b and 4.1.2b when the definable sets are not pairwise disjoint and they do not necessarily cover the universe.

**4.4 Proposition.** Let  $\langle \underline{\varepsilon}, \overline{\varepsilon} \rangle$  be a Pawlak's approximation pair on  $U$ . Then

(1) the formulae 4.1.1b and 4.1.1c are equivalent, i.e.

$$\bigcup \{Y \mid Y \in U/\varepsilon, Y \subseteq X\} = \bigcup \{D \mid D \in \mathfrak{D}_{U/\varepsilon}, D \subseteq X\};$$

(2) the formulae 4.1.2b and 4.1.2c are equivalent, i.e.

$$\bigcup \{Y \mid Y \in U/\varepsilon, Y \cap X \neq \emptyset\} = \bigcap \{D \mid D \in \mathfrak{D}_{U/\varepsilon}, X \subseteq D\}.$$

**Proof.** (1) It follows from the fact that every  $D(\subseteq X) \in \mathfrak{D}_{U/\varepsilon}$  is of the form  $D = \bigcup \{Y \mid Y \in U/\varepsilon, Y \subseteq X\}$ .

(2) Since  $Y \cap D = Y$  or  $\emptyset$  for any  $Y \in U/\varepsilon$  and  $D \in \mathfrak{D}_{U/\varepsilon}$ , then

$$\bigcup \{Y \mid Y \in U/\varepsilon, Y \cap X \neq \emptyset\} \subseteq \bigcup \{Y \mid Y \in U/\varepsilon, Y \cap D \neq \emptyset\} = D$$

for any  $D \in \mathfrak{D}_{U/\varepsilon}$  where  $X \subseteq D$ .

In addition, by definition,  $X \subseteq \bigcup \{Y \mid Y \in U/\varepsilon, Y \cap X \neq \emptyset\} \in \mathfrak{D}_{U/\varepsilon}$ .  $\square$

**4.5 Remark.** In the equation 4.1.2c, contrary to the expectations, the formula  $\bigcap \{D \mid D \in \mathfrak{D}_{U/\varepsilon}, X \subseteq D\}$  cannot be replaced with the formula  $\bigcup \{D \mid D \in \mathfrak{D}_{U/\varepsilon}, X \cap D \neq \emptyset\}$  as the next example shows. Let

$$\begin{aligned} U &= \{x_1, x_2\}, \\ U/\varepsilon &= \{\{x_1\}, \{x_2\}\}, \\ \mathfrak{D}_{U/\varepsilon} &= \{\emptyset, \{x_1\}, \{x_2\}, \{x_1, x_2\}\}. \end{aligned}$$

Then

$$\bar{\varepsilon}(\{x_1\}) = \bigcup \{\{x_1\}\} = \{x_1\} \quad (\text{according to 4.1.2b}),$$

$$\bar{\varepsilon}(\{x_1\}) = \bigcap \{\{x_1\}, \{x_1, x_2\}\} = \{x_1\} \quad (\text{according to 4.1.2c}),$$

however, according to the formula  $\bigcup \{D \mid D \in \mathfrak{D}_{U/\varepsilon}, X \cap D \neq \emptyset\}$  we do not obtain the correct result:

$$\bar{\varepsilon}(\{x_1\}) \neq \bigcup \{\{x_1\}, \{x_1, x_2\}\} = \{x_1, x_2\}.$$

**4.6 Proposition.** Let  $\langle \underline{\varepsilon}, \bar{\varepsilon} \rangle$  be a Pawlak's approximation pair on  $U$ . Then

(0)  $\underline{\varepsilon}(2^U), \bar{\varepsilon}(2^U) \subseteq \mathfrak{D}_{U/\varepsilon}$  (definability of  $\underline{\varepsilon}$  and  $\bar{\varepsilon}$ ), and  $\underline{\varepsilon}, \bar{\varepsilon}$  are total and generally many-to-one;

(1) if  $X \subseteq Y$ , then  $\underline{\varepsilon}(X) \subseteq \underline{\varepsilon}(Y)$  and  $\bar{\varepsilon}(X) \subseteq \bar{\varepsilon}(Y)$  ( $\underline{\varepsilon}, \bar{\varepsilon}$  are monotone);

(2)  $\underline{\varepsilon}(\emptyset) = \bar{\varepsilon}(\emptyset) = \emptyset$  ( $\underline{\varepsilon}, \bar{\varepsilon}$  are normalized).

**Proof.** (0) It is straightforward by Definition 4.2.

(1) [55] Proposition 2.2 points 5), 6).

(2) [55] Proposition 2.2 point 2).  $\square$

According to Proposition 4.6 points (0), (1), (2), the rough set theory fulfills the conditions (C0), (C1), (C2) of the set theoretic approximation framework.

In the following we will investigate the fulfillment of the conditions (C3) and (C4) of the set theoretic approximation framework.

**4.7 Proposition.** ([55], Proposition 2.1, point a)) *Let  $\langle \underline{\varepsilon}, \bar{\varepsilon} \rangle$  be a Pawlak's approximation pair on  $U$ .*

*Then  $X \in \mathfrak{D}_{U/\varepsilon}$  if and only if  $\underline{\varepsilon}(X) = \bar{\varepsilon}(X)$ .*

**4.8 Corollary.** *If  $D \in \mathfrak{D}_{U/\varepsilon}$ , then  $\underline{\varepsilon}(D) = D$ .*

**Proof.** Since  $\underline{\varepsilon}(D) \in \mathfrak{D}_{U/\varepsilon}$ , then  $\underline{\varepsilon}(D) = D$  by Proposition 4.6 point 3) and Proposition 4.7.  $\square$

According to Corollary 4.8, the rough set theory fulfills the conditions (C3) of the set theoretic approximation framework.

The next statement is a characteristic feature of the rough set theory.

**4.9 Corollary.**  *$\underline{\varepsilon}(X) = X$  if and only if  $X = \bar{\varepsilon}(X)$ .*

**Proof.** Since  $\underline{\varepsilon}(X) \in \mathfrak{D}_{U/\varepsilon}$  ( $\bar{\varepsilon}(X) \in \mathfrak{D}_{U/\varepsilon}$ ), then  $X = \underline{\varepsilon}(X) \in \mathfrak{D}_{U/\varepsilon}$  ( $X = \bar{\varepsilon}(X) \in \mathfrak{D}_{U/\varepsilon}$ ) by Proposition 4.6 point 3) and Proposition 4.7, and so  $X = \underline{\varepsilon}(X) = \bar{\varepsilon}(X)$  ( $X = \bar{\varepsilon}(X) = \underline{\varepsilon}(X)$ ) by Proposition 4.7.  $\square$

**4.10 Proposition.** *Let  $\langle \underline{\varepsilon}, \bar{\varepsilon} \rangle$  be a Pawlak's approximation pair on  $U$ . Then*

$$\forall X \in 2^U (\underline{\varepsilon}(X) \subseteq X \subseteq \bar{\varepsilon}(X)) \quad (\underline{\varepsilon} \text{ is contractive, } \bar{\varepsilon} \text{ is extensive}).$$

**Proof.** [55] Proposition 2.2 point 1).  $\square$

According to Proposition 4.10, the rough set theory fulfills the condition (C4) of the set theoretic approximation framework. At the same time, this proposition also prove that the rough set theory fulfills the condition (C7), too.

Summing up the above results, in the language of set theoretic approximation framework, a Pawlak's approximation pair  $\langle \underline{\varepsilon}, \bar{\varepsilon} \rangle$  is a strong approximation pair. Consequently, the quadruple  $\langle U, \mathfrak{D}_{U/\varepsilon}, \underline{\varepsilon}, \bar{\varepsilon} \rangle$  forms a strong approximation space. It is also called *Pawlak's approximation space*.

**4.11 Remark.** Note that the idea of approximation *space* is a bit older than Pawlak's initial works. For the evolutionary survey of approximation spaces, see [68].

The next properties of  $\underline{\varepsilon}$  and  $\bar{\varepsilon}$  partly follows from Proposition 3.9. Of course, they can easily be proved by Definition 4.2 directly.



**4.12 Proposition.** Let  $\langle U, \mathfrak{D}_{U/\varepsilon}, \underline{\varepsilon}, \bar{\varepsilon} \rangle$  be a Pawlak's approximation space.

- (1)  $\underline{\varepsilon}(U) = \bar{\varepsilon}(U) = U$  ( $\underline{\varepsilon}, \bar{\varepsilon}$  are co-normalized);
- (2)  $\forall X \in 2^U$  ( $\underline{\varepsilon}(\underline{\varepsilon}(X)) = \underline{\varepsilon}(X) \wedge \bar{\varepsilon}(\bar{\varepsilon}(X)) = \bar{\varepsilon}(X)$ ) ( $\underline{\varepsilon}, \bar{\varepsilon}$  are idempotent);
- (3)  $\underline{\varepsilon}(2^U) = \bar{\varepsilon}(2^U) = \mathfrak{D}_{U/\varepsilon}$ .

**Proof.** (1) [55] Proposition 2.2 point 2).

(2) [55] Proposition 2.2 points 11), 12).

(3) It follows from Proposition 4.7 □

**4.13 Definition.** Let  $\langle U, \mathfrak{D}_{U/\varepsilon}, \underline{\varepsilon}, \bar{\varepsilon} \rangle$  be a Pawlak's approximation space and  $X \subseteq U$ .

The  $\varepsilon$ -boundary of  $X$  is

$$B_\varepsilon(X) = \bar{\varepsilon}(X) \setminus \underline{\varepsilon}(X).$$

$X$  is  $\varepsilon$ -crisp, if  $B_\varepsilon(X) = \emptyset$ , otherwise  $X$  is  $\varepsilon$ -rough.

**4.14 Proposition. (Csajbók [16], Proposition 4.14)** Let  $\langle U, \mathfrak{D}_{U/\varepsilon}, \underline{\varepsilon}, \bar{\varepsilon} \rangle$  be a Pawlak's approximation space and  $X \subseteq U$ .

- (1)  $X$  is  $\varepsilon$ -crisp if and only if  $X$  is  $\varepsilon$ -definable.
- (2)  $X$  is  $\varepsilon$ -rough if and only if  $X$  is  $\varepsilon$ -undefinable.

**Proof.** (1)  $(\Rightarrow)$   $X$  is  $\varepsilon$ -crisp  $\Leftrightarrow B_\varepsilon(X) = \bar{\varepsilon}(X) \setminus \underline{\varepsilon}(X) = \emptyset \Leftrightarrow \bar{\varepsilon}(X) \subseteq \underline{\varepsilon}(X)$ . However,  $\underline{\varepsilon}(X) \subseteq \bar{\varepsilon}(X)$ , and so  $\underline{\varepsilon}(X) = \bar{\varepsilon}(X)$ . According to Proposition 4.7,  $\underline{\varepsilon}(X) = \bar{\varepsilon}(X) \Leftrightarrow X \in \mathfrak{D}_{U/\varepsilon}$ .

$(\Leftarrow)$  Since  $X \in \mathfrak{D}_{U/\varepsilon} \Leftrightarrow \underline{\varepsilon}(X) = \bar{\varepsilon}(X)$ , then  $B_\varepsilon(X) = \bar{\varepsilon}(X) \setminus \underline{\varepsilon}(X) = \emptyset$  trivially satisfies.

(2) It is the contrapositive version of (1). □

As a consequence of Proposition 4.14, in Pawlak's approximation spaces the notions ' $\varepsilon$ -crisp' and ' $\varepsilon$ -definable' are synonymous to each other, and so are ' $\varepsilon$ -rough' and ' $\varepsilon$ -undefinable'. However, the notions ' $\varepsilon$ -crisp' and ' $\varepsilon$ -definable' are two *different* notions, they are inherently one and the same only in Pawlak's approximation spaces. As we will see, in partial approximation of sets this compound notion splits into two parts.

## 4.2 Granularity Aspects of Rough Set Theory

The following statement is elementary, however, in the context of Pawlak's rough set theory it is an important fact. For the sake of simple reference, it is formulated in a lemma. It follows from just the fact that the partition  $U/\varepsilon$  consists of nonempty pairwise disjoint subsets of  $U$ .

**4.15 Lemma.**  $\forall \mathfrak{X} \in 2^{U/\varepsilon} \forall X \in U/\varepsilon (X \subseteq \bigcup \mathfrak{X} \Leftrightarrow X \in \mathfrak{X})$ .

**4.16 Proposition.** (Csajbók [14], Theorem 8) *Let  $\langle U, \mathfrak{D}_{U/\varepsilon}, \underline{\varepsilon}, \bar{\varepsilon} \rangle$  be a Pawlak's approximation space.*

*Then the posets  $(2^{U/\varepsilon}, \subseteq)$  and  $(\mathfrak{D}_{U/\varepsilon}, \subseteq)$  are order isomorphic via the map  $i_\varepsilon : 2^{U/\varepsilon} \rightarrow \mathfrak{D}_{U/\varepsilon}, \mathfrak{X} \mapsto \bigcup \mathfrak{X}$ .*

**Proof.** We will show that the map  $i_\varepsilon$  is a bijection and both  $i_\varepsilon$  and  $i_\varepsilon^{-1}$  are monotone.

Let  $\mathfrak{D}_1, \mathfrak{D}_2 \in 2^{U/\varepsilon}$  be such that  $\bigcup \mathfrak{D}_1 = \bigcup \mathfrak{D}_2 \in \mathfrak{D}_{U/\varepsilon}$ . By Lemma 4.15,  $\forall X \in 2^U (X \in \mathfrak{D}_1 \Leftrightarrow X \subseteq \bigcup \mathfrak{D}_1 = \bigcup \mathfrak{D}_2 \Leftrightarrow X \in \mathfrak{D}_2)$ , i.e.,  $\mathfrak{D}_1 = \mathfrak{D}_2$ , thus  $i_\varepsilon$  is injective. By definition of  $\mathfrak{D}_{U/\varepsilon}$ ,  $i_\varepsilon$  is surjective. Consequently,  $i_\varepsilon$  is a bijection.

Clearly, the map  $i_\varepsilon$  is monotone, since  $\mathfrak{X}_1, \mathfrak{X}_2 \in 2^{U/\varepsilon}$ ,  $\mathfrak{X}_1 \subseteq \mathfrak{X}_2$  immediately implies  $\bigcup \mathfrak{X}_1 \subseteq \bigcup \mathfrak{X}_2$ .

Now, let  $D_1, D_2 \in \mathfrak{D}_{U/\varepsilon}$  be so that  $D_1 \subseteq D_2$ . Since  $i_\varepsilon$  is a bijection, there exist unique  $i_\varepsilon^{-1}(D_1) = \mathfrak{X}_1, i_\varepsilon^{-1}(D_2) = \mathfrak{X}_2 \in 2^{U/\varepsilon}$  so that  $D_1 = \bigcup \mathfrak{X}_1, D_2 = \bigcup \mathfrak{X}_2$ . By Lemma 4.15,  $\forall X \in 2^U (X \in \mathfrak{X}_1 \Leftrightarrow X \subseteq \bigcup \mathfrak{X}_1 \subseteq \bigcup \mathfrak{X}_2 \Leftrightarrow X \in \mathfrak{X}_2)$ , i.e.,  $\mathfrak{X}_1 \subseteq \mathfrak{X}_2$ , and so  $i_\varepsilon^{-1}$  is also monotone.  $\square$

**4.17 Corollary.** (Csajbók [16], Corollary 3.5) *Any  $\varepsilon$ -definable subset  $D$  of  $U$  can be written uniquely in the following form:*

$$D = \bigcup \mathfrak{X}, \text{ where } \mathfrak{X} = \{X \mid X \in U/\varepsilon, X \subseteq D\} \in 2^{U/\varepsilon},$$

*that is, there is no other  $\mathfrak{X}' \in 2^{U/\varepsilon}$  satisfying  $D = \bigcup \mathfrak{X}'$ .*

**Proof.** Since  $D \in \mathfrak{D}_{U/\varepsilon}$ , thus  $D = \bigcup \{X \mid X \in U/\varepsilon, X \subseteq D\}$  immediately holds. However,  $i_\varepsilon$  is a bijection, and so  $i_\varepsilon^{-1}(D)$  always exists and  $i_\varepsilon^{-1}(D) = \{X \mid X \in U/\varepsilon, X \subseteq D\} = \mathfrak{X} \in 2^{U/\varepsilon}$  is unique.  $\square$

**4.18 Proposition.** (Csajbók [16], Proposition 3.7) *Let  $\langle U, \mathfrak{D}_{U/\varepsilon}, \underline{\varepsilon}, \bar{\varepsilon} \rangle$  be a Pawlak's approximation space and  $X$  be a subset of  $U$ .*

Then the sets  $\underline{\varepsilon}(X)$ ,  $\bar{\varepsilon}(X)$  can be written uniquely in the following forms:

$$\begin{aligned}\underline{\varepsilon}(X) &= \bigcup \underline{\mathfrak{X}}, \text{ where } \underline{\mathfrak{X}} = \{Y \mid Y \in U/\varepsilon, Y \subseteq X\} \in 2^{U/\varepsilon}, \\ \bar{\varepsilon}(X) &= \bigcup \bar{\mathfrak{X}}, \text{ where } \bar{\mathfrak{X}} = \{Y \mid Y \in U/\varepsilon, Y \cap X \neq \emptyset\} \in 2^{U/\varepsilon},\end{aligned}$$

that is, there are no other  $\underline{\mathfrak{X}}', \bar{\mathfrak{X}}' \in 2^{U/\varepsilon}$  satisfying  $\underline{\varepsilon}(X) = \bigcup \underline{\mathfrak{X}}'$  and  $\bar{\varepsilon}(X) = \bigcup \bar{\mathfrak{X}}'$ .

**Proof.** According to Definition 4.2 equations (4.1.1b), (4.1.2b), we only have to prove the uniqueness.

$\underline{\varepsilon}(X)$ ,  $\bar{\varepsilon}(X) \in \mathfrak{D}_{U/\varepsilon}$ , and so, by Proposition 4.16,  $i_\varepsilon^{-1}(\underline{\varepsilon}(X))$  and  $i_\varepsilon^{-1}(\bar{\varepsilon}(X))$  are unique and, by Lemma 4.15, we get

$$\begin{aligned}i_\varepsilon^{-1}(\underline{\varepsilon}(X)) &= \{Y \mid Y \in U/\varepsilon, Y \subseteq \underline{\varepsilon}(X)\} \\ &= \{Y \mid Y \in U/\varepsilon, Y \subseteq \bigcup \{Y' \mid Y' \in U/\varepsilon, Y' \subseteq X\}\} \\ &= \{Y \mid Y \in U/\varepsilon, Y \in \{Y' \mid Y' \in U/\varepsilon, Y' \subseteq X\}\} \\ &= \{Y \mid Y \in U/\varepsilon, Y \subseteq X\} = \underline{\mathfrak{X}}. \\ i_\varepsilon^{-1}(\bar{\varepsilon}(X)) &= \{Y \mid Y \in U/\varepsilon, Y \subseteq \bar{\varepsilon}(X)\} \\ &= \{Y \mid Y \in U/\varepsilon, Y \subseteq \bigcup \{Y' \mid Y' \in U/\varepsilon, Y' \cap X \neq \emptyset\}\} \\ &= \{Y \mid Y \in U/\varepsilon, Y \in \bigcup \{Y' \mid Y' \in U/\varepsilon, Y' \cap X \neq \emptyset\}\} \\ &= \{Y \mid Y \in U/\varepsilon, Y \cap X \neq \emptyset\} = \bar{\mathfrak{X}}.\end{aligned}$$

□

### 4.3 Galois Connection of Upper and Lower Approximations

It is well known fact that  $\bar{\varepsilon}$  and  $\underline{\varepsilon}$  form a  $\mathbb{G}(2^U, \bar{\varepsilon}, \underline{\varepsilon}, 2^U)$  Galois connection. Now, let us investigate this connection in a wider context.

Lower and upper  $\varepsilon$ -approximations can be generalized via their element based definitions (4.1.1a) and (4.1.2a) relying on *arbitrary* binary relations  $\epsilon$  on  $U$  [32].

**4.19 Definition.** Let  $\epsilon$  be an arbitrary binary relation on  $U$  and  $X \in 2^U$ .

The *lower  $\epsilon$ -approximation* of  $X$  is

$$\underline{\epsilon}(X) = \{x \in U \mid [x]_\epsilon \subseteq X\},$$

and the *upper  $\epsilon$ -approximation* of  $X$  is

$$\bar{\epsilon}(X) = \{x \in U \mid [x]_\epsilon \cap X \neq \emptyset\}.$$

If  $\epsilon^{-1}$  denotes the inverse relation of  $\epsilon$ , in the same manner one can also define the lower and upper  $\epsilon^{-1}$ -approximations of  $X$ .

**4.20 Proposition.** ([33], Proposition 134) *Let  $\epsilon$  be an arbitrary binary relation on  $U$ .*

*Then  $\mathbb{G}(2^U, \bar{\epsilon}, \underline{\epsilon^{-1}}, 2^U)$  and  $\mathbb{G}(2^U, \overline{\epsilon^{-1}}, \underline{\epsilon}, 2^U)$  are Galois connections on  $(2^U, \subseteq)$ .*

Next corollary immediately comes from Proposition 4.20.

**4.21 Corollary.** *Let  $\epsilon$  be an arbitrary binary relation on  $U$ .*

*The pair  $(\bar{\epsilon}, \underline{\epsilon})$  is a Galois connection on  $(2^U, \subseteq)$  if and only if  $\epsilon$  is symmetric.*

*In particular, if  $\epsilon$  is an equivalence relation on  $U$ , then  $\mathbb{G}(2^U, \bar{\epsilon}, \underline{\epsilon}, 2^U)$  is a Galois connection on  $2^U$ .*

The next examples show that even if the relation  $\epsilon$  is symmetric, it is not sufficient that the upper and lower  $\epsilon$ -approximations relying on point-free definitions form Galois connection.

**4.22 Example.** (Csajbók [12], Example 3.10) Let  $U = \{x_1, x_2, x_3\}$  and

$$\epsilon = \{(x_1, x_1), (x_1, x_2), (x_2, x_1), (x_2, x_3), (x_3, x_2)\} \subset U \times U$$

be a symmetric binary relation on  $U$ .

We define the straightforward generalizations of  $U/\epsilon$ ,  $\mathfrak{D}_{U/\epsilon}$  as follows.

$$\begin{aligned} [x_1]_\epsilon &= \{u \in U \mid (x_1, u) \in \epsilon\} = \{x_1, x_2\}, \\ [x_2]_\epsilon &= \{u \in U \mid (x_2, u) \in \epsilon\} = \{x_1, x_3\}, \\ [x_3]_\epsilon &= \{u \in U \mid (x_3, u) \in \epsilon\} = \{x_2\}, \\ U/\epsilon &= \{[x_1]_\epsilon, [x_2]_\epsilon, [x_3]_\epsilon\} = \{\{x_1, x_2\}, \{x_1, x_3\}, \{x_2\}\}, \\ \mathfrak{D}_{U/\epsilon} &= \{\emptyset, [x_1]_\epsilon, [x_2]_\epsilon, [x_3]_\epsilon, [x_1]_\epsilon \cup [x_2]_\epsilon, [x_1]_\epsilon \cup [x_3]_\epsilon, [x_2]_\epsilon \\ &\quad \cup [x_3]_\epsilon, [x_1]_\epsilon \cup [x_2]_\epsilon \cup [x_3]_\epsilon\} \\ &= \{\emptyset, \{x_1, x_2\}, \{x_1, x_3\}, \{x_2\}, \underbrace{\{x_1, x_2\} \cup \{x_1, x_3\}}_{\{x_1, x_2, x_3\}}, \underbrace{\{x_1, x_2\} \cup \{x_2\}}_{\{x_1, x_2\}}, \\ &\quad \underbrace{\{x_1, x_3\} \cup \{x_2\}}_{\{x_1, x_2, x_3\}}, \underbrace{\{x_1, x_2\} \cup \{x_1, x_3\} \cup \{x_2\}}_{\{x_1, x_2, x_3\}}\} \\ &= \{\emptyset, \{x_1, x_2\}, \{x_1, x_3\}, \{x_2\}, \{x_1, x_2, x_3\}\}. \end{aligned}$$

Note that  $\mathfrak{D}_{U/\epsilon} \subseteq 2^U$  is a subsystem but not a  $\sigma$ -algebra. For example,  $\{x_1, x_2\} \cap \{x_1, x_3\} = \{x_1\} \notin \mathfrak{D}_{U/\epsilon}$ .

Case 1. *Elementary set based definitions relying on  $U/\epsilon$ .*

Let us define the lower and upper  $\epsilon$ -approximations taking the pattern by the equations 4.1.1b and 4.1.2b, respectively:

$$\begin{aligned}\underline{\epsilon}_e : 2^U &\rightarrow \mathfrak{D}_{U/\epsilon}, & X &\mapsto \bigcup \{Y \mid Y \in U/\epsilon, Y \subseteq X\}, \\ \bar{\epsilon}_e : 2^U &\rightarrow \mathfrak{D}_{U/\epsilon}, & X &\mapsto \bigcup \{Y \mid Y \in U/\epsilon, Y \cap X \neq \emptyset\}.\end{aligned}$$

Clearly, the maps  $\underline{\epsilon}_e$  and  $\bar{\epsilon}_e$  are monotone.

For instance, for the set  $\{x_2\} \in 2^U$ :

$$\begin{aligned}\underline{\epsilon}_e(\{x_2\}) &= \bigcup \{Y \mid Y \in U/\epsilon, Y \subseteq \{x_2\}\} = \bigcup \{\{x_2\}\} = \{x_2\}, \\ \bar{\epsilon}_e(\{x_2\}) &= \bigcup \{Y \mid Y \in U/\epsilon, Y \cap \{x_2\} \neq \emptyset\} \\ &= \bigcup \{\{x_1, x_2\}, \{x_2\}\} = \{x_1, x_2\}.\end{aligned}$$

Do the following relations  $\{x_2\} \subseteq \underline{\epsilon}_e(\bar{\epsilon}_e(\{x_2\}))$  and/or  $\bar{\epsilon}_e(\underline{\epsilon}_e(\{x_2\})) \subseteq \{x_2\}$  hold?

$$\begin{aligned}\underline{\epsilon}_e(\bar{\epsilon}_e(\{x_2\})) &= \underline{\epsilon}_e(\{x_1, x_2\}) = \bigcup \{Y \mid Y \in U/\epsilon, Y \subseteq \{x_1, x_2\}\} \\ &= \bigcup \{\{x_1, x_2\}, \{x_2\}\} = \{x_1, x_2\} \supsetneq \{x_2\}, \\ \bar{\epsilon}_e(\underline{\epsilon}_e(\{x_2\})) &= \bar{\epsilon}_e(\{x_2\}) = \{x_1, x_2\} \not\subseteq \{x_2\}.\end{aligned}$$

That is, by Proposition 2.2,  $(2^U, \bar{\epsilon}_e, \underline{\epsilon}_e, 2^U)$  does not form Galois connection.

Case 2. *Subsystem based definitions relying on  $\mathfrak{D}_{U/\epsilon}$ .*

Let us define the lower and upper  $\epsilon$ -approximations after the pattern of the equations 4.1.1c and 4.1.2c, respectively (note that,  $\mathfrak{D}_{U/\epsilon}$  is closed under unions, but not closed under intersections):

$$\begin{aligned}\underline{\epsilon}_s : 2^U &\rightarrow \mathfrak{D}_{U/\epsilon}, & X &\mapsto \bigcup \{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, Y \subseteq X\}, \\ \bar{\epsilon}_s : 2^U &\rightarrow 2^U, & X &\mapsto \bigcap \{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, X \subseteq Y\}.\end{aligned}$$

Clearly, the map  $\underline{\epsilon}_s$  is monotone.

The map  $\bar{\epsilon}_s$  is also monotone. Namely, let  $X_1 \subseteq X_2$  be subsets of  $U$ .

- If  $\{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, X_1 \subseteq Y\} = \emptyset$ , then  $\{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, X_2 \subseteq Y\} = \emptyset$  also holds, and so  $\bar{\epsilon}_s(X_1) = \bar{\epsilon}_s(X_2) = \bigcap \emptyset = U$ .
- If  $\{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, X_1 \subseteq Y\} \neq \emptyset$  and  $\{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, X_2 \subseteq Y\} = \emptyset$ , then  $\bar{\epsilon}_s(X_1) \subseteq \bar{\epsilon}_s(X_2) = \bigcap \emptyset = U$ .

- If  $\{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, X_1 \subseteq Y\}, \{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, X_2 \subseteq Y\} \neq \emptyset$ , then

$$\{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, X_2 \subseteq Y\} \subseteq \{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, X_1 \subseteq Y\},$$

and so

$$\begin{aligned} \bar{\epsilon}_s(X_1) &= \bigcap \{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, X_1 \subseteq Y\} \\ &\subseteq \bigcap \{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, X_2 \subseteq Y\} = \bar{\epsilon}_s(X_2). \end{aligned}$$

For instance, for the set  $\{x_1\}$ :

$$\begin{aligned} \underline{\epsilon}_s(\{x_1\}) &= \bigcup \{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, Y \subseteq \{x_1\}\} = \bigcup \{\emptyset\} = \emptyset, \\ \bar{\epsilon}_s(\{x_1\}) &= \bigcap \{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, \{x_1\} \subseteq Y\} \\ &= \bigcap \{\{x_1, x_2\}, \{x_1, x_3\}, \{x_1, x_2, x_3\}\} = \{x_1\}. \end{aligned}$$

Do the following relations  $\{x_1\} \subseteq \underline{\epsilon}_s(\bar{\epsilon}_s(\{x_1\}))$  and/or  $\bar{\epsilon}_s(\underline{\epsilon}_s(\{x_1\})) \subseteq \{x_1\}$  hold?

$$\begin{aligned} \underline{\epsilon}_s(\bar{\epsilon}_s(\{x_1\})) &= \underline{\epsilon}_s(\{x_1\}) = \emptyset \not\subseteq \{x_1\}, \\ \bar{\epsilon}_s(\underline{\epsilon}_s(\{x_1\})) &= \bar{\epsilon}_s(\emptyset) = \bigcap \{Y \mid Y \in \mathfrak{D}_{U/\epsilon}, \emptyset \subseteq Y\} = \bigcap \mathfrak{D}_{U/\epsilon} = \emptyset \subseteq \{x_1\}. \end{aligned}$$

That is, by Proposition 2.2,  $(2^U, \bar{\epsilon}_s, \underline{\epsilon}_s, 2^U)$  does not form a Galois connection.

Case 3. *Point-wise definitions.* Now let us check and see that the sets  $\{x_1\}$  and  $\{x_2\}$  fulfill the conditions of Proposition 2.2 in the case of point-wise definitions of the approximations.

Let us define the lower and upper  $\epsilon$ -approximations of  $\{x_1\}$  and  $\{x_2\}$  in the point-wise manner due to equations 4.1.1a and 4.1.2a:

$$\begin{aligned} \underline{\epsilon}_p(\{x_1\}) &= \{x \in U \mid [x]_\epsilon \subseteq \{x_1\}\} = \emptyset, \\ \bar{\epsilon}_p(\{x_1\}) &= \{x \in U \mid [x]_\epsilon \cap \{x_1\} \neq \emptyset\} = \{x_1, x_2\}, \\ \underline{\epsilon}_p(\{x_2\}) &= \{x \in U \mid [x]_\epsilon \subseteq \{x_2\}\} = \{x_3\}, \\ \bar{\epsilon}_p(\{x_2\}) &= \{x \in U \mid [x]_\epsilon \cap \{x_2\} \neq \emptyset\} = \{x_1, x_3\}. \end{aligned}$$

Of course, by Corollary 4.21, the maps  $\underline{\epsilon}_p$  and  $\bar{\epsilon}_p$  are monotone. Moreover, the formulae

$$\begin{aligned} \{x_1\} &\subseteq \underline{\epsilon}_p(\bar{\epsilon}_p(\{x_1\})) \text{ and } \bar{\epsilon}_p(\underline{\epsilon}_p(\{x_1\})) \subseteq \{x_1\}, \\ \{x_2\} &\subseteq \underline{\epsilon}_p(\bar{\epsilon}_p(\{x_2\})) \text{ and } \bar{\epsilon}_p(\underline{\epsilon}_p(\{x_3\})) \subseteq \{x_3\} \end{aligned}$$

must hold. Indeed,

$$\begin{aligned}\{x_1\} &\subseteq \epsilon_p(\bar{\epsilon}_p(\{x_1\})) = \epsilon_p(\{x_1, x_2\}) = \{x \in U \mid [x]_\epsilon \subseteq \{x_1, x_2\}\} = \{x_1, x_3\}, \\ \bar{\epsilon}_p(\epsilon_p(\{x_1\})) &= \bar{\epsilon}_p(\emptyset) = \{x \in U \mid [x]_\epsilon \cap \emptyset \neq \emptyset\} = \emptyset \subseteq \{x_1\},\end{aligned}$$

and

$$\begin{aligned}\{x_2\} &\subseteq \epsilon_p(\bar{\epsilon}_p(\{x_2\})) = \epsilon_p(\{x_1, x_3\}) = \{x \in U \mid [x]_\epsilon \subseteq \{x_1, x_3\}\} = \{x_2\}, \\ \bar{\epsilon}_p(\epsilon_p(\{x_2\})) &= \bar{\epsilon}_p(\{x_3\}) = \{x \in U \mid [x]_\epsilon \cap \{x_3\} \neq \emptyset\} = \{x_2\} \subseteq \{x_2\}.\end{aligned}$$





## Chapter 5

# Approximation of Sets Based on Partial Covering

### 5.1 Introduction

In practice, there are objects which cannot be characterized by certain features directly.

Some illustrative examples:

- Bald men cannot be characterized with the property ‘color of hair’.
- An infinite set is investigated via a finite family of its finite subsets. For instance, a number theorist studies the regularities of natural numbers using computers.
- Security policies are partial-natured in corporate information security. Typically some policies may only apply to specific hardware appliances, software applications or type of information.

Moreover, there are features with which *a set and its complement* cannot be treated simultaneously. For instance, complements of recursively enumerable sets are not necessarily recursively enumerable. The membership of recursively enumerable sets can effectively be determined by a *finite amount* of information, while the determination of their non-membership requires an *infinite amount* of information [50]. That is, the complement of a recursively enumerable set cannot necessarily be determined effectively. In other words, the recursively enumerable sets can be managed by computers (e.g., via a special rewriting system, the Markov algorithm [65]), while

its complement not necessarily. Thus, this is an important *practical* partial approximation problem: how can we approximate an arbitrary set with recursively enumerable sets?

Another question is the point-freeness. Let us suppose that we study a collection of groups of individuals. In some cases it is important to distinguish individuals in these groups, whereas in other cases it is irrelevant. For instance, in genotype-phenotype investigations for understanding evolution it is reasonable to distinguish individuals (for a generalized point-set topological theory, see, e.g. [71]). On the other hand, during the investigation of spreading of different types of floral zones in a given geographical area, the distinction of the individuals has no relevance.

Moreover, these floral zones overlap each other. In addition they generally do not cover the entire area, e.g. on lands of desert, or when we investigate the spreading of woodlands excluding the underwood. As another example, in the game of go there are two groups of stones, black and white. Black stones are inherently undistinguishable, so are the white ones. In addition, the black and white zones overlap each other, and even together they never cover the entire game table.

Throughout this section let  $U$  be a nonempty set called the *universe of discourse*.

## 5.2 Base Systems

According to the general set theoretic approximation framework, let  $\mathfrak{B} \subseteq 2^U$  be a base system, i.e. a nonempty family of nonempty subsets of  $U$ . Its members, the  $\mathfrak{B}$ -sets, are considered as our *primary tools* because we want the subsets of  $U$  to be approximated with their help.

Now, let us define our *derived tools*, i.e. an extension of  $\mathfrak{B}$  as follows.

**5.1 Definition. (Csajbók [12], Definition 4.1)** A nonempty subset  $X \in 2^U$  is  $\mathfrak{B}$ -*definable* if there exists a family of sets  $\mathfrak{D} \subseteq \mathfrak{B}$  such that  $X = \bigcup \mathfrak{D}$ , otherwise  $X$  is  $\mathfrak{B}$ -*undefinable*.

The empty set is considered to be a  $\mathfrak{B}$ -definable set.

Let  $\mathfrak{D}_{\mathfrak{B}}$  denote the family of  $\mathfrak{B}$ -definable sets of  $U$ .

## 5.3 Single-Layered Base Systems

Some properties of the rough set theory can partly be preserved with the help of the next constrained version of the base system.

**5.2 Definition.** (Csajbók [12], Definition 4.2) The base system  $\mathfrak{B} \subseteq 2^U$  is *single-layered*, if

$$\forall B \in \mathfrak{B} \quad \forall \mathfrak{B}' \subseteq \mathfrak{B} \setminus \{B\} \quad (B \cap \bigcup \mathfrak{B}' \neq B),$$

and *one-layered*, if

$$\forall B \in \mathfrak{B} \quad \forall \mathfrak{B}' \subseteq \mathfrak{B} \setminus \{B\} \quad (B \cap \bigcup \mathfrak{B}' = \emptyset).$$

Informally, a base system  $\mathfrak{B}$  is single-layered if every nonempty  $\mathfrak{B}$ -definable subset has at least one element which can be characterized by exactly one primary tool, whereas  $\mathfrak{B}$  is one-layered if every element of the universe can be characterized by at most one primary tool.

An important question is how can we form a single-layered base system from an arbitrary one. In general, this problem is reduced by the practice to finite base systems ( $|\mathfrak{B}| < \infty$ ).

The simplest way to construct a single/one-layered base system from an arbitrary one is to form its *intersection structure*. Formally, a nonempty family  $\mathfrak{S}$  of subsets of the universe  $U$  is an intersection structure if  $\forall \mathfrak{S}' (\neq \emptyset) \subseteq \mathfrak{S} \quad (\bigcap \mathfrak{S}' \in \mathfrak{S})$ , i.e. it is closed under intersection [21].

Let us take an arbitrary base system  $\mathfrak{B}$  and create its intersection structure  $C(\mathfrak{B})$  as the smallest set which satisfies the following two properties:

- (1)  $\mathfrak{B} \subseteq C(\mathfrak{B})$ ;
- (2) if  $\mathfrak{B}', \mathfrak{B}'' \subseteq C(\mathfrak{B})$ , then  $\mathfrak{B}' \cap \mathfrak{B}'' \in C(\mathfrak{B})$ .

Note that any intersections of primary tools are also considered primary tools, i.e. new ‘combined’ primary tools appear in  $C(\mathfrak{B})$ . In other words, the intersection structure  $C(\mathfrak{B})$  is a collection of all original and all possible ‘combined’ primary tools.

Having given the intersection structure  $C(\mathfrak{B})$ , first, we can create a single-layered base system  $SC(\mathfrak{B})$  as the smallest set which satisfies the following two properties:

- (1)  $SC(\mathfrak{B}) = \emptyset$  is a single-layered base system;
- (2) if  $B, B' \in C(\mathfrak{B})$  such that  $B \subset B'$ , then let  $B, B' \setminus B \in SC(\mathfrak{B})$ .

Next, having given a single-layered base system  $SC(\mathfrak{B})$ , we can create a one-layered base system  $OSC(\mathfrak{B})$  as the smallest set which satisfies the following two properties:

- (1)  $OSC(\mathfrak{B}) = \emptyset$  is a one-layered base system;
- (2) if  $B, B' \in SC(\mathfrak{B})$  such that  $B \cap B' \neq \emptyset$ , then let the differences  $B \setminus B', B' \setminus B \in OSC(\mathfrak{B})$ .

**5.3 Proposition.** (Csajbók [12], Proposition 4.3) *Let  $\mathfrak{B} \subseteq 2^U$  be a base system.*

*Then the map  $i_{\mathfrak{B}} : 2^{\mathfrak{B}} \rightarrow \mathfrak{D}_{\mathfrak{B}}, \mathfrak{D} \mapsto \bigcup \mathfrak{D}$  is a bijection if and only if  $\mathfrak{B}$  is single-layered.*

**Proof.** If  $\#\mathfrak{B} = 1$ , the base system  $\mathfrak{B} = \{B\}$  is single-layered and  $i_{\mathfrak{B}} : \{\emptyset, \{B\}\} \rightarrow \{\emptyset, \{B\}\}, \emptyset \mapsto \bigcup \emptyset = \emptyset, B \mapsto \bigcup \{B\} = B$  is a bijection evidently. Now, let us suppose that  $\#\mathfrak{B} > 1$ .

( $\Rightarrow$ ) Let us assume, by contradiction, that the base system  $\mathfrak{B}$  is not single-layered. If so,

$$\exists B \in \mathfrak{B} \exists \mathfrak{B}' \subseteq \mathfrak{B} \setminus \{B\} (B \subseteq \bigcup \mathfrak{B}').$$

Hence,  $\mathfrak{B}', \mathfrak{B}' \cup \{B\} \in 2^{\mathfrak{B}}$  and  $\bigcup \mathfrak{B}' = \bigcup (\mathfrak{B}' \cup \{B\}) \in \mathfrak{D}_{\mathfrak{B}}$ , but  $\mathfrak{B}' \neq \mathfrak{B}' \cup \{B\}$  because of  $\mathfrak{B}' \subseteq \mathfrak{B} \setminus \{B\}$ . This, however, contradicts the assumption that the map  $i_{\mathfrak{B}}$  is injective.

( $\Leftarrow$ ) Clearly, by Definition 5.1., the map  $i_{\mathfrak{B}}$  is onto.

By contradiction, let us suppose that the map  $i_{\mathfrak{B}}$  is not injective. In this case,

$$\exists \mathfrak{B}_1, \mathfrak{B}_2 \subseteq \mathfrak{B} (\mathfrak{B}_1 \neq \mathfrak{B}_2 \wedge \bigcup \mathfrak{B}_1 = \bigcup \mathfrak{B}_2).$$

Since  $\mathfrak{B}_1 \neq \mathfrak{B}_2$ , there exists  $B \in \mathfrak{B}$  such that  $B$  is an element of either one or the other. Without any loss of generality we can assume that  $B \in \mathfrak{B}_1$  and  $B \notin \mathfrak{B}_2$ . Clearly,  $B \subseteq \bigcup \mathfrak{B}_1 = \bigcup \mathfrak{B}_2$ . Hence,  $B \in \mathfrak{B}$ ,  $\mathfrak{B}_2 \subseteq \mathfrak{B} \setminus \{B\}$  but  $B \cap \bigcup \mathfrak{B}_2 = B$ , which, however, contradicts the assumption that the base system  $\mathfrak{B}$  is single-layered.  $\square$

The following two statements, provided that the base system is single-layered, present certain properties that Pawlak's rough set theory has.

**5.4 Lemma.** (Csajbók [12], Lemma 4.3) *For a base system  $\mathfrak{B} \subseteq 2^U$*

$$\forall B \in \mathfrak{B} \forall \mathfrak{B}' \subseteq \mathfrak{B} (B \subseteq \bigcup \mathfrak{B}' \Leftrightarrow B \in \mathfrak{B}')$$

*if and only if the base system  $\mathfrak{B}$  is single-layered.*

**Proof.** ( $\Rightarrow$ ) Let us suppose, by contradiction, that the base system  $\mathfrak{B}$  is not single-layered, that is  $\exists B \in \mathfrak{B} \wedge \exists \mathfrak{B}' \subseteq \mathfrak{B} \setminus \{B\} (B \subseteq \bigcup \mathfrak{B}')$ .

Hence  $B \subseteq \bigcup \mathfrak{B}'$  but  $B \notin \mathfrak{B}'$ . This contradicts the assumption that  $\forall B \in \mathfrak{B} \forall \mathfrak{B}' \subseteq \mathfrak{B} (B \subseteq \bigcup \mathfrak{B}' \Rightarrow B \in \mathfrak{B}')$ .

( $\Leftarrow$ ) Of course, the statement  $B \in \mathfrak{B}' \Rightarrow B \subseteq \bigcup \mathfrak{B}'$  is trivial. Thus we have to prove that  $\forall B \in \mathfrak{B} \forall \mathfrak{B}' \subseteq \mathfrak{B} (B \subseteq \bigcup \mathfrak{B}' \Rightarrow B \in \mathfrak{B}')$ . Contrary to this statement, let us assume that  $\exists B \in \mathfrak{B} \exists \mathfrak{B}' \subseteq \mathfrak{B} (B \subseteq \bigcup \mathfrak{B}' \wedge B \notin \mathfrak{B}')$ .

Hence  $\mathfrak{B}' \subseteq \mathfrak{B} \setminus \{B\}$  and  $B \subseteq \bigcup \mathfrak{B}'$  which, however, contradicts the assumption that the base system  $\mathfrak{B}$  is single-layered.  $\square$

**5.5 Proposition.** (Csajbók [12], Proposition 4.5) *Let  $\mathfrak{B} \subseteq 2^U$  be a base system.*

*Then the posets  $(2^{\mathfrak{B}}, \subseteq)$  and  $(\mathfrak{D}_{\mathfrak{B}}, \subseteq)$  are order isomorphic via the map  $\mathbf{i}_{\mathfrak{B}} : 2^{\mathfrak{B}} \rightarrow \mathfrak{D}_{\mathfrak{B}}, \mathfrak{X} \mapsto \bigcup \mathfrak{X}$  if and only if the base system  $\mathfrak{B}$  is single-layered.*

**Proof.** By Proposition 5.3, the map  $\mathbf{i}_{\mathfrak{B}}$  is a bijection if and only if the base system  $\mathfrak{B}$  is single-layered.

The monotonicity of  $\mathbf{i}_{\mathfrak{B}}$  is trivial. The monotonicity of  $\mathbf{i}_{\mathfrak{B}}^{-1}$  can similarly be proven to Proposition 4.16 changing the reference to Lemma 4.15 for the reference to Lemma 5.4.  $\square$

## 5.4 Lower and Upper $\mathfrak{B}$ -Approximations

Let us define the lower and upper approximations based on partial covering. Recall that  $\mathfrak{B}$  does not cover the universe necessarily.

**5.6 Definition.** (Csajbók [12], Definition 4.6) Let  $\mathfrak{B} \subseteq 2^U$  be a base system and  $X$  be any subset of  $U$ .

The *lower  $\mathfrak{B}$ -approximation of  $X$*  (Fig. 5.1) is

$$\mathfrak{C}_{\mathfrak{B}}^{\flat}(X) = \bigcup \{Y \mid Y \in \mathfrak{B}, Y \subseteq X\},$$

the *upper  $\mathfrak{B}$ -approximation of  $X$*  (Fig. 5.2) is

$$\mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) = \bigcup \{Y \mid Y \in \mathfrak{B}, Y \cap X \neq \emptyset\}.$$

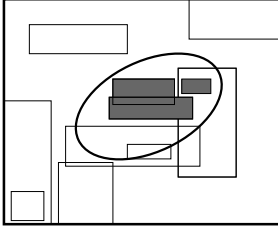


Figure 5.1. Lower approximation

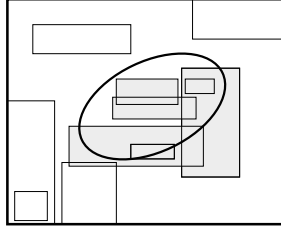


Figure 5.2. Upper approximation

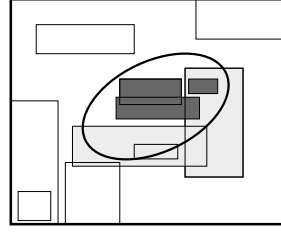


Figure 5.3. Lower and upper approximations

Notice that  $\mathfrak{C}_{\mathfrak{B}}^b$  and  $\mathfrak{C}_{\mathfrak{B}}^\#$  are the straightforward point-free generalizations of lower and upper  $\varepsilon$ -approximations relying on  $\varepsilon$ -elementary sets.

Clearly,  $\mathfrak{C}_{\mathfrak{B}}^b(X), \mathfrak{C}_{\mathfrak{B}}^\#(X) \in \mathfrak{D}_{\mathfrak{B}}$ , and the maps  $\mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\# : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}$  are total and generally many-to-one.

**5.7 Proposition.** *Let  $\langle \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\# \rangle$  be a  $\mathfrak{B}$ -approximation pair of maps  $\mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\# : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}$  on  $U$ . Then*

- (1)  $\langle \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\# \rangle$  is a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong approximation pair on  $U$ ;
- (2)  $\langle \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\# \rangle$  is a strong approximation pair on  $U$  if and only if the base system  $\mathfrak{B}$  covers the universe  $U$ .

*In other words, the maps  $\mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\#$  fulfill the following conditions:*

- (C0)  $\mathfrak{C}_{\mathfrak{B}}^b(2^U), \mathfrak{C}_{\mathfrak{B}}^\#(2^U) \subseteq \mathfrak{D}_{\mathfrak{B}}$  (definability of  $\mathfrak{C}_{\mathfrak{B}}^b$  and  $\mathfrak{C}_{\mathfrak{B}}^\#$ ).
- (C1)  $\mathfrak{C}_{\mathfrak{B}}^b$  and  $\mathfrak{C}_{\mathfrak{B}}^\#$  are monotone (monotonicity of  $\mathfrak{C}_{\mathfrak{B}}^b$  and  $\mathfrak{C}_{\mathfrak{B}}^\#$ ).
- (C2)  $\mathfrak{C}_{\mathfrak{B}}^\#(\emptyset) = \emptyset$  (normality of  $\mathfrak{C}_{\mathfrak{B}}^\#$ ).
- (C3) If  $D \in \mathfrak{D}_{\mathfrak{B}}$ , then  $\mathfrak{C}_{\mathfrak{B}}^b(D) = D$  ( $\mathfrak{C}_{\mathfrak{B}}^b$  is standard or Pawlak's type).
- (C4) If  $S \in 2^U$ , then  $\mathfrak{C}_{\mathfrak{B}}^b(S) \subseteq \mathfrak{C}_{\mathfrak{B}}^\#(S)$  (approximation property).
- (C5)  $\mathfrak{C}_{\mathfrak{B}}^b$  is contractive.
- (C6)  $\mathfrak{C}_{\mathfrak{B}}^\#$  is extensive if and only if  $\mathfrak{B}$  covers the universe  $U$ .

**Proof.** The conditions (C0), (C1), (C2) and (C4), (C5) are straightforward by the definition of lower and upper  $\mathfrak{B}$ -approximations.

(C3) Clearly, if  $\emptyset \in \mathfrak{D}_{\mathfrak{B}}$ , then  $\mathfrak{C}_{\mathfrak{B}}^b(\emptyset) = \emptyset$ .

If  $\emptyset \neq D \in \mathfrak{D}_{\mathfrak{B}}$ , there exists at least one nonempty family of sets  $\mathfrak{B}' \subseteq \mathfrak{B}$  such that

$$D = \bigcup \mathfrak{B}' = \bigcup \{B \mid B \in \mathfrak{B}', B \subseteq D\} \subseteq \bigcup \{B \mid B \in \mathfrak{B}, B \subseteq D\} = \mathfrak{C}_{\mathfrak{B}}^b(D).$$

On the other hand, we have  $\mathfrak{C}_{\mathfrak{B}}^b(D) \subseteq D$ . Thus  $\mathfrak{C}_{\mathfrak{B}}^b(D) = D$ .

(C6) ( $\Rightarrow$ ) If  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}$  is extensive, then

$$U \subseteq \mathfrak{C}_{\mathfrak{B}}^{\sharp}(U) = \bigcup \{B \mid B \in \mathfrak{B}, B \subseteq U\} = \bigcup \mathfrak{B}.$$

Of course,  $\bigcup \mathfrak{B} \subseteq U$ , and so  $\bigcup \mathfrak{B} = U$ .

( $\Leftarrow$ ) If  $\mathfrak{B}$  covers the universe, then  $\forall S \in 2^U (S \subseteq U = \bigcup \mathfrak{B})$ . Thus we get

$$\begin{aligned} S &\subseteq \bigcup (\mathfrak{B} \setminus \{B \mid B \in \mathfrak{B}, B \cap S = \emptyset\}) \\ &= \bigcup \{B \mid B \in \mathfrak{B}, B \cap S \neq \emptyset\} = \mathfrak{C}_{\mathfrak{B}}^{\sharp}(S). \end{aligned} \quad \square$$

In the language of the set theoretic approximation framework, by Proposition 5.7,  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^{\sharp} \rangle$  is a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong approximation framework, and it is a strong one if and only if the base system  $\mathfrak{B}$  covers the universe.

The next properties of  $\mathfrak{C}_{\mathfrak{B}}^b$  and  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}$  immediately follows from Proposition 3.9. Of course, they can easily be proven by Definition 5.6 directly.

**5.8 Proposition.** *Let  $\langle \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^{\sharp} \rangle$  be a  $\mathfrak{B}$ -approximation pair of maps  $\mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^{\sharp} : 2^U \rightarrow \mathfrak{B}_{\mathfrak{D}}$  on  $U$ . Then*

- (1)  $\mathfrak{C}_{\mathfrak{B}}^b(\emptyset) = \emptyset$  (normality of  $\mathfrak{C}_{\mathfrak{B}}^b$ ).
- (2)  $\forall S \in 2^U (\mathfrak{C}_{\mathfrak{B}}^b(\mathfrak{C}_{\mathfrak{B}}^b(S)) = \mathfrak{C}_{\mathfrak{B}}^b(S))$  (idempotency of  $\mathfrak{C}_{\mathfrak{B}}^b$ ).
- (3)  $\mathfrak{C}_{\mathfrak{B}}^b(2^U) = \mathfrak{D}_{\mathfrak{B}}$  ( $\mathfrak{C}_{\mathfrak{B}}^b$  is surjective).
- (4)  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}(2^U) \subseteq \mathfrak{C}_{\mathfrak{B}}^b(2^U) = \mathfrak{D}_{\mathfrak{B}}$ .

**Proof.** (4) We have to show that the inclusion  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}(2^U) \subseteq \mathfrak{C}_{\mathfrak{B}}^b(2^U)$  may be proper because of the particular constructions of lower and upper approximation maps. To do this, let  $U = \{a, b\}$ ,  $\mathfrak{B} = \{\{a\}, \{a, b\}\}$  and  $\mathfrak{D}_{\mathfrak{B}} = \{\emptyset, \{a\}, \{a, b\}\}$ . Then

$$\mathfrak{C}_{\mathfrak{B}}^b : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}, \quad X \mapsto \mathfrak{C}_{\mathfrak{B}}^b(X) = \begin{cases} \emptyset, & \text{if } X = \emptyset, \{b\}; \\ \{a\}, & \text{if } X = \{a\}; \\ \{a, b\}, & \text{if } X = \{a, b\}, \end{cases}$$

and

$$\mathfrak{C}_{\mathfrak{B}}^{\sharp} : 2^U \rightarrow \mathfrak{D}_{\mathfrak{B}}, \quad X \mapsto \mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) = \begin{cases} \emptyset, & \text{if } X = \emptyset; \\ \{a, b\}, & \text{if } X = \{a\}, \{b\}, \{a, b\}. \end{cases}$$

Conditions (C1)–(C5) can easily be checked, however,  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}$  is not surjective:  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}(2^U) = \{\emptyset, \{a, b\}\} \subsetneq \{\emptyset, \{a\}, \{a, b\}\} = \mathfrak{C}_{\mathfrak{B}}^b(2^U) = \mathfrak{D}_{\mathfrak{B}}$ .  $\square$

Unlike Pawlak's approximation spaces (cf. Proposition 4.7), the  $\mathfrak{B}$ -definable property is generally not equivalent to the condition  $\mathfrak{C}_{\mathfrak{B}}^b(X) = \mathfrak{C}_{\mathfrak{B}}^{\sharp}(X)$ .

**5.9 Proposition.** (Csajbók [12], Proposition 4.7) *Let  $\mathfrak{B} \subseteq 2^U$  be a base system. Then*

(1)  *$X \in 2^U$  is  $\mathfrak{B}$ -definable if and only if  $\mathfrak{C}_{\mathfrak{B}}^b(X) = X$ .*

(2)  *$X \in 2^U$  is  $\mathfrak{B}$ -undefinable if and only if  $\mathfrak{C}_{\mathfrak{B}}^b(X) \neq X$ .*

**Proof.** (1) It is straightforward, when  $X = \emptyset$ . Let  $X \neq \emptyset$ .

( $\Rightarrow$ ) If  $X \in \mathfrak{D}_{\mathfrak{B}}$ , there exists at least one nonempty family of sets  $\mathfrak{B}' \subseteq \mathfrak{B}$  such that

$$X = \bigcup \mathfrak{B}' = \bigcup \{Y \mid Y \in \mathfrak{B}', Y \subseteq X\} \subseteq \bigcup \{Y \mid Y \in \mathfrak{B}, Y \subseteq X\} = \mathfrak{C}_{\mathfrak{B}}^b(X).$$

On the other hand,  $\mathfrak{C}_{\mathfrak{B}}^b(X) \subseteq X$ , thus  $X = \mathfrak{C}_{\mathfrak{B}}^b(X)$ .

( $\Leftarrow$ )  $\emptyset \neq X = \mathfrak{C}_{\mathfrak{B}}^b(X) \in \mathfrak{D}_{\mathfrak{B}}$ .

(2) It is the contrapositive version of 1.  $\square$

## 5.5 Representation of Sets

Clearly, for a  $\mathfrak{B}$ -definable subset  $D \in \mathfrak{D}_{\mathfrak{B}}$  there may exist two or more families of  $\mathfrak{B}$ -sets such that their unions are equal to  $D$ . For instance, let  $\mathfrak{B} = \{B_1, B_2\}$  ( $B_1, B_2 \in 2^U$ ) be a base system such that  $B_1 \subsetneq B_2$ . If  $\mathfrak{F}_1 = \{B_1, B_2\}$ ,  $\mathfrak{F}_2 = \{B_2\}$ , then  $\mathfrak{F}_1 \neq \mathfrak{F}_2$  but  $\bigcup \mathfrak{F}_1 = \bigcup \mathfrak{F}_2 = B_2$ . Of course, the same is true for lower and upper  $\mathfrak{B}$ -approximations in general.

If  $D \in \mathfrak{D}_{\mathfrak{B}}$  is a  $\mathfrak{B}$ -definable set, then let  $\mathfrak{F}_{\mathfrak{B}}(D) \subseteq \mathfrak{B}$  denote a possible family of  $\mathfrak{B}$ -sets so that  $\bigcup \mathfrak{F}_{\mathfrak{B}}(D) = D$ .  $\mathfrak{F}_{\mathfrak{B}}(D)$  is called a (possible)  $\mathfrak{B}$ -composition of  $D$ . Unlike Pawlak's approximation spaces, the  $\mathfrak{B}$ -compositions of  $\mathfrak{B}$ -definable sets are generally not unique.



**5.10 Definition.** Let  $\mathfrak{B} \subseteq 2^U$  be a base system.

The  $\mathfrak{B}$ -definable set  $D \in \mathfrak{D}_{\mathfrak{B}}$  is  $\mathfrak{B}$ -representable, if there exists exactly one  $\mathfrak{B}$ -composition  $\mathfrak{F}_{\mathfrak{B}}(D)(\subseteq \mathfrak{B})$  of  $D$  such that  $D = \bigcup \mathfrak{F}_{\mathfrak{B}}(D)$ .

In this case, it is said that  $\mathfrak{F}_{\mathfrak{B}}(D)$  is the  $\mathfrak{B}$ -representation of  $D$ .

**5.11 Proposition.** Let  $\mathfrak{B} \subseteq 2^U$  be a base system.

All  $\mathfrak{B}$ -definable subsets  $D \in \mathfrak{D}_{\mathfrak{B}}$  of  $U$  are  $\mathfrak{B}$ -representable if and only if the base system  $\mathfrak{B}$  is single-layered.

**Proof.** All  $\mathfrak{B}$ -definable subsets of  $U$  are  $\mathfrak{B}$ -representable if and only if the map  $\mathfrak{F}_{\mathfrak{B}} : \mathfrak{D}_{\mathfrak{B}} \rightarrow 2^{\mathfrak{B}}, \bigcup \mathfrak{D} \mapsto \mathfrak{D}$  is the inverse of  $\mathfrak{i}_{\mathfrak{B}} : 2^{\mathfrak{B}} \rightarrow \mathfrak{D}_{\mathfrak{B}}, \mathfrak{D} \mapsto \bigcup \mathfrak{D}$ . A map has an inverse map if and only if it is a bijection. Consequently, all  $\mathfrak{B}$ -definable subsets of  $U$  are  $\mathfrak{B}$ -representable if and only if the map  $\mathfrak{i}_{\mathfrak{B}} : 2^{\mathfrak{B}} \rightarrow \mathfrak{D}_{\mathfrak{B}}, \mathfrak{D} \mapsto \bigcup \mathfrak{D}$  is a bijection. And so, this proposition is just a restatement of Proposition 5.3.  $\square$

**5.12 Corollary.** Let  $\mathfrak{B} \subseteq 2^U$  be a base system.

All  $\mathfrak{B}$ -definable subsets  $D \in \mathfrak{D}_{\mathfrak{B}}$  of  $U$  are  $\mathfrak{B}$ -representable in the following form

$$D = \bigcup \mathfrak{F}_{\mathfrak{B}}(D), \text{ where } \mathfrak{F}_{\mathfrak{B}}(D) = \{Y \mid Y \in \mathfrak{B}, Y \subseteq D\},$$

if and only if the base system  $\mathfrak{B}$  is single-layered.

**Proof.** According to Proposition 5.11,  $\mathfrak{B}$  is single-layered if and only if all  $\mathfrak{B}$ -definable subsets are  $\mathfrak{B}$ -representable. And so, we only have to show that the  $\mathfrak{B}$ -representations of all  $\mathfrak{B}$ -definable subsets are of the form

$$\mathfrak{F}_{\mathfrak{B}}(D) = \{Y \mid Y \in \mathfrak{B}, Y \subseteq D\}.$$

Since  $D = \bigcup \{Y \mid Y \in \mathfrak{B}, Y \subseteq D\}$  satisfies for all  $\mathfrak{B}$ -definable subsets  $D \in \mathfrak{D}_{\mathfrak{B}}$  by definition, the claim immediately follows from the uniqueness of  $\mathfrak{B}$ -representation.  $\square$

**5.13 Proposition.** Let  $\mathfrak{B} \subseteq 2^U$  be a base system,  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^{\#} \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong approximation space and  $X$  be a subset of  $U$ .

Then the sets  $\mathfrak{C}_{\mathfrak{B}}^b(X)$  and  $\mathfrak{C}_{\mathfrak{B}}^{\#}(X)$  are  $\mathfrak{B}$ -representable in the forms

$$\begin{aligned} \mathfrak{C}_{\mathfrak{B}}^b(X) &= \bigcup \mathfrak{F}_{\mathfrak{B}}^b(X), \text{ where } \mathfrak{F}_{\mathfrak{B}}^b(X) = \{Y \mid Y \in \mathfrak{B}, Y \subseteq X\}, \\ \mathfrak{C}_{\mathfrak{B}}^{\#}(X) &= \bigcup \mathfrak{F}_{\mathfrak{B}}^{\#}(X), \text{ where } \mathfrak{F}_{\mathfrak{B}}^{\#}(X) = \{Y \mid Y \in \mathfrak{B}, Y \cap X \neq \emptyset\}, \end{aligned}$$

if and only if the base system  $\mathfrak{B}$  is single-layered.

**5.14 Remark.** Of course, the equations

$$\mathfrak{C}_{\mathfrak{B}}^b(X) = \bigcup \mathfrak{F}_{\mathfrak{B}}^b(X) \text{ and } \mathfrak{C}_{\mathfrak{B}}^\sharp(X) = \bigcup \mathfrak{F}_{\mathfrak{B}}^\sharp(X)$$

trivially satisfy, they are just the definition of lower and upper  $\mathfrak{B}$ -approximations. Proposition 5.13, therefore, claims nothing else that there are no other set families  $\mathfrak{X}_1, \mathfrak{X}_2 \subseteq \mathfrak{B}$  satisfying the equations  $\mathfrak{C}_{\mathfrak{B}}^b(X) = \bigcup \mathfrak{X}_1$  and  $\mathfrak{C}_{\mathfrak{B}}^\sharp(X) = \bigcup \mathfrak{X}_2$  if and only if the base system  $\mathfrak{B}$  is single-layered.

**Proof.** Since  $\mathfrak{C}_{\mathfrak{B}}^b(X), \mathfrak{C}_{\mathfrak{B}}^\sharp(X) \in \mathfrak{D}_{\mathfrak{B}}$ , by Corollary 5.12, they are  $\mathfrak{B}$ -representable if and only if the base system  $\mathfrak{B}$  is single layered. And so, we only have to show that  $\mathfrak{F}_{\mathfrak{B}}(\mathfrak{C}_{\mathfrak{B}}^b(X))$  and  $\mathfrak{F}_{\mathfrak{B}}(\mathfrak{C}_{\mathfrak{B}}^\sharp(X))$  are of the forms

$$\begin{aligned} \mathfrak{F}_{\mathfrak{B}}(\mathfrak{C}_{\mathfrak{B}}^b(X)) &= \{Y \mid Y \in \mathfrak{B}, Y \subseteq X\}, \\ \mathfrak{F}_{\mathfrak{B}}(\mathfrak{C}_{\mathfrak{B}}^\sharp(X)) &= \{Y \mid Y \in \mathfrak{B}, Y \cap X \neq \emptyset\}. \end{aligned}$$

By Corollary 5.12 and Lemma 5.4, we have

$$\begin{aligned} \mathfrak{F}_{\mathfrak{B}}(\mathfrak{C}_{\mathfrak{B}}^b(X)) &= \{Y \mid Y \in \mathfrak{B}, Y \subseteq \mathfrak{C}_{\mathfrak{B}}^b(X)\} \\ &= \{Y \mid Y \in \mathfrak{B}, Y \subseteq \bigcup \{Y' \mid Y' \in \mathfrak{B}, Y' \subseteq X\}\} \\ &= \{Y \mid Y \in \mathfrak{B}, Y \in \{Y' \mid Y' \in \mathfrak{B}, Y' \subseteq X\}\} \\ &= \{Y \mid Y \in \mathfrak{B}, Y \subseteq X\}, \\ \mathfrak{F}_{\mathfrak{B}}(\mathfrak{C}_{\mathfrak{B}}^\sharp(X)) &= \{Y \mid Y \in \mathfrak{B}, Y \subseteq \mathfrak{C}_{\mathfrak{B}}^\sharp(X)\} \\ &= \{Y \mid Y \in \mathfrak{B}, Y \subseteq \bigcup \{Y' \mid Y' \in \mathfrak{B}, Y' \cap X \neq \emptyset\}\} \\ &= \{Y \mid Y \in \mathfrak{B}, Y \in \{Y' \mid Y' \in \mathfrak{B}, Y' \cap X \neq \emptyset\}\} \\ &= \{Y \mid Y \in \mathfrak{B}, Y \cap X \neq \emptyset\}. \end{aligned} \quad \square$$

## 5.6 Exactness in $\mathfrak{B}$ -approximation spaces

In Pawlak's approximation spaces the notions of ' $\varepsilon$ -crisp' and ' $\varepsilon$ -definable' are inherently one and the same, they are synonymous to each other.

The  $R$ -definable sets are those subsets of the universe which can be exactly defined in the knowledge base  $K$ , whereas the  $R$ -undefinable sets cannot be defined in this knowledge base.

The  $R$ -definable sets will be also called *R-exact* sets, and  $R$ -undefinable sets will be also said to be *R-inexact* or *R-rough*. ([55], p. 9. *The italics are the author's*. Here,  $R$  is an equivalence relation on a finite universe  $U$ , pp. 3–4.)

The equivalence of ‘ $\varepsilon$ -crisp’ and ‘ $\varepsilon$ -definable’ formally is drawn up by Proposition 4.14, point (1). Moreover, a subset  $X \subseteq U$  is  $\varepsilon$ -definable, and consequently  $\varepsilon$ -crisp as well, if and only if its lower  $\varepsilon$ -approximation is equal to its upper  $\varepsilon$ -approximation according to Proposition 4.14.

In our approach, however, the compound notion of ‘crisp’ and ‘definable’ splits into two parts.

**5.15 Definition.** Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\sharp \rangle$  be a  $\mathfrak{B}$ -approximation space and  $X \subseteq U$ . The subset  $X$  is  $\mathfrak{B}$ -crisp, if  $\mathfrak{C}_{\mathfrak{B}}^b(X) = \mathfrak{C}_{\mathfrak{B}}^\sharp(X)$ , otherwise  $X$  is  $\mathfrak{B}$ -rough.

**5.16 Definition.** Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\sharp \rangle$  be a  $\mathfrak{B}$ -approximation space and  $X \subseteq U$ . The set  $\mathfrak{N}_{\mathfrak{B}}(X) = \mathfrak{C}_{\mathfrak{B}}^\sharp(X) \setminus \mathfrak{C}_{\mathfrak{B}}^b(X)$  is called the  $\mathfrak{B}$ -boundary of  $X$ .

The  $\mathfrak{B}$ -boundary  $\mathfrak{N}_{\mathfrak{B}}(X)$  is not necessarily  $\mathfrak{B}$ -definable.

The next elementary facts are formulated in propositions for the sake of simple reference.

**5.17 Proposition.** Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\sharp \rangle$  be a  $\mathfrak{B}$ -approximation space and  $X \subseteq U$ .

The subset  $X$  is  $\mathfrak{B}$ -crisp if and only if the  $\mathfrak{B}$ -boundary  $\mathfrak{N}_{\mathfrak{B}}(X) = \emptyset$ .

**Proof.**  $\mathfrak{N}_{\mathfrak{B}}(X) = \mathfrak{C}_{\mathfrak{B}}^\sharp(X) \setminus \mathfrak{C}_{\mathfrak{B}}^b(X) = \emptyset \Leftrightarrow \mathfrak{C}_{\mathfrak{B}}^\sharp(X) \subseteq \mathfrak{C}_{\mathfrak{B}}^b(X)$ . However,  $\mathfrak{C}_{\mathfrak{B}}^b(X) \subseteq \mathfrak{C}_{\mathfrak{B}}^\sharp(X)$  always fulfills by the approximation property (C4), and so  $\mathfrak{C}_{\mathfrak{B}}^\sharp(X) \setminus \mathfrak{C}_{\mathfrak{B}}^b(X) = \emptyset \Leftrightarrow \mathfrak{C}_{\mathfrak{B}}^b(X) = \mathfrak{C}_{\mathfrak{B}}^\sharp(X)$ .  $\square$

**5.18 Proposition.** Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\sharp \rangle$  be a strong  $\mathfrak{B}$ -approximation space and  $X \subseteq U$ .

The subset  $X$  is  $\mathfrak{B}$ -crisp if and only if  $\mathfrak{C}_{\mathfrak{B}}^b(X) = \mathfrak{C}_{\mathfrak{B}}^\sharp(X) = X$ .

**Proof.** In strong  $\mathfrak{B}$ -approximation spaces every subset  $X \subseteq U$  is bounded by its lower and upper  $\mathfrak{B}$ -approximations:  $\mathfrak{C}_{\mathfrak{B}}^b(X) \subseteq X \subseteq \mathfrak{C}_{\mathfrak{B}}^\sharp(X)$ . And so, in strong  $\mathfrak{B}$ -approximation spaces a subset  $X$  is crisp  $\Leftrightarrow \mathfrak{C}_{\mathfrak{B}}^b(X) = \mathfrak{C}_{\mathfrak{B}}^\sharp(X) \Leftrightarrow \mathfrak{C}_{\mathfrak{B}}^b(X) = \mathfrak{C}_{\mathfrak{B}}^\sharp(X) = X$ .  $\square$

**5.19 Proposition.** Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\sharp \rangle$  be a strong  $\mathfrak{B}$ -approximation space and  $X \subseteq U$ . If  $X$  is  $\mathfrak{B}$ -crisp, then  $X$  is  $\mathfrak{B}$ -definable.

**Proof.** By Proposition 5.18  $X$  is  $\mathfrak{B}$ -crisp if and only if  $\mathfrak{C}_{\mathfrak{B}}^b(X) = \mathfrak{C}_{\mathfrak{B}}^\sharp(X) = X$ , and  $X$  is  $\mathfrak{B}$ -definable if and only if  $X = \mathfrak{C}_{\mathfrak{B}}^b(X)$  by Proposition 5.9 point (1).  $\square$

A subset  $X \in 2^U$  is  $\mathfrak{B}$ -definable if and only if  $X = \mathfrak{C}_{\mathfrak{B}}^b(X)$  by Proposition 5.9 point (1). However, as the next simple example shows,  $X = \mathfrak{C}_{\mathfrak{B}}^b(X)$  generally does not imply  $X = \mathfrak{C}_{\mathfrak{B}}^\sharp(X)$  even though the  $\mathfrak{B}$ -approximation space is a strong one. Let  $\mathfrak{B} = \{B_1, B_2\}$  be a base system, where  $B_1 \subsetneq B_2$  ( $B_1, B_2 \in 2^U$ ). Then  $\mathfrak{C}_{\mathfrak{B}}^b(B_1) = B_1 \subsetneq B_2 = \mathfrak{C}_{\mathfrak{B}}^\sharp(B_1)$ .

In other words, a  $\mathfrak{B}$ -definable subset is not necessarily  $\mathfrak{B}$ -crisp not even in strong  $\mathfrak{B}$ -approximation spaces. The converse statement only holds in strong  $\mathfrak{B}$ -approximation spaces by Proposition 5.19. Consequently, in our approach, the notions of ‘definable’ and ‘crisp’ are not synonymous to each other in Pawlak’s sense.

## 5.7 A Possible Interpretation of Our Approach

Let us suppose that we observe a collection of objects which is modelled as an abstract set, called the universe of discourse.

In real life, when we observe objects we cannot decide directly whether an object possesses a certain feature or not. Therefore we need a *tool* to be at our disposal with which we are able to judge easily and unambiguously whether an object possesses a *property* ascertained by the tool or not. It is expected that all tools can be used simply and quickly. The objects which are classified by a tool are modelled as a crisp subset of the universe. With a slight abuse of terminology, these subsets are simply called tools as well.

In sum, we model an object of interest as the element of an abstract set, called the universe, and the fact that ‘it possesses a property’ as ‘it is the element of a suitable crisp subset of the universe’.

Different tools usually form different subsets, but they are not necessarily disjoint. Notice that the complement of a tool is not necessarily a tool at the same time because the complement may not be used simply and quickly. For instance, let us take the tools being recursively enumerable. However, the complement of a recursively enumerable set is not necessarily recursively enumerable [50]. This significant fact confirms the partial nature of our approach [45].

Properties in  $\mathfrak{B}$  are our primary tools which serve as fundamental building blocks of knowledge about the universe. Properties in  $\mathfrak{D}_{\mathfrak{B}}$  are our derived tools which are formed from primary tools. To characterize any subset of the universe we want to use  $\mathfrak{D}_{\mathfrak{B}}$ . It is said that a property  $D \in \mathfrak{D}_{\mathfrak{B}}$  *characterizes* a subset  $X$  of the universe, if  $D \subseteq X$ , and  $X$  is *characterized in terms of*  $\mathfrak{D}_{\mathfrak{B}}$ , if  $X$  is  $\mathfrak{B}$ -definable.

However, apart from the derived tools themselves, any other subsets cannot be characterized in terms of  $\mathfrak{D}_{\mathfrak{B}}$ . Therefore, their description is replaced by a pair of derived tools, in particular, their lower and upper approximations.

The universe can be divided into the following parts by means of lower and upper approximations concerning a subset  $X \subseteq U$  [55, 58]:

- *$\mathfrak{B}$ -positive region of  $X$* :  $\mathfrak{C}_{\mathfrak{B}}^b(X) = \bigcup\{Y \mid Y \in \mathfrak{B}, Y \subseteq X\}$ , i.e. the lower  $\mathfrak{B}$ -approximation of  $X$ .  
 $\{Y \mid Y \in \mathfrak{B}, Y \subseteq X\}$  is the family of all properties which *certainly* characterize  $X$  with respect to the current derived tools  $\mathfrak{D}_{\mathfrak{B}}$ .
- *Upper  $\mathfrak{B}$ -approximation of  $X$* :  $\mathfrak{C}_{\mathfrak{B}}^\sharp(X) = \bigcup\{Y \mid Y \in \mathfrak{B}, Y \cap X \neq \emptyset\}$ .  
 $\{Y \mid Y \in \mathfrak{B}, Y \cap X \neq \emptyset\}$  is the family of all properties which *possibly* characterize  $X$  with respect to the current derived tools  $\mathfrak{D}_{\mathfrak{B}}$ .

- $\mathfrak{B}$ -negative region of  $X$ :  $\bigcup(\mathfrak{D}_{\mathfrak{B}} \setminus \{Y \mid Y \in \mathfrak{B}, Y \cap X \neq \emptyset\})$ .

$\mathfrak{D}_{\mathfrak{B}} \setminus \{Y \mid Y \in \mathfrak{B}, Y \cap X \neq \emptyset\}$  is the family of all properties which *certainly do not characterize*  $X$  with respect to the current derived tools  $\mathfrak{D}_{\mathfrak{B}}$ .

- $\mathfrak{B}$ -borderline region of  $X$ :

$$\bigcup(\{Y \mid Y \in \mathfrak{B}, Y \cap X \neq \emptyset\} \setminus \{Y \mid Y \in \mathfrak{B}, Y \subseteq X\}).$$

$\{Y \mid Y \in \mathfrak{B}, Y \cap X \neq \emptyset\} \setminus \{Y \mid Y \in \mathfrak{B}, Y \subseteq X\}$  is the family of all properties which cannot be classified with certainty either as characterizing  $X$  or as not characterizing  $X$  with respect to the current derived tools  $\mathfrak{D}_{\mathfrak{B}}$ .

## Chapter 6

# Galois Connections

Recall that for any arbitrary binary relation  $\epsilon$  on  $U$ , the pairs of maps  $(2^U, \bar{\epsilon}, \underline{\epsilon^{-1}}, 2^U)$  and  $(2^U, \overline{\epsilon^{-1}}, \underline{\epsilon}, 2^U)$  are Galois connections (Proposition 4.20). Especially, when  $\varepsilon$  is an equivalence relation on  $U$ , the upper and lower  $\varepsilon$ -approximations form a  $\mathbb{G}(2^U, \bar{\varepsilon}, \underline{\varepsilon}, 2^U)$  Galois connection. Note that the left adjoint is the upper  $\varepsilon$ -approximation  $\bar{\varepsilon}$  and the right adjoint is the lower  $\varepsilon$ -approximation  $\underline{\varepsilon}$ . Some further observations about upper and lower approximations as Galois connections see, e.g. [31, 32, 34, 51].

Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\# \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong  $\mathfrak{B}$ -approximation space. In this Section we will investigate what conditions have to be satisfied by a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong  $\mathfrak{B}$ -approximation space so that the pair of maps  $(2^U, \mathfrak{C}_{\mathfrak{B}}^\#, \mathfrak{C}_{\mathfrak{B}}^b, 2^U)$  forms a Galois connection on  $(2^U, \subseteq)$ . To do this, we take up the assertions of Proposition 2.2 and examine the conditions under which they hold point by point.

### 6.1 Regular Galois Connection

Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\# \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong  $\mathfrak{B}$ -approximation space.

The maps  $\mathfrak{C}_{\mathfrak{B}}^\#$  and  $\mathfrak{C}_{\mathfrak{B}}^b$  are trivially monotone, i.e. the point (2) in Proposition 2.2 immediately holds. Thus we have to examine only the point (1) in Proposition 2.2 in detail.

Next proposition answers the *first half* of the point (1) in Proposition 2.2.

**6.1 Proposition. (Csajbók [14], Theorem 20)** *Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\# \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong  $\mathfrak{B}$ -approximation space.*

Then  $\forall X \in 2^U (X \subseteq \mathfrak{C}_{\mathfrak{B}}^b(\mathfrak{C}_{\mathfrak{B}}^\sharp(X)))$  if and only if  $\bigcup \mathfrak{B} = U$ .

**6.2 Remark.** The proposition does not require that the base system  $\mathfrak{B}$  should be single-layered.

**Proof.** ( $\Rightarrow$ ) By contradiction, let us assume that  $\bigcup \mathfrak{B} \neq U$ . Accordingly,  $\exists X' (\neq \emptyset) \subseteq U \setminus \bigcup \mathfrak{B}$ . Hence,  $\mathfrak{C}_{\mathfrak{B}}^b(\mathfrak{C}_{\mathfrak{B}}^\sharp(X')) = \emptyset$ , which gives  $\emptyset \neq X' \subseteq \mathfrak{C}_{\mathfrak{B}}^b(\mathfrak{C}_{\mathfrak{B}}^\sharp(X')) = \emptyset$ , a contradiction.

( $\Leftarrow$ )  $\mathfrak{C}_{\mathfrak{B}}^\sharp(X) \in \mathfrak{D}_{\mathfrak{B}}$ , and so, by Proposition 5.8 point (3),  $\mathfrak{C}_{\mathfrak{B}}^b(\mathfrak{C}_{\mathfrak{B}}^\sharp(X)) = \mathfrak{C}_{\mathfrak{B}}^\sharp(X)$ . Since  $\bigcup \mathfrak{B} = U$ , by Proposition 5.7 condition (C6),  $\mathfrak{C}_{\mathfrak{B}}^\sharp$  is extensive, thus  $X \subseteq \mathfrak{C}_{\mathfrak{B}}^\sharp(X) = \mathfrak{C}_{\mathfrak{B}}^b(\mathfrak{C}_{\mathfrak{B}}^\sharp(X))$ .  $\square$

Let us take up the question of the *second half* of the point (1) in Proposition 2.2. In general, it also does not hold.

**6.3 Proposition.** (Csajbók [14], Theorem 21) *Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\sharp \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong  $\mathfrak{B}$ -approximation space, and let us assume that the base system  $\mathfrak{B}$  is single-layered.*

*Then*

$$\forall X \in 2^U (\mathfrak{C}_{\mathfrak{B}}^\sharp(\mathfrak{C}_{\mathfrak{B}}^b(X)) \subseteq X)$$

*if and only if the  $\mathfrak{B}$ -sets are pairwise disjoint.*

**Proof.** ( $\Rightarrow$ ) Let us suppose, by contradiction, that the  $\mathfrak{B}$ -sets are not pairwise disjoint. If so,

$$\exists B_1, B_2 \in \mathfrak{B} (B_1 \neq B_2 \wedge B_1 \cap B_2 \neq \emptyset),$$

where neither  $B_1 \subseteq B_2$  nor  $B_2 \subseteq B_1$  holds because of the base system  $\mathfrak{B}$  is single-layered. Hence, e.g. for  $B_1$ , we get

$$\mathfrak{C}_{\mathfrak{B}}^\sharp(\mathfrak{C}_{\mathfrak{B}}^b(B_1)) = \mathfrak{C}_{\mathfrak{B}}^\sharp(B_1) \supseteq B_1 \cup B_2 \not\subseteq B_1, \text{ a contradiction.}$$

( $\Leftarrow$ ) If  $X = \emptyset$ , then  $\mathfrak{C}_{\mathfrak{B}}^\sharp(\mathfrak{C}_{\mathfrak{B}}^b(\emptyset)) = \mathfrak{C}_{\mathfrak{B}}^\sharp(\emptyset) = \emptyset \subseteq \emptyset$  trivially holds (independently of the  $\mathfrak{B}$ -sets are pairwise disjoint or not).

Let  $\emptyset \neq X \in 2^U$ .

If  $\mathfrak{C}_{\mathfrak{B}}^b(X) = \emptyset$ , then  $\mathfrak{C}_{\mathfrak{B}}^\sharp(\emptyset) = \emptyset \subseteq X$ .

Let  $\emptyset \neq \mathfrak{C}_{\mathfrak{B}}^b(X) = \bigcup \mathfrak{B}' \subseteq X$  for a family of  $\mathfrak{B}$ -sets  $\mathfrak{B}' \subseteq \mathfrak{B}$  (such a  $\mathfrak{B}'$  exists because  $\mathfrak{C}_{\mathfrak{B}}^b(X)$  is  $\mathfrak{B}$ -definable and  $\mathfrak{C}_{\mathfrak{B}}^b$  is contractive). Since the  $\mathfrak{B}$ -sets are pairwise disjoint,

$$\{Y \mid Y \in \mathfrak{B}, Y \cap \mathfrak{C}_{\mathfrak{B}}^b(X) \neq \emptyset\} = \{Y \mid Y \in \mathfrak{B}, Y \subseteq \bigcup \mathfrak{C}_{\mathfrak{B}}^b(X)\}.$$



Hence, we get

$$\begin{aligned}
\mathfrak{C}_{\mathfrak{B}}^{\sharp}(\mathfrak{C}_{\mathfrak{B}}^{\flat}(X)) &= \bigcup \{Y \mid Y \in \mathfrak{B}, Y \cap \mathfrak{C}_{\mathfrak{B}}^{\flat}(X) \neq \emptyset\} \\
&= \bigcup \{Y \mid Y \in \mathfrak{B}, Y \subseteq \mathfrak{C}_{\mathfrak{B}}^{\flat}(X)\} \\
&= \mathfrak{C}_{\mathfrak{B}}^{\flat}(\mathfrak{C}_{\mathfrak{B}}^{\sharp}(X)) = \mathfrak{C}_{\mathfrak{B}}^{\flat}(X) \subseteq X.
\end{aligned}
\quad \square$$

**6.4 Proposition.** (Csajbók [14], Theorem 22) *Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^{\flat}, \mathfrak{C}_{\mathfrak{B}}^{\sharp} \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^{\flat}$ -semi-strong  $\mathfrak{B}$ -approximation space, and let us assume that the base system  $\mathfrak{B}$  is single-layered.*

*The pair of maps  $(2^U, \mathfrak{C}_{\mathfrak{B}}^{\sharp}, \mathfrak{C}_{\mathfrak{B}}^{\flat}, 2^U)$  forms a Galois connection on  $(2^U, \subseteq)$  if and only if the base system  $\mathfrak{B}$  is a partition of  $U$ .*

**Proof.** The maps  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}$  and  $\mathfrak{C}_{\mathfrak{B}}^{\flat}$  are monotone, and so by Proposition 6.1 and Proposition 6.3, the conditions in Proposition 2.2 satisfy.  $\square$

According to Proposition 6.4, the Galois connection between the pair of maps  $(2^U, \mathfrak{C}_{\mathfrak{B}}^{\sharp}, \mathfrak{C}_{\mathfrak{B}}^{\flat}, 2^U)$  was proved under the condition that the base system  $\mathfrak{B}$  is single-layered. However, as we have seen in Proposition 6.1, the fulfilment of the first half of the point (1) in Proposition 2.2 does not require that the base system  $\mathfrak{B}$  to be single-layered. Now we examine whether the condition that the base system  $\mathfrak{B}$  is single-layered can be removed from Proposition 6.3.

First we need the following lemma.

**6.5 Lemma.** (Csajbók [12], Lemma 4.11) *Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^{\flat}, \mathfrak{C}_{\mathfrak{B}}^{\sharp} \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^{\flat}$ -semi-strong  $\mathfrak{B}$ -approximation space. If*

$$\forall X \in 2^U (\mathfrak{C}_{\mathfrak{B}}^{\sharp}(\mathfrak{C}_{\mathfrak{B}}^{\flat}(X)) \subseteq X),$$

*the base system  $\mathfrak{B}$  is singled-layered.*

**6.6 Remark.** The converse statement does not hold. Let  $\mathfrak{B} = \{B_1, B_2\}$  be a base system such that  $B_1 \cap B_2 \neq \emptyset$  but  $B_1 \not\subseteq B_2 \wedge B_2 \not\subseteq B_1$ . Clearly,  $\mathfrak{B}$  is single-layered, and, e.g.  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}(\mathfrak{C}_{\mathfrak{B}}^{\flat}(B_1)) = \mathfrak{C}_{\mathfrak{B}}^{\sharp}(B_1) = B_1 \cup B_2 \not\subseteq B_1$ .

**Proof.** Let us suppose, by contradiction, that  $\mathfrak{B}$  is not singled-layered. If so,  $\exists B \in \mathfrak{B} \wedge \exists \mathfrak{B}' \subseteq \mathfrak{B} \setminus \{B\} (B \subseteq \bigcup \mathfrak{B}')$ . Hence,  $B \subseteq \bigcup \mathfrak{B}'$  but  $B \not\subseteq \mathfrak{B}'$ , and so there exists at least one  $B \neq B' \in \mathfrak{B}'$  such that  $B' \cap B \neq \emptyset$ .

We have to distinguish three cases:

Case (1)  $B \subsetneq B'$ :  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}(\mathfrak{C}_{\mathfrak{B}}^{\flat}(B)) = \mathfrak{C}_{\mathfrak{B}}^{\sharp}(B) \supseteq B' \not\supseteq B$ , a contradiction.

Case (2)  $B' \subsetneq B$ :  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}(\mathfrak{C}_{\mathfrak{B}}^{\flat}(B')) = \mathfrak{C}_{\mathfrak{B}}^{\sharp}(B') \supseteq B \not\supseteq B'$ , a contradiction.

Case (3)  $B' \cap B \neq \emptyset$ , but neither  $B \subsetneq B'$  nor  $B' \subsetneq B$  holds:  
 $\mathfrak{C}_{\mathfrak{B}}^{\sharp}(\mathfrak{C}_{\mathfrak{B}}^{\flat}(B')) = \mathfrak{C}_{\mathfrak{B}}^{\sharp}(B') \supseteq B \cup B' \not\supseteq B'$ , a contradiction.  $\square$

**6.7 Proposition.** (Csajbók [12], Proposition 4.13) *Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^{\flat}, \mathfrak{C}_{\mathfrak{B}}^{\sharp} \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^{\flat}$ -semi-strong  $\mathfrak{B}$ -approximation space.*

*Then*

$$\forall X \in 2^U (\mathfrak{C}_{\mathfrak{B}}^{\sharp}(\mathfrak{C}_{\mathfrak{B}}^{\flat}(X)) \subseteq X)$$

*if and only if the  $\mathfrak{B}$ -sets are pairwise disjoint.*

**Proof.** ( $\Rightarrow$ ) The base system  $\mathfrak{B}$  is single-layered by Lemma 6.5. Hereafter the proof is the same as in Proposition 6.3.

( $\Leftarrow$ ) The  $\mathfrak{B}$ -sets are pairwise disjoint which immediately implies that the the base system  $\mathfrak{B}$  is single-layered. Hereafter the proof is the same as in Proposition 6.3.  $\square$

**6.8 Theorem.** (Csajbók [12], Theorem 4.14) *Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^{\flat}, \mathfrak{C}_{\mathfrak{B}}^{\sharp} \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^{\flat}$ -semi-strong  $\mathfrak{B}$ -approximation space.*

*The pair of maps  $(2^U, \mathfrak{C}_{\mathfrak{B}}^{\sharp}, \mathfrak{C}_{\mathfrak{B}}^{\flat}, 2^U)$  forms a Galois connection on  $(2^U, \subseteq)$  if and only if the base system  $\mathfrak{B}$  is a partition of  $U$ .*

**Proof.** The maps  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}$  and  $\mathfrak{C}_{\mathfrak{B}}^{\flat}$  are monotone, and so by Proposition 6.1 and Proposition 6.7, the conditions in Proposition 2.2 satisfy.  $\square$

## 6.2 Partial Galois Connection

### 6.2.1 On Partial Lower $\mathfrak{B}$ -approximations

If a nonempty  $X \in 2^U$  does not contain nonempty  $\mathfrak{B}$ -definable subsets, then  $\mathfrak{C}_{\mathfrak{B}}^{\flat}(X) = \bigcup \emptyset = \emptyset \subseteq X$  holds—which, however, does not provide new information about the relationship between  $X$  and  $\mathfrak{B}$ . This phenomenon appears in Pawlak's classic rough set theory, too.

**6.9 Definition.** Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^{\flat}, \mathfrak{C}_{\mathfrak{B}}^{\sharp} \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^{\flat}$ -semi-strong  $\mathfrak{B}$ -approximation space, and  $X$  be any subset of  $U$ .

The *partial lower  $\mathfrak{B}$ -approximation* of  $X$  is

$$\partial \mathfrak{C}_{\mathfrak{B}}^{\flat}(X) = \begin{cases} \mathfrak{C}_{\mathfrak{B}}^{\flat}(X), & \text{if } X = \emptyset \vee (X \neq \emptyset \wedge \mathfrak{C}_{\mathfrak{B}}^{\flat}(X) \neq \emptyset); \\ \text{undefined,} & \text{otherwise.} \end{cases} \quad (6.2.1)$$

**6.10 Remark.** Note that the formula  $X = \emptyset \vee (X \neq \emptyset \wedge \mathfrak{C}_{\mathfrak{B}}^b(X) \neq \emptyset)$  is equal to the formula  $X = \emptyset \vee \mathfrak{C}_{\mathfrak{B}}^b(X) \neq \emptyset$ . The latter is used in the following because it is a simpler formula than the former.

There exists at least one nonempty  $B \in \mathfrak{B}$   $\mathfrak{B}$ -set by Definition 5.1. Then  $\mathfrak{C}_{\mathfrak{B}}^b(B) = B \neq \emptyset$  according to Definition 5.6. Hence,  $\partial\mathfrak{C}_{\mathfrak{B}}^b$  is defined on at least one nonempty subset of  $U$ .

If  $X \in 2^U$  is nonempty and its lower  $\mathfrak{B}$ -approximation  $\mathfrak{C}_{\mathfrak{B}}^b(X)$  is empty at the same time, then its partial lower  $\mathfrak{B}$ -approximation  $\partial\mathfrak{C}_{\mathfrak{B}}^b(X)$  is undefined by Definition 6.9. This implies that the map  $\partial\mathfrak{C}_{\mathfrak{B}}^b$  is total only if the base system  $\mathfrak{B}$  contains all singleton sets  $\{x\}$  ( $x \in U$ ), in other words, if all singletons are  $\mathfrak{B}$ -definable. This is a rather special situation as well. That is to exclude that we allow the empty set to be the lower  $\mathfrak{B}$ -approximation of a nonempty subset of  $U$  is problematic as well.

A natural total extension of  $\partial\mathfrak{C}_{\mathfrak{B}}^b$  is the lower  $\mathfrak{B}$ -approximation  $\mathfrak{C}_{\mathfrak{B}}^b$ . That is the map  $\partial\mathfrak{C}_{\mathfrak{B}}^b$  can be made total if it is allowed that the empty set may be the lower  $\mathfrak{B}$ -approximation of a nonempty subset of  $U$ . Of course, any extension  $\mathfrak{C}_{\mathfrak{B}}^*$  of  $\partial\mathfrak{C}_{\mathfrak{B}}^b$  also has to be  $\mathfrak{B}$ -definable and contractive, i.e. formally, the condition  $\forall X \in 2^U (\mathfrak{C}_{\mathfrak{B}}^*(X) \in \mathfrak{D}_{\mathfrak{B}} \wedge \mathfrak{C}_{\mathfrak{B}}^*(X) \subseteq X)$  has to be fulfilled by  $\mathfrak{C}_{\mathfrak{B}}^*$ . Under the previous assumptions, we will show that any extension of this type is unique.

**6.11 Proposition.** *Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\# \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong  $\mathfrak{B}$ -approximation space, and  $X$  be any subset of  $U$ .*

*The total extension  $\mathfrak{C}_{\mathfrak{B}}^b$  of  $\partial\mathfrak{C}_{\mathfrak{B}}^b$  is unique under the conditions that*

- (1) *the empty set may be the lower  $\mathfrak{B}$ -approximation of nonempty subsets of  $U$ , and*
- (2)  *$\forall X \in 2^U (\mathfrak{C}_{\mathfrak{B}}^*(X) \in \mathfrak{D}_{\mathfrak{B}} \wedge \mathfrak{C}_{\mathfrak{B}}^*(X) \subseteq X)$  has to be fulfilled by any total extension  $\mathfrak{C}_{\mathfrak{B}}^*$  of  $\mathfrak{C}_{\mathfrak{B}}^b$ .*

**Proof.** It is straightforward that  $\mathfrak{C}_{\mathfrak{B}}^b$  is a total extension of  $\partial\mathfrak{C}_{\mathfrak{B}}^b$  from  $\text{dom } \partial\mathfrak{C}_{\mathfrak{B}}^b$  to  $2^U$ , and the points (1) and (2) automatically satisfy.

In order to prove the uniqueness, let us suppose, by contradiction, that  $\mathfrak{C}_{\mathfrak{B}}^*$  is an extension of  $\partial\mathfrak{C}_{\mathfrak{B}}^b$  from  $\text{dom } \partial\mathfrak{C}_{\mathfrak{B}}^b$  to  $2^U$  which differs from  $\mathfrak{C}_{\mathfrak{B}}^b$  and

$$\forall X \in 2^U (\mathfrak{C}_{\mathfrak{B}}^*(X) \in \mathfrak{D}_{\mathfrak{B}} \wedge \mathfrak{C}_{\mathfrak{B}}^*(X) \subseteq X)$$

holds.

Since  $\mathfrak{C}_{\mathfrak{B}}^*$  is an extension of  $\partial\mathfrak{C}_{\mathfrak{B}}^b$ , thus  $\mathfrak{C}_{\mathfrak{B}}^* = \partial\mathfrak{C}_{\mathfrak{B}}^b = \mathfrak{C}_{\mathfrak{B}}^b$  on  $\text{dom } \partial\mathfrak{C}_{\mathfrak{B}}^b$ , i.e. when  $X = \emptyset \vee \mathfrak{C}_{\mathfrak{B}}^b(X) \neq \emptyset$  satisfies (see Remark 6.10). On the other hand,  $\mathfrak{C}_{\mathfrak{B}}^*$

differs from  $\mathfrak{C}_{\mathfrak{B}}^b$ , thus there exists at least one nonempty  $X' \in 2^U \setminus \text{dom } \partial\mathfrak{C}_{\mathfrak{B}}^b$  such that  $\mathfrak{C}_{\mathfrak{B}}^*(X') \neq \mathfrak{C}_{\mathfrak{B}}^b(X')$ .

From the formula

$$X \in 2^U \setminus \text{dom } \partial\mathfrak{C}_{\mathfrak{B}}^b \Leftrightarrow \neg(X = \emptyset \vee \mathfrak{C}_{\mathfrak{B}}^b(X) \neq \emptyset) \Leftrightarrow X \neq \emptyset \wedge \mathfrak{C}_{\mathfrak{B}}^b(X) = \emptyset,$$

we get that  $\mathfrak{C}_{\mathfrak{B}}^b(X) = \emptyset$  for every nonempty subset  $X \in 2^U \setminus \text{dom } \partial\mathfrak{C}_{\mathfrak{B}}^b$ . In particular,  $\mathfrak{C}_{\mathfrak{B}}^*(X') \neq \mathfrak{C}_{\mathfrak{B}}^b(X') = \emptyset$ .

Since  $\emptyset \neq \mathfrak{C}_{\mathfrak{B}}^*(X') \in \mathfrak{D}_{\mathfrak{B}}$ , there exists a nonempty family of sets  $\mathfrak{B}' \subseteq \mathfrak{B}$  such that  $\mathfrak{C}_{\mathfrak{B}}^*(X') = \bigcup \mathfrak{B}' \subseteq X'$ . Hence,

$$\begin{aligned} \emptyset \neq \mathfrak{C}_{\mathfrak{B}}^*(X') = \bigcup \mathfrak{B}' &= \bigcup \{Y \mid Y \in \mathfrak{B}', Y \subseteq X'\} \\ &\subseteq \bigcup \{Y \mid Y \in \mathfrak{B}, Y \subseteq X'\} \\ &= \mathfrak{C}_{\mathfrak{B}}^b(X') = \emptyset, \end{aligned}$$

which is a contradiction.  $\square$

### 6.2.2 Partial Upper $\mathfrak{B}$ -approximations

According to Proposition 5.7,  $\mathfrak{C}_{\mathfrak{B}}^\#$  is extensive if and only if the base system  $\mathfrak{B}$  covers the universe. Hence, if  $\bigcup \mathfrak{B} \neq U$ , then

$$\forall X \subseteq U \setminus \bigcup \mathfrak{B} \quad \forall B \in \mathfrak{B} \quad (X \cap B = \emptyset).$$

Consequently,  $\mathfrak{C}_{\mathfrak{B}}^\#(X) = \bigcup \emptyset = \emptyset$  for all subsets of this type. In other words, the empty set may be the upper  $\mathfrak{B}$ -approximation of certain nonempty subsets of  $U$ . Indeed, if  $\mathfrak{C}_{\mathfrak{B}}^\#(X) \neq \emptyset$ , then  $X \not\subseteq \mathfrak{C}_{\mathfrak{B}}^\#(X)$  is also possible.

**6.12 Definition.**  $X$  is  $\mathfrak{B}$ -approximatable if  $X \subseteq \mathfrak{C}_{\mathfrak{B}}^\#(X)$ , otherwise it is said that  $X$  has a  $\mathfrak{B}$ -approximation gap.

The  $\mathfrak{B}$ -approximation gap may be interpreted so that our knowledge about the universe encoded in the base system is incomplete and not enough to approximate  $X$ . This phenomenon may be natural/necessary or not. In the latter case, in order to fulfill the inclusion  $X \subseteq \mathfrak{C}_{\mathfrak{B}}^\#(X)$  as far as possible, the base system  $\mathfrak{B}$  has to be augmented via taking into account additional features concerning the observed system. In both former and latter cases, another possible solution is that the upper  $\mathfrak{B}$ -approximation map is defined as a partial one excluding the  $\mathfrak{B}$ -approximation gaps.

**6.13 Definition.** Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\# \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong  $\mathfrak{B}$ -approximation space, and  $X$  be any subset of  $U$ .

The *partial upper  $\mathfrak{B}$ -approximation* of  $X$  is

$$\partial \mathfrak{C}_{\mathfrak{B}}^\#(X) = \begin{cases} \mathfrak{C}_{\mathfrak{B}}^\#(X), & \text{if } X \text{ is } \mathfrak{B}\text{-approximatable;} \\ \text{undefined,} & \text{otherwise.} \end{cases} \quad (6.2.2)$$

There exists at least one nonempty  $B \in \mathfrak{B}$   $\mathfrak{B}$ -set by Definition 5.1. Then  $B \subseteq \mathfrak{C}_{\mathfrak{B}}^\#(B)$  according to Definition 5.6. Hence,  $\partial \mathfrak{C}_{\mathfrak{B}}^\#$  is defined on at least one nonempty subset of  $U$ .

Notice that  $\mathfrak{C}_{\mathfrak{B}}^b(X) \subseteq X \subseteq \partial \mathfrak{C}_{\mathfrak{B}}^\#(X)$  holds provided  $X$  is  $\mathfrak{B}$ -approximatable, i.e. on  $\text{dom } \partial \mathfrak{C}_{\mathfrak{B}}^\#$ .

As Theorem 6.8 shows, the pair of maps  $(2^U, \mathfrak{C}_{\mathfrak{B}}^\#, \mathfrak{C}_{\mathfrak{B}}^b, 2^U)$  forms a Galois connection on  $(2^U, \subseteq)$  if and only if the base system  $\mathfrak{B}$  is a partition of  $U$ . The question naturally arises whether the Galois connection generalize so that the pair of maps  $(2^U, \partial \mathfrak{C}_{\mathfrak{B}}^\#, \mathfrak{C}_{\mathfrak{B}}^b, 2^U)$  may form a Galois connection in some sense. Moreover, if the answer is yes, what conditions have to be fulfilled by a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong  $\mathfrak{B}$ -approximation space  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\# \rangle$  so that  $(2^U, \partial \mathfrak{C}_{\mathfrak{B}}^\#, \mathfrak{C}_{\mathfrak{B}}^b, 2^U)$  forms a Galois connection of this special type. Recall that  $\mathfrak{C}_{\mathfrak{B}}^b$  is a total and  $\partial \mathfrak{C}_{\mathfrak{B}}^\#$  is a partial map on  $2^U$ , and so the notion of the *partial* Galois connection which is drawn up in Definition 2.4 may be suitable for our purpose. In the following, we take up the points (1)–(4) in Definition 2.4 and examine the conditions under which they hold point by point.

Clearly, the map  $\partial \mathfrak{C}_{\mathfrak{B}}^\#$  is a monotone partial map and  $\mathfrak{C}_{\mathfrak{B}}^b$  is a monotone total map. Thus the points (1) and (2) in Definition 2.4 immediately holds. Thus we only have to examine the point (3) and (4) in Definition 2.4 in detail.

Next proposition answers the condition (3) in Definition 2.4.

**6.14 Proposition. (Csajbók [14], Theorem 25)** Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\# \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong  $\mathfrak{B}$ -approximation space.

Then  $\partial \mathfrak{C}_{\mathfrak{B}}^\#(\mathfrak{C}_{\mathfrak{B}}^b(X))$  is defined for all  $X \in 2^U$ .

**Proof.** Let  $X \in 2^U$  be an arbitrary subset of  $U$ . By the idempotency property of  $\mathfrak{C}_{\mathfrak{B}}^b(X)$  (Proposition 5.8 point (2)),  $\mathfrak{C}_{\mathfrak{B}}^b(\mathfrak{C}_{\mathfrak{B}}^b(X)) = \mathfrak{C}_{\mathfrak{B}}^b(X)$ . Thus,

$$\begin{aligned} \mathfrak{C}_{\mathfrak{B}}^b(X) = \mathfrak{C}_{\mathfrak{B}}^b(\mathfrak{C}_{\mathfrak{B}}^b(X)) &= \bigcup \{Y \mid Y \in \mathfrak{B}, Y \subseteq \mathfrak{C}_{\mathfrak{B}}^b(X)\} \\ &\subseteq \bigcup \{Y \mid Y \in \mathfrak{B}, Y \cap \mathfrak{C}_{\mathfrak{B}}^b(X) \neq \emptyset\} \\ &= \mathfrak{C}_{\mathfrak{B}}^\#(\mathfrak{C}_{\mathfrak{B}}^b(X)), \end{aligned}$$

that is, by Definition 6.13,  $\partial\mathfrak{C}_{\mathfrak{B}}^{\sharp}(\mathfrak{C}_{\mathfrak{B}}^b(X))$  is defined.  $\square$

Next two propositions deal with the condition (4) in Definition 2.4.

**6.15 Proposition.** (Csajbók [14], Theorem 26) *Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^{\sharp} \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong  $\mathfrak{B}$ -approximation space.*

*Then for all  $\mathfrak{B}$ -approximatable subsets  $X \in 2^U$  and all subsets  $Y \in 2^U$ :*

$$\partial\mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) \subseteq Y \Rightarrow X \subseteq \mathfrak{C}_{\mathfrak{B}}^b(Y).$$

**Proof.** Let  $X, Y \in 2^U$  be two subsets of  $U$  such that  $X$  is  $\mathfrak{B}$ -approximatable. Then  $X \subseteq \mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) = \partial\mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) \subseteq Y$ . Hence, by the Pawlak's type and monotonicity properties of  $\mathfrak{C}_{\mathfrak{B}}^b$ , we get

$$X \subseteq \mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) = \mathfrak{C}_{\mathfrak{B}}^b(\mathfrak{C}_{\mathfrak{B}}^{\sharp}(X)) \subseteq \mathfrak{C}_{\mathfrak{B}}^b(Y). \quad \square$$

**6.16 Lemma.** (Csajbók [12], Lemma 4.19) *Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^{\sharp} \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong  $\mathfrak{B}$ -approximation space. If*

$$X \subseteq \mathfrak{C}_{\mathfrak{B}}^b(Y) \Rightarrow \partial\mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) \subseteq Y$$

*holds for all  $\mathfrak{B}$ -approximatable subsets  $X \in 2^U$  and all subsets  $Y \in 2^U$ , the base system  $\mathfrak{B}$  is singled-layered.*

**6.17 Remark.** The converse statement does not hold. Let  $\mathfrak{B} = \{B_1, B_2\}$  be a base system such that  $B_1 \cap B_2 \neq \emptyset$  but  $B_1 \not\subseteq B_2 \wedge B_2 \not\subseteq B_1$ . Clearly,  $\mathfrak{B}$  is single-layered. Let  $X \in 2^U$  such that  $X \subsetneq B_1$ ,  $X \cap B_2 \neq \emptyset$  but  $X \not\subseteq B_2$ . Then  $X \subseteq B_1 \cup B_2 = \mathfrak{C}_{\mathfrak{B}}^{\sharp}(X)$ , i.e.  $X$  is  $\mathfrak{B}$ -approximatable. Hence,  $X \subseteq \mathfrak{C}_{\mathfrak{B}}^b(B_1) = B_1$ , but  $\partial\mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) = \mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) = B_1 \cup B_2 \not\subseteq B_1$ .

**Proof.** First, we note that, if for all  $\mathfrak{B}$ -approximatable subsets  $X \in 2^U$  and all subsets  $Y \in 2^U$ , the relationship  $X \subseteq \mathfrak{C}_{\mathfrak{B}}^b(Y) \Rightarrow \partial\mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) \subseteq Y$  is satisfied, then, of course, for all  $\mathfrak{B}$ -approximatable subsets  $X \in 2^U$ ,

$$X \subseteq \mathfrak{C}_{\mathfrak{B}}^b(X) \Rightarrow \partial\mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) \subseteq X$$

also has to be satisfied.

Since  $\mathfrak{C}_{\mathfrak{B}}^b$  is contractive, then  $\mathfrak{C}_{\mathfrak{B}}^b(X) \subseteq X$  holds for all subsets  $X \in 2^U$ . Thus, if  $X \subseteq \mathfrak{C}_{\mathfrak{B}}^b(X)$ , then  $\mathfrak{C}_{\mathfrak{B}}^b(X) = X$ . Consequently, by Proposition 5.9 point (1),  $X$  is  $\mathfrak{B}$ -definable, i.e.  $X \in \mathfrak{D}_{\mathfrak{B}}$ .

On the other hand, for all  $\mathfrak{B}$ -approximatable subsets  $X \in 2^U$ ,  $X \subseteq \mathfrak{C}_{\mathfrak{B}}^{\sharp}(X)$ . Thus if for all  $\mathfrak{B}$ -approximatable subsets  $X \in 2^U$ ,  $\partial\mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) = \mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) \subseteq X$  also holds, then  $X = \mathfrak{C}_{\mathfrak{B}}^{\sharp}(X)$ .

For all these reasons, we can restate the previous statement as follows. For all  $\mathfrak{B}$ -approximatable subsets  $X \in 2^U$ ,

$$X \in \mathfrak{D}_{\mathfrak{B}} \Rightarrow X = \mathfrak{C}_{\mathfrak{B}}^{\sharp}(X).$$

has to be satisfied.

Now, let us suppose, by contradiction, that  $\mathfrak{B}$  is not singled-layered, that is,  $\exists B \in \mathfrak{B} \wedge \exists \mathfrak{B}' \subseteq \mathfrak{B} \setminus \{B\} (B \subseteq \bigcup \mathfrak{B}')$ . Hence,  $B \subseteq \bigcup \mathfrak{B}'$ , but  $B \notin \mathfrak{B}'$ , and so there exists at least one  $B \neq B' \in \mathfrak{B}'$  such that  $B' \cap B \neq \emptyset$ . Of course,  $B, B' \in \mathfrak{D}_{\mathfrak{B}}$ , and  $B \subseteq \mathfrak{C}_{\mathfrak{B}}^{\sharp}(B)$ ,  $B' \subseteq \mathfrak{C}_{\mathfrak{B}}^{\sharp}(B')$ , i.e.  $B, B'$  are  $\mathfrak{B}$ -approximatable.

We have to distinguish three cases:

**Case (1)** If  $B \subsetneq B'$ , then  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}(B) \supseteq B' \supsetneq B$ , a contradiction.

**Case (2)** If  $B' \subsetneq B$ , then  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}(B') \supseteq B \supsetneq B'$ , a contradiction.

**Case (3)** If  $B' \cap B \neq \emptyset$ , but neither  $B \subsetneq B'$  nor  $B' \subsetneq B$  holds, then  $\mathfrak{C}_{\mathfrak{B}}^{\sharp}(B') \supseteq B \cup B' \supsetneq B'$ , a contradiction.  $\square$

**6.18 Proposition.** (Csajbók [12], Proposition 4.21) *Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^{\flat}, \mathfrak{C}_{\mathfrak{B}}^{\sharp} \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^{\flat}$ -semi-strong  $\mathfrak{B}$ -approximation space.*

*Then for all  $\mathfrak{B}$ -approximatable subsets  $X \in 2^U$  and all subsets  $Y \in 2^U$ ,*

$$X \subseteq \mathfrak{C}_{\mathfrak{B}}^{\flat}(Y) \Rightarrow \partial\mathfrak{C}_{\mathfrak{B}}^{\sharp}(X) \subseteq Y,$$

*if and only if the  $\mathfrak{B}$ -sets are pairwise disjoint.*

**Proof.** ( $\Rightarrow$ ) Let us suppose, by contradiction, that the  $\mathfrak{B}$ -sets are not pairwise disjoint, If so,

$$\exists B_1, B_2 \in \mathfrak{B} (B_1 \neq B_2 \wedge B_1 \cap B_2 \neq \emptyset).$$

By Lemma 6.16, the base system  $\mathfrak{B}$  is single-layered, and so neither  $B_1 \subseteq B_2$  nor  $B_2 \subseteq B_1$  holds. Clearly, e.g.  $B_1 \subseteq \mathfrak{C}_{\mathfrak{B}}^{\sharp}(B_1)$ , i.e.  $B_1$  is  $\mathfrak{B}$ -approximatable. Hence, we get

$$B_1 \subseteq \mathfrak{C}_{\mathfrak{B}}^{\flat}(B_1), \text{ but } \partial\mathfrak{C}_{\mathfrak{B}}^{\sharp}(B_1) = \mathfrak{C}_{\mathfrak{B}}^{\sharp}(B_1) \supseteq B_1 \cup B_2 \not\subseteq B_1,$$

a contradiction.

( $\Leftarrow$ ) Let  $X, Y \in 2^U$  such that  $X$  is  $\mathfrak{B}$ -approximatable and  $X \subseteq \mathfrak{C}_{\mathfrak{B}}^b(Y)$ . Then, by the monotonicity of  $\mathfrak{C}_{\mathfrak{B}}^\sharp$  and Proposition 6.3,

$$\partial \mathfrak{C}_{\mathfrak{B}}^\sharp(X) = \mathfrak{C}_{\mathfrak{B}}^\sharp(X) \subseteq \mathfrak{C}_{\mathfrak{B}}^\sharp(\mathfrak{C}_{\mathfrak{B}}^b(Y)) \subseteq Y.$$

□

**6.19 Theorem.** (Csajbók [12], Theorem 4.22) *Let  $\langle 2^U, \mathfrak{D}_{\mathfrak{B}}, \mathfrak{C}_{\mathfrak{B}}^b, \mathfrak{C}_{\mathfrak{B}}^\sharp \rangle$  be a  $\mathfrak{C}_{\mathfrak{B}}^b$ -semi-strong  $\mathfrak{B}$ -approximation space.*

*The pair of maps  $(2^U, \partial \mathfrak{C}_{\mathfrak{B}}^\sharp, \mathfrak{C}_{\mathfrak{B}}^b, 2^U)$  forms a partial Galois connection on  $(2^U, \subseteq)$  if and only if the  $\mathfrak{B}$ -sets are pairwise disjoint.*

**Proof.** Clearly,  $\partial \mathfrak{C}_{\mathfrak{B}}^\sharp$  is a monotone partial map, and  $\mathfrak{C}_{\mathfrak{B}}^b$  is a monotone total map. Thus the conditions (1) and (2) in Definition 2.4 are trivially satisfied. Proposition 6.14 implies condition (3) in Definition 2.4, Propositions 6.15 and 6.18 implies condition (4) in Definition 2.4. □



## Chapter 7

# Applications

To demonstrate the effectiveness of our approach let us see its three real life applications.

The first application will demonstrate the relationship of our approach with natural computing [37] via a biological example.

Natural computing is the field of research that investigates models and computational techniques inspired by nature and, dually, *attempts to understand the world around us in terms of information processing.* ([37], p. 72, *The italics are mine.*)

In particular, we will show how our approach helps us *to understand* some behavioral features of the natural vegetation heritage of Hungary. This presentation is based on the so-called MÉTA program which is a recognition and evaluation system of the state of the natural and semi-natural vegetation heritage of Hungary [1, 47].

The second application presents a general tool-based approximation framework. We observe a class of objects and, as usual, we suppose that there are some well-defined features which an object possesses or not. In practice, two relevant groups of objects can be separated. A group whose elements really possess some features in question and another group whose elements do not substantially possess the same features.

In general, the features of objects cannot directly be observed. We need tools to be at our disposal with which we are able to judge easily and unambiguously whether an object possesses a feature in question or not. However, as a rule, a property ascertained by a tool never coincides with the feature observed by the tool completely.

In the framework, the class of objects is modelled as an abstract set called the universe of discourse. The two separate groups of objects correspond two crisp subsets of the universe. They are disjoint and, in general, their union does not add up to the whole universe. For obvious reasons, the former can be marked with the adjective *positive*, whereas the latter with *negative*.

The objects classified by a tool can also be modelled as one or more crisp subsets of the universe. These subsets are simply called tools. Notice that the complement of a tool is not necessarily a tool at the same time. We also distinguish two types of tools: the *positive* and *negative* ones. Positive (resp., negative) tools provide the opportunity to locate the positive (resp., negative) subset. It is a natural assumption that the union of positive tools and the union of negative tools are disjoint and their union does not add up to the whole universe.

In the proposed tool-based approximation framework, two approximation spaces are defined, a positive and a negative one. Any proportion of the observed objects can *simultaneously* be approximated in the two approximation spaces.

The third example applies the tool-based approximation framework to model the Intrusion Detection Systems (IDS) in computer security. In this framework anomalies and misuses can be detected simultaneously.

## 7.1 Natural Computing—A Biological Example

### 7.1.1 A Brief Outline of the MÉTA program

The biological example is relying on the MÉTA program which is a grid-based, landscape-ecology-oriented, satellite-image supported, field vegetation mapping method of Hungarian habitats (MÉTA stands for Magyarországi Élőhelyek Térképi Adatbázisa: GIS Database of the Hungarian Habitats) [1, 9, 30, 47, 72]. Its main goals include a nationwide survey of the actual state of (semi-)natural vegetation heritage of Hungary and the evaluation of the present state of Hungarian landscapes from a vegetation point of view.

The survey in MÉTA program was carried out on three spatial levels which are nested units of the survey: 1. quadrant, 2. hexagon, 3. habitat type inside the hexagon.

The basic units of the survey are the hexagons. A *hexagon grid* consists of cells of 35 hectares covering the territory of Hungary comprehensively. 267,813 hexagons cover the whole country.

For organizational reasons around 100 hexagons form a *quadrant*. Quadrants are also used for collecting certain vegetation data. The quadrants are the quarters of the base units of the European Flora Survey. Their territory is approximately 35 km<sup>2</sup> and there are 2834 quadrants in Hungary.

In 1996 a new habitat classification system was developed in Hungary, called Á-NÉR (the Hungarian abbreviation stands for General National Habitat Classification System). This system has 112 *habitat types*, all with detailed and standardized descriptions [26]. For the MÉTA method the Á-NÉR system was partly extended and thoroughly revised [8, 48]. These Á-NÉR habitat types are recorded as a list for each hexagon.

The data is mainly collected by a single field survey of the hexagons. The mapper estimates the actual status on the spot. Hexagons with more than 25% natural or semi-natural vegetation are compulsory to survey and to be thoroughly documented. In most cases satellite images and maps help to decide whether a hexagon is compulsory or not. During the field mapping each compulsory hexagon has to be examined by thematically travelling through the area that it covers. Its most dominant habitat type is recorded, as well as those types covering at least 25% of the hexagon. Moreover, the vegetation patches found “on the way” should also be recorded. Vegetation data of noncompulsory hexagons should be documented if these hexagons are crossed by the mapping route or the data can be derived from the satellite image. Collected data are stored in an MS-SQL 2000 database and are mainly recorded as codes.

The data for each habitat type collected by the MÉTA method at the hexagon level are as follows [47]:

- The *areal cover* of each recorded habitat type has to be given as a proportion of the hexagon using the categories < 1, 1, 10, 50, 100%. Satellite images help the observers make the estimation.
- *Spatial pattern* of each type should be documented so that it forms only 1-2, 3 or several distinct patches, or it has a diffuse spatial pattern in the hexagon.
- In order to establish the *naturalness-based habitat quality* of each vegetation type in the hexagon, the following standardized naturalness-based habitat evaluation was used: (1) totally degraded state; (2) heavily degraded state; (3) moderately degraded state; (4) semi-natural state; (5) natural state.

- In each hexagon for each occurring habitat the most characteristic ones from the 28 *threat types* ( $Th_1-Th_{28}$ ) had to be selected that actually threaten the survival and maintenance of the habitat type in the MÉTA hexagon in the next 10-15 years [49]. The strength of the threats is not recorded. The presence of the discernible threats in each case has been documented. Maximum four threats could be given, others were to be written in notes column.

The threatening factors are as follows [47]: improper water management, improper pasturing or mowing, drainage, encroachment of shrubs and trees, burning, afforestation with improper species, woodland patches managed homogeneously, improper selection of trees for timber extraction, logging trees at low age, inappropriate plantation, keeping high densities of game, colonization by invasive plant species, tillage, building and construction, gardening, mining, establishment of a pond, trampling, pollution, rubbish, commercial collection of plants.

- Prediction of future changes of vegetation patches can be supported by the evaluation of the direct effect of the *neighborhood* ( $< 200$  m) on the mapped stands. This evaluation defines whether the neighboring patches will aid or hinder the survival of the particular patch in the next few (10-15) years [49]. The categories are: (1) definitely positive (sustaining neighborhood), (2) slightly positive, (3) indifferent, (4) slightly negative, (5) definitely negative (destructive neighborhood).

The neighborhood is negative, e.g., if there is an intensively used arable field (chemicals, infiltration of fertilizer), expanding settlement, or spreading populations of invasive species surrounding the patch. Neighborhood is positive, if it serves as a source of species, provides proper micro-climate, buffers against degrading factors.

- The *connectedness* is the potential of dispersal of the species of one vegetation stand compared to the surrounding areas. It is documented at two spatial scales: within the distance of several hundred meters (hexagon), and several kilometers (quadrant). It is recorded whether the patches are (1) isolated (typical species of the habitat are not present in the surroundings), (2) connected (species are abundant) or (3) the connectedness is intermediate.

*Connectedness* is also documented at the quadrant level. Categories indicate whether stands are properly connected, moderately connected

or isolated. The first two categories denote any possibilities for dispersal through quadrant whereas the third category shows whether the dispersal is hindered.

Additional data for each habitat type at quadrant level (invasive species, connectedness, regeneration potential), and data for the landscape at hexagon level (potential natural vegetation, area of invasive species and old fields, land-use type, landscape health status) are collected, as well (see [47] for details).

### 7.1.2 Model of Behavioral Features of Natural Vegetation

In this section we present a model for the behavioral features of natural vegetation of Hungary. The model is relying on the results of MÉTA program and the set approximations including both Pawlak's rough set theory and partial approximation of sets.

Let  $\mathfrak{H}$  denote the set of all hexagons of Hungarian landscapes. The hexagons are disjoint and cover the whole country, i.e. they form a partition. Let  $\pi$  denote the equivalence relation corresponding to this partition. If  $A$  denotes an arbitrary area of the country, one can approximate  $A$  in Pawlak's framework. So, Pawlak's lower  $\underline{\pi}(A)$  and upper  $\overline{\pi}(A)$   $\pi$ -approximations can be determined in the universe  $\mathfrak{H}$ .

Now, we want to investigate the area  $A$  in relation to the threat types. First of all, we need the following classification of threat types which is applied by the MÉTA program [49]:  $Th_1$  = water shortage,  $Th_2$  = access water,  $Th_3$  = improper water dynamics,  $Th_4$  = overgrazing,  $Th_5$  = undergrazing,  $Th_6$  = improper grazing regime,  $Th_7$  = abandonment from grazing,  $Th_8$  = improper mowing,  $Th_9$  = abandonment from mowing,  $Th_{10}$  = melioration,  $Th_{11}$  = encroachment of shrubs and trees,  $Th_{12}$  = non-natural burning,  $Th_{13}$  = afforestation with improper species,  $Th_{14}$  = woodland patches managed homogeneously,  $Th_{15}$  = improper selection of trees for timber extraction,  $Th_{16}$  = logging trees at low age,  $Th_{17}$  = new plantations on grasslands,  $Th_{18}$  = overpopulated game,  $Th_{19}$  = colonization by invasive plant species,  $Th_{20}$  = tillage,  $Th_{21}$  = building and construction,  $Th_{22}$  = spread of gardens threatens vegetation,  $Th_{23}$  = mines destroying vegetation,  $Th_{24}$  = establishment of a pond destroying vegetation,  $Th_{25}$  = trampling,  $Th_{26}$  = pollution,  $Th_{27}$  = rubbish,  $Th_{28}$  = commercial collection of plants.

According to the MÉTA method, any threat type determine a well-defined subset of hexagons in  $\mathfrak{H}$ . Let  $\mathfrak{h}_{Th_1}, \mathfrak{h}_{Th_2}, \dots, \mathfrak{h}_{Th_{28}} \in 2^{\mathfrak{H}}$  denote the sets of all hexagons in which threat types  $Th_1, Th_2, \dots, Th_{28}$  are found,

respectively. For instance,  $\mathfrak{h}_{Th_1}$  contains *all hexagons* which are threatened with the threat type  $Th_1$ . Of course,  $\mathfrak{h}_{Th_1}, \mathfrak{h}_{Th_2}, \dots, \mathfrak{h}_{Th_{28}}$  do not necessarily disjoint and their union do not necessarily cover the whole country.

Let  $\mathfrak{B} = \{\mathfrak{h}_{Th_1}, \mathfrak{h}_{Th_2}, \dots, \mathfrak{h}_{Th_{28}}\} \subseteq 2^{\mathfrak{H}}$  be the base system in the universe  $2^{\mathfrak{H}}$ . The base system  $\mathfrak{B}$  can directly be applied only in the case when  $A$  is exactly built up by hexagons, which is rather an extreme case. It is self-explanatory that for the first time we apply Pawlak's  $\pi$ -approximations to  $A$  in the universe  $\mathfrak{H}$ . All the lower and upper  $\pi$ -approximations and  $\pi$ -boundary of  $A$  are already made up of hexagons. Thus we can next apply the partial approximation of sets to the three sets in the universe  $2^{\mathfrak{H}}$ .

Now, let us consider all possible cases one by one.

Case (1) The lower  $\mathfrak{B}$ -approximation of the lower  $\pi$ -approximation of  $A$  is:

$$\mathfrak{C}_{\mathfrak{B}}^b(\underline{\pi}(A)) = \bigcup \{\mathfrak{h} \mid \mathfrak{h} \in \mathfrak{B}, \mathfrak{h} \subseteq \underline{\pi}(A)\}.$$

If  $\mathfrak{C}_{\mathfrak{B}}^b(\underline{\pi}(A)) \neq \emptyset$ , then  $\{\mathfrak{h} \mid \mathfrak{h} \in \mathfrak{B}, \mathfrak{h} \subseteq \underline{\pi}(A)\}$  contains the threat types (to be more exact, the hexagons threatened with these threat types) which *certainly* and *exclusively* appear in  $A$ .

Case (2) The upper  $\mathfrak{B}$ -approximation of the lower  $\pi$ -approximation of  $A$  is:

$$\mathfrak{C}_{\mathfrak{B}}^{\#}(\underline{\pi}(A)) = \bigcup \{\mathfrak{h} \mid \mathfrak{h} \in \mathfrak{B}, \mathfrak{h} \cap \underline{\pi}(A) \neq \emptyset\}.$$

If  $\mathfrak{C}_{\mathfrak{B}}^{\#}(\underline{\pi}(A)) \neq \emptyset$ , then  $\{\mathfrak{h} \mid \mathfrak{h} \in \mathfrak{B}, \mathfrak{h} \cap \underline{\pi}(A) \neq \emptyset\}$  contains the threat types which *certainly* but *not exclusively* appear in  $A$ .

Case (3) The lower  $\mathfrak{B}$ -approximation of the upper  $\pi$ -approximation of  $A$  is:

$$\mathfrak{C}_{\mathfrak{B}}^b(\overline{\pi}(A)) = \bigcup \{\mathfrak{h} \mid \mathfrak{h} \in \mathfrak{B}, \mathfrak{h} \subseteq \overline{\pi}(A)\}.$$

If  $\mathfrak{C}_{\mathfrak{B}}^b(\overline{\pi}(A)) \neq \emptyset$ , then  $\{\mathfrak{h} \mid \mathfrak{h} \in \mathfrak{B}, \mathfrak{h} \subseteq \overline{\pi}(A)\}$  contains the threat types which *perhaps exclusively* but *not certainly* appear in  $A$ .

Case (4) The upper  $\mathfrak{B}$ -approximation of the upper  $\pi$ -approximation of  $A$  is:

$$\mathfrak{C}_{\mathfrak{B}}^{\#}(\overline{\pi}(A)) = \bigcup \{\mathfrak{h} \mid \mathfrak{h} \in \mathfrak{B}, \mathfrak{h} \cap \overline{\pi}(A) \neq \emptyset\}.$$

If  $\mathfrak{C}_{\mathfrak{B}}^{\#}(\overline{\pi}(A)) \neq \emptyset$ , then  $\{\mathfrak{h} \mid \mathfrak{h} \in \mathfrak{B}, \mathfrak{h} \cap \overline{\pi}(A) \neq \emptyset\}$  contains all the threat types which may appear in  $A$  at all.

The  $\mathfrak{B}$ -approximations of the  $\pi$ -boundary of  $A$  provide information about the *spread of the thread types* around  $A$ .

Case (5) The lower  $\mathfrak{B}$ -approximation of the  $\pi$ -boundary of  $A$  is:

$$\mathfrak{C}_{\mathfrak{B}}^b(\pi(A) \setminus \underline{\pi}(A)) = \bigcup \{\mathfrak{h} \mid \mathfrak{h} \in \mathfrak{B}, \mathfrak{h} \subseteq \pi(A) \setminus \underline{\pi}(A)\}.$$

If  $\mathfrak{C}_{\mathfrak{B}}^b(\pi(A) \setminus \underline{\pi}(A)) \neq \emptyset$ , then  $\{\mathfrak{h} \mid \mathfrak{h} \in \mathfrak{B}, \mathfrak{h} \subseteq \pi(A) \setminus \underline{\pi}(A)\}$  contains the threat types which partly belong to  $A$  and partly not. These are the threat types which *spread from inside of  $A$  to outwards*.

Case (6) The upper  $\mathfrak{B}$ -approximation of the  $\pi$ -boundary of  $A$  is:

$$\mathfrak{C}_{\mathfrak{B}}^\sharp(\pi(A) \setminus \underline{\pi}(A)) = \bigcup \{\mathfrak{h} \mid \mathfrak{h} \in \mathfrak{B}, \mathfrak{h} \cap (\pi(A) \setminus \underline{\pi}(A)) \neq \emptyset\}.$$

If  $\mathfrak{C}_{\mathfrak{B}}^\sharp(\pi(A) \setminus \underline{\pi}(A)) \neq \emptyset$ , then  $\{\mathfrak{h} \mid \mathfrak{h} \in \mathfrak{B}, \mathfrak{h} \cap (\pi(A) \setminus \underline{\pi}(A)) \neq \emptyset\}$  contains the threat types which may partly belong to  $A$ . These are the threat types which *spread from outside of  $A$  to inwards* (“Hannibal ante portas”).

## 7.2 A General Tool-Based Approximation Framework

Let  $U$  be a nonempty set.

Let  $A^+, A^- \in 2^U$  be nonempty subsets of  $U$  such that  $A^+ \cap A^- = \emptyset$ .  $A^+$  and  $A^-$  are called the *positive reference set* and *negative reference set*, respectively. In general,  $A^+ \cap A^- = \emptyset$  is the only requirement for  $A^+$  and  $A^-$ . Of course, additional relations between them may be supposed.

Furthermore, let  $\mathfrak{T}^+, \mathfrak{T}^- \subseteq 2^U$  be two nonempty families of subsets of  $U$  such that  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- = \emptyset$ .  $\mathfrak{T}^+$  is called *positive* or  $\mathfrak{T}^+$ -tools,  $\mathfrak{T}^-$  is called *negative* or  $\mathfrak{T}^-$ -tools. For each subset  $T^+ \in \mathfrak{T}^+$  (resp.,  $T^- \in \mathfrak{T}^-$ ) it is *easy* to decide whether an element of  $U$  belongs to  $T^+$  (resp.,  $T^-$ ) or not.

The sets in  $\mathfrak{T}^+$  are not necessarily pairwise disjoint, so they are not in  $\mathfrak{T}^-$ . Neither  $\bigcup \mathfrak{T}^+$  nor  $\bigcup \mathfrak{T}^-$  covers  $U$ .

Note that, the adjectives *positive* and *negative* claim nothing else but that the sets  $A^+$  (resp.,  $\mathfrak{T}^+$ ) and  $A^-$  (resp.,  $\mathfrak{T}^-$ ) are well separated.

According to the general set theoretic approximation framework,  $\mathfrak{T}^+$  and  $\mathfrak{T}^-$  are *primary* tools and let  $\mathfrak{D}_{\mathfrak{T}^+}$  and  $\mathfrak{D}_{\mathfrak{T}^-}$  denote their *derived* tools as usual. Then the quadruples  $\langle U, \mathfrak{D}_{\mathfrak{T}^+}, \mathfrak{C}_{\mathfrak{T}^+}^b, \mathfrak{C}_{\mathfrak{T}^+}^\sharp \rangle$  and  $\langle U, \mathfrak{D}_{\mathfrak{T}^-}, \mathfrak{C}_{\mathfrak{T}^-}^b, \mathfrak{C}_{\mathfrak{T}^-}^\sharp \rangle$  form a  $\mathfrak{C}_{\mathfrak{T}^+}^b$ -semi-strong  $\mathfrak{T}^+$ -approximation space and a  $\mathfrak{C}_{\mathfrak{T}^-}^b$ -semi-strong  $\mathfrak{T}^-$ -approximation space, respectively.

Borrowing the terminology from the inductive logic programming [41], the mutual relationships between  $A^+$  and  $A^-$  can be characterized by available  $\mathfrak{T}^+$ -tools and  $\mathfrak{T}^-$ -tools as follows.

It is said that

- $A^+$  is  $\mathfrak{T}^+$ -complete if  $A^+ \subseteq \mathfrak{C}_{\mathfrak{T}^+}^\sharp(A^+)$ , otherwise  $A^+$  is  $\mathfrak{T}^+$ -incomplete;
- $A^+$  and  $\mathfrak{T}^-$  are consistent, or  $A^+$  is  $\mathfrak{T}^-$ -consistent for short, if  $\mathfrak{C}_{\mathfrak{T}^-}^\sharp(A^+) = \emptyset$ , otherwise  $A^+$  and  $\mathfrak{T}^-$  are inconsistent, or  $A^+$  is  $\mathfrak{T}^-$ -inconsistent for short.

It is said that

- $A^-$  is  $\mathfrak{T}^-$ -complete if  $A^- \subseteq \mathfrak{C}_{\mathfrak{T}^-}^\sharp(A^-)$ , otherwise  $A^-$  is  $\mathfrak{T}^-$ -incomplete;
- $A^-$  and  $\mathfrak{T}^+$  are consistent, or  $A^-$  is  $\mathfrak{T}^+$ -consistent for short, if  $\mathfrak{C}_{\mathfrak{T}^+}^\sharp(A^-) = \emptyset$ , otherwise  $A^-$  and  $\mathfrak{T}^+$  are inconsistent, or  $A^-$  is  $\mathfrak{T}^+$ -inconsistent for short.

From a pure combinatorial point of view, according to the previous terminology, a positive reference set  $A^+$  may be

- $\mathfrak{T}^+$ -complete and  $\mathfrak{T}^-$ -consistent,
- $\mathfrak{T}^+$ -complete and  $\mathfrak{T}^-$ -inconsistent,
- $\mathfrak{T}^+$ -incomplete and  $\mathfrak{T}^-$ -consistent,
- $\mathfrak{T}^+$ -incomplete and  $\mathfrak{T}^-$ -inconsistent;

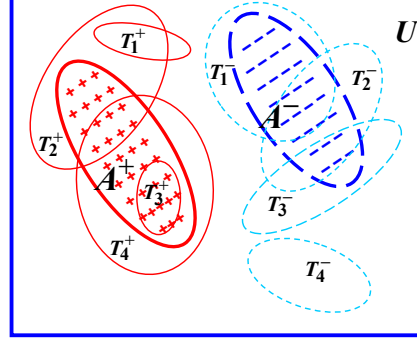
a negative reference set  $A^-$  may be

- $\mathfrak{T}^-$ -complete and  $\mathfrak{T}^+$ -consistent,
- $\mathfrak{T}^-$ -complete and  $\mathfrak{T}^+$ -inconsistent,
- $\mathfrak{T}^-$ -incomplete and  $\mathfrak{T}^+$ -consistent,
- $\mathfrak{T}^-$ -incomplete and  $\mathfrak{T}^+$ -inconsistent.

There may be in sum  $4 \cdot 4 = 16$  different compound situations. However, some of them are impossible by the constraint  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- = \emptyset$ . Now, let us consider all the possible and impossible cases one by one.



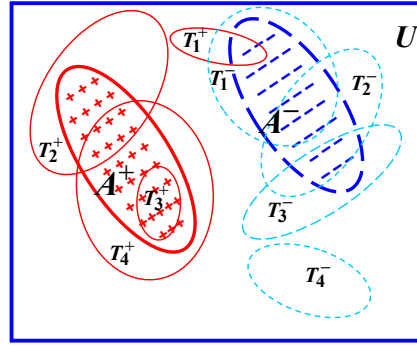
Case (1)  $A^+$  is  $\mathfrak{T}^+$ -complete,  $\mathfrak{T}^-$ -consistent;  
 $A^-$  is  $\mathfrak{T}^-$ -complete,  $\mathfrak{T}^+$ -consistent.  
 It is a possible case (see, e.g., Figure 7.1).



$A^+$  is  $\mathfrak{T}^+$ -complete:  
 $A^+ \subseteq \mathfrak{C}_{\mathfrak{T}^+}^\#(A^+) = T_2^+ \cup T_3^+ \cup T_4^+$ ,  
 $A^+$  is  $\mathfrak{T}^-$ -consistent:  
 $\mathfrak{C}_{\mathfrak{T}^-}^\#(A^+) = \emptyset$ ;  
 $A^-$  is  $\mathfrak{T}^-$ -complete:  
 $A^- \subseteq \mathfrak{C}_{\mathfrak{T}^-}^\#(A^-) = T_1^- \cup T_2^- \cup T_3^-$ ,  
 $A^-$  is  $\mathfrak{T}^+$ -consistent:  
 $\mathfrak{C}_{\mathfrak{T}^+}^\#(A^-) = \emptyset$ .

Figure 7.1: Case (1) It is a possible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- = \emptyset$ .

Case (2)  $A^+$  is  $\mathfrak{T}^+$ -complete,  $\mathfrak{T}^-$ -consistent;  
 $A^-$  is  $\mathfrak{T}^-$ -complete,  $\mathfrak{T}^+$ -inconsistent.  
 It is an impossible case (see, e.g., Figure 7.2).



$A^+$  is  $\mathfrak{T}^+$ -complete:  
 $A^+ \subseteq \mathfrak{C}_{\mathfrak{T}^+}^\#(A^+) = T_2^+ \cup T_3^+ \cup T_4^+$ ,  
 $A^+$  is  $\mathfrak{T}^-$ -consistent:  
 $\mathfrak{C}_{\mathfrak{T}^-}^\#(A^+) = \emptyset$ ;  
 $A^-$  is  $\mathfrak{T}^-$ -complete:  
 $A^- \subseteq \mathfrak{C}_{\mathfrak{T}^-}^\#(A^-) = T_1^- \cup T_2^- \cup T_3^-$ ,  
 $A^-$  is  $\mathfrak{T}^+$ -inconsistent:  
 $\mathfrak{C}_{\mathfrak{T}^+}^\#(A^-) = T_1^+ \neq \emptyset$ .

Figure 7.2: Case (2) It is an impossible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- \neq \emptyset$ .

Case (3)  $A^+$  is  $\mathfrak{T}^+$ -complete,  $\mathfrak{T}^-$ -consistent;  
 $A^-$  is  $\mathfrak{T}^-$ -incomplete,  $\mathfrak{T}^+$ -consistent.  
 It is a possible case (see, e.g., Figure 7.3).

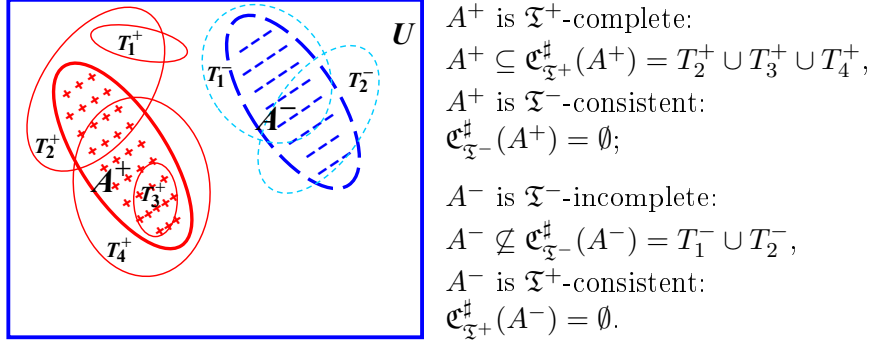


Figure 7.3: Case (3) It is a possible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- = \emptyset$ .

Case (4)  $A^+$  is  $\mathfrak{T}^+$ -complete,  $\mathfrak{T}^-$ -consistent;  
 $A^-$  is  $\mathfrak{T}^-$ -incomplete,  $\mathfrak{T}^+$ -inconsistent.  
 It is a possible case (see, e.g., Figure 7.4).

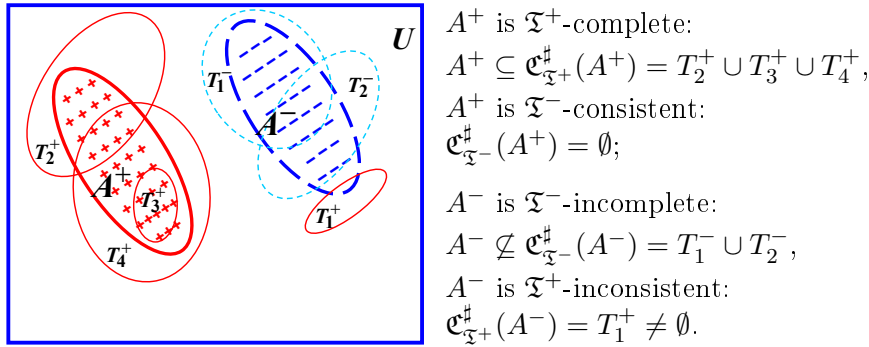
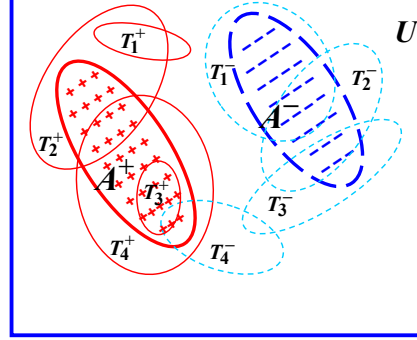


Figure 7.4: Case (4) It is a possible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- = \emptyset$ .

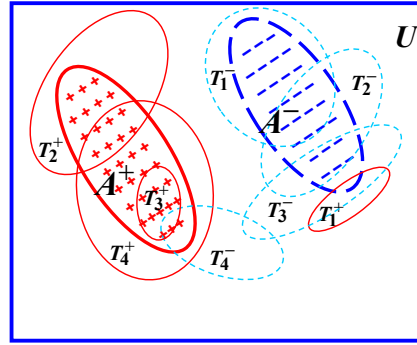
Case (5)  $A^+$  is  $\mathfrak{T}^+$ -complete,  $\mathfrak{T}^-$ -inconsistent;  
 $A^-$  is  $\mathfrak{T}^-$ -complete,  $\mathfrak{T}^+$ -consistent.  
 It is an impossible case (see, e.g., Figure 7.5).



$A^+$  is  $\mathfrak{T}^+$ -complete:  
 $A^+ \subseteq \mathfrak{C}_{\mathfrak{T}^+}^\#(A^+) = T_2^+ \cup T_3^+ \cup T_4^+$ ,  
 $A^+$  is  $\mathfrak{T}^-$ -inconsistent:  
 $\mathfrak{C}_{\mathfrak{T}^-}^\#(A^+) = T_4^- \neq \emptyset$ ;  
 $A^-$  is  $\mathfrak{T}^-$ -complete:  
 $A^- \subseteq \mathfrak{C}_{\mathfrak{T}^-}^\#(A^-) = T_1^- \cup T_2^- \cup T_3^-$ ,  
 $A^-$  is  $\mathfrak{T}^+$ -consistent:  
 $\mathfrak{C}_{\mathfrak{T}^+}^\#(A^-) = \emptyset$ .

Figure 7.5: Case (5) It is an impossible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- \neq \emptyset$ .

Case (6)  $A^+$  is  $\mathfrak{T}^+$ -complete,  $\mathfrak{T}^-$ -inconsistent;  
 $A^-$  is  $\mathfrak{T}^-$ -complete,  $\mathfrak{T}^+$ -inconsistent.  
 It is an impossible case (see, e.g., Figure 7.6).



$A^+$  is  $\mathfrak{T}^+$ -complete:  
 $A^+ \subseteq \mathfrak{C}_{\mathfrak{T}^+}^\#(A^+) = T_2^+ \cup T_3^+ \cup T_4^+$ ,  
 $A^+$  is  $\mathfrak{T}^-$ -inconsistent:  
 $\mathfrak{C}_{\mathfrak{T}^-}^\#(A^+) = T_4^- \neq \emptyset$ ;  
 $A^-$  is  $\mathfrak{T}^-$ -complete:  
 $A^- \subseteq \mathfrak{C}_{\mathfrak{T}^-}^\#(A^-) = T_1^- \cup T_2^- \cup T_3^-$ ,  
 $A^-$  is  $\mathfrak{T}^+$ -inconsistent:  
 $\mathfrak{C}_{\mathfrak{T}^+}^\#(A^-) = T_1^+ \neq \emptyset$ .

Figure 7.6: Case (6) It is an impossible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- \neq \emptyset$ .

Case (7)  $A^+$  is  $\mathfrak{T}^+$ -complete,  $\mathfrak{T}^-$ -inconsistent;  
 $A^-$  is  $\mathfrak{T}^-$ -incomplete,  $\mathfrak{T}^+$ -consistent.  
 It is an impossible case (see, e.g., Figure 7.7).

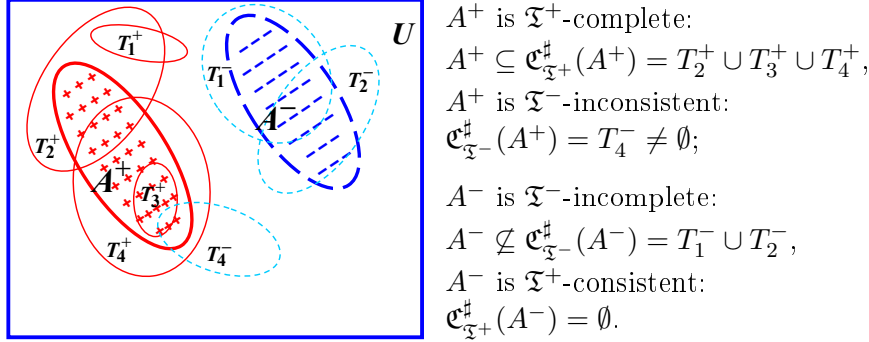


Figure 7.7: Case (7) It is an impossible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- \neq \emptyset$ .

Case (8)  $A^+$  is  $\mathfrak{T}^+$ -complete,  $\mathfrak{T}^-$ -inconsistent;  
 $A^-$  is  $\mathfrak{T}^-$ -incomplete,  $\mathfrak{T}^+$ -inconsistent.  
 It is an impossible case (see, e.g., Figure 7.8).

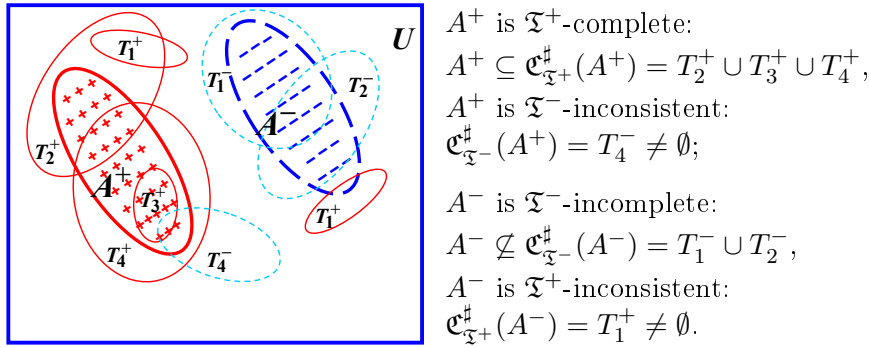


Figure 7.8: Case (8) It is an impossible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- \neq \emptyset$ .

Case (9)  $A^+$  is  $\mathfrak{T}^+$ -incomplete,  $\mathfrak{T}^-$ -consistent;  
 $A^-$  is  $\mathfrak{T}^-$ -complete,  $\mathfrak{T}^+$ -consistent.  
 It is a possible case (see, e.g., Figure 7.9).

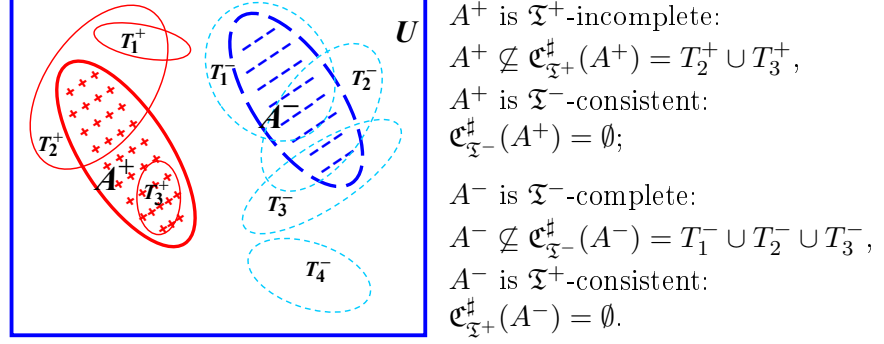


Figure 7.9: Case (9) It is a possible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- = \emptyset$ .

Case (10)  $A^+$  is  $\mathfrak{T}^+$ -incomplete,  $\mathfrak{T}^-$ -consistent;  
 $A^-$  is  $\mathfrak{T}^-$ -complete,  $\mathfrak{T}^+$ -inconsistent.  
 It is an impossible case (see, e.g., Figure 7.10).

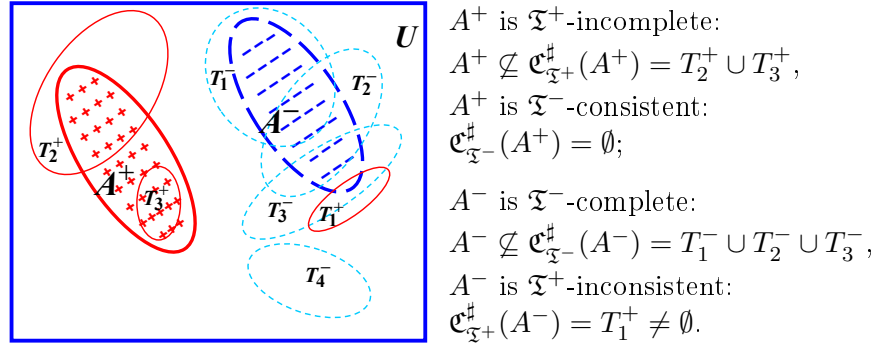
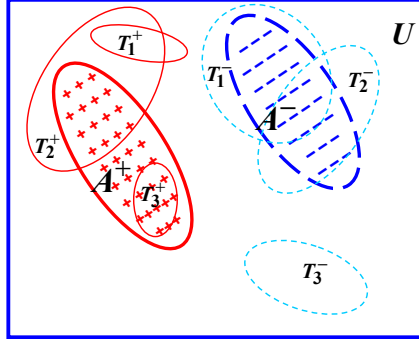


Figure 7.10: Case (10) It is an impossible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- \neq \emptyset$ .

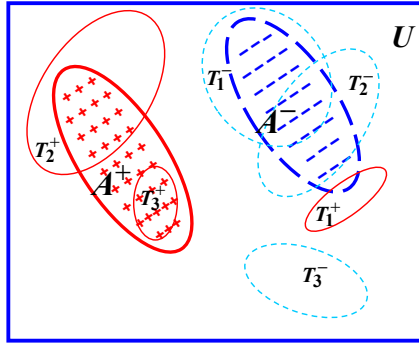
Case (11)  $A^+$  is  $\mathfrak{T}^+$ -incomplete,  $\mathfrak{T}^-$ -consistent;  
 $A^-$  is  $\mathfrak{T}^-$ -incomplete,  $\mathfrak{T}^+$ -consistent.  
 It is a possible case (see, e.g., Figure 7.11).



$A^+$  is  $\mathfrak{T}^+$ -incomplete:  
 $A^+ \not\subseteq \mathfrak{C}_{\mathfrak{T}^+}^\#(A^+) = T_2^+ \cup T_3^+$ ,  
 $A^+$  is  $\mathfrak{T}^-$ -consistent:  
 $\mathfrak{C}_{\mathfrak{T}^-}^\#(A^+) = \emptyset$ ;  
 $A^-$  is  $\mathfrak{T}^-$ -incomplete:  
 $A^- \not\subseteq \mathfrak{C}_{\mathfrak{T}^-}^\#(A^-) = T_1^- \cup T_2^-$ ,  
 $A^-$  is  $\mathfrak{T}^+$ -consistent:  
 $\mathfrak{C}_{\mathfrak{T}^+}^\#(A^-) = \emptyset$ .

Figure 7.11: Case (11) It is a possible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- = \emptyset$ .

Case (12)  $A^+$  is  $\mathfrak{T}^+$ -incomplete,  $\mathfrak{T}^-$ -consistent;  
 $A^-$  is  $\mathfrak{T}^-$ -incomplete,  $\mathfrak{T}^+$ -inconsistent.  
 It is a possible case (see, e.g., Figure 7.12).



$A^+$  is  $\mathfrak{T}^+$ -incomplete:  
 $A^+ \not\subseteq \mathfrak{C}_{\mathfrak{T}^+}^\#(A^+) = T_2^+ \cup T_3^+$ ,  
 $A^+$  is  $\mathfrak{T}^-$ -consistent:  
 $\mathfrak{C}_{\mathfrak{T}^-}^\#(A^+) = \emptyset$ ;  
 $A^-$  is  $\mathfrak{T}^-$ -incomplete:  
 $A^- \not\subseteq \mathfrak{C}_{\mathfrak{T}^-}^\#(A^-) = T_1^- \cup T_2^-$ ,  
 $A^-$  is  $\mathfrak{T}^+$ -inconsistent:  
 $\mathfrak{C}_{\mathfrak{T}^+}^\#(A^-) = T_1^+ \neq \emptyset$ .

Figure 7.12: Case (12) It is a possible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- = \emptyset$ .

Case (13)  $A^+$  is  $\mathfrak{T}^+$ -incomplete,  $\mathfrak{T}^-$ -inconsistent;  
 $A^-$  is  $\mathfrak{T}^-$ -complete,  $\mathfrak{T}^+$ -consistent.  
 It is a possible case (see, e.g., Figure 7.13).

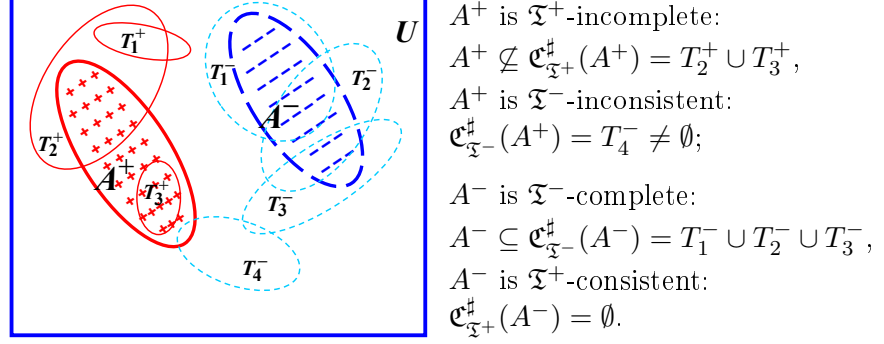


Figure 7.13: Case (13) It is a possible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- = \emptyset$ .

Case (14)  $A^+$  is  $\mathfrak{T}^+$ -incomplete,  $\mathfrak{T}^-$ -inconsistent;  
 $A^-$  is  $\mathfrak{T}^-$ -complete,  $\mathfrak{T}^+$ -inconsistent.  
 It is an impossible case (see, e.g., Figure 7.14).

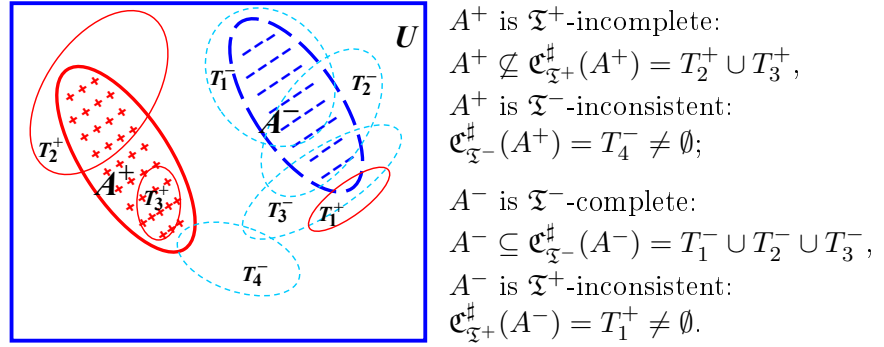
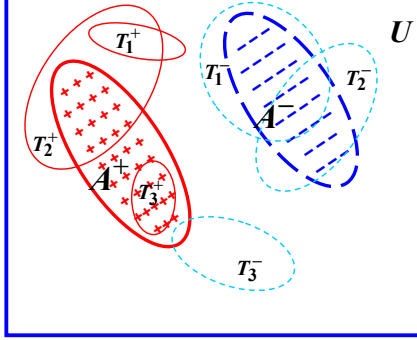


Figure 7.14: Case (14) It is an impossible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- \neq \emptyset$ .

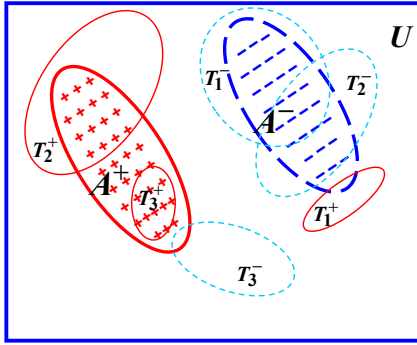
Case (15)  $A^+$  is  $\mathfrak{T}^+$ -incomplete,  $\mathfrak{T}^-$ -inconsistent;  
 $A^-$  is  $\mathfrak{T}^-$ -incomplete,  $\mathfrak{T}^+$ -consistent.  
 It is a possible case (see, e.g., Figure 7.15).



$A^+$  is  $\mathfrak{T}^+$ -incomplete:  
 $A^+ \not\subseteq \mathfrak{C}_{\mathfrak{T}^+}^\#(A^+) = T_2^+ \cup T_3^+$ ,  
 $A^+$  is  $\mathfrak{T}^-$ -inconsistent:  
 $\mathfrak{C}_{\mathfrak{T}^-}^\#(A^+) = T_3^- \neq \emptyset$ ;  
 $A^-$  is  $\mathfrak{T}^-$ -incomplete:  
 $A^- \not\subseteq \mathfrak{C}_{\mathfrak{T}^-}^\#(A^-) = T_1^- \cup T_2^-$ ,  
 $A^-$  is  $\mathfrak{T}^+$ -consistent:  
 $\mathfrak{C}_{\mathfrak{T}^+}^\#(A^-) = \emptyset$ .

Figure 7.15: Case (15) It is a possible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- \neq \emptyset$ .

Case (16)  $A^+$  is  $\mathfrak{T}^+$ -incomplete,  $\mathfrak{T}^-$ -inconsistent;  
 $A^-$  is  $\mathfrak{T}^-$ -incomplete,  $\mathfrak{T}^+$ -inconsistent.  
 It is a possible case (see, e.g., Figure 7.16).



$A^+$  is  $\mathfrak{T}^+$ -incomplete:  
 $A^+ \not\subseteq \mathfrak{C}_{\mathfrak{T}^+}^\#(A^+) = T_2^+ \cup T_3^+$ ,  
 $A^+$  is  $\mathfrak{T}^-$ -inconsistent:  
 $\mathfrak{C}_{\mathfrak{T}^-}^\#(A^+) = T_3^- \neq \emptyset$ ;  
 $A^-$  is  $\mathfrak{T}^-$ -incomplete:  
 $A^- \not\subseteq \mathfrak{C}_{\mathfrak{T}^-}^\#(A^-) = T_1^- \cup T_2^-$ ,  
 $A^-$  is  $\mathfrak{T}^+$ -inconsistent:  
 $\mathfrak{C}_{\mathfrak{T}^+}^\#(A^-) = T_1^+ \neq \emptyset$ .

Figure 7.16: Case (16) It is a possible case because  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- \neq \emptyset$ .



In order to build up the general tool-based approximation framework, at the beginning, let us assume that a positive reference set  $A^+$  and a negative reference set  $A^-$  are at our disposal together with the suitable positive tools  $\mathfrak{T}^+$  and negative tools  $\mathfrak{T}^-$ . Initially, we only presuppose that  $A^+ \cap A^- = \emptyset$  and  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- = \emptyset$ . The framework can be built up and used in the following three consecutive steps.

(1) *Mutual justifying the reference sets and tools*

Step (1) is intended to reveal consistencies/inconsistencies between positive (resp., negative) reference sets and negative (resp., positive) tools and the completeness/incompleteness of reference sets in terms of  $\mathfrak{T}^+$  and  $\mathfrak{T}^-$ -approximation spaces.

(2) *Rebuilding positive and negative tools*

Step (2) is intended to resolve inconsistencies completely and eliminate incompleteness as far as possible.

**The case of consistency.** There is nothing to be done.

**The case of inconsistency.** We have to *decide* within the context of the observed field

- if  $A^+$  or the concerned negative tools is reasonable or not, and/or
- whether  $A^-$  or the concerned positive tools is reasonable or not.

**The case of completeness.** We remove the covered positive and/or negative reference sets from the framework, because they do not cause any problem from the security point of view.

**The case of incompleteness.** We may *decide* within the context of the observed field either to remove the uncovered subset from  $A^+$  (resp.,  $A^-$ ) on the whole or in part or to augment the positive (resp., negative) tools with new subsets whose elements are patterned upon one or more elements of the uncovered subset of  $A^+$  (resp.,  $A^-$ ).

These new subsets may contain any elements of the universe, provided that they can easily be determined.

For the new tools  $\mathfrak{T}^+$  and/or  $\mathfrak{T}^-$ ,  $\bigcup \mathfrak{T}^+ \cap \bigcup \mathfrak{T}^- = \emptyset$  should also be fulfilled.

In this step, all decisions should be made by domain experts on professional criteria within the context of the observed field.

We obtain new rebuilt tools, denoted by  $\mathfrak{T}_r^+$  and  $\mathfrak{T}_r^-$ , by the end of the steps (1) and (2).

(3) *Applying rebuilt tools*

Step (3) is intended to justify any subset of the universe in terms of partial approximation of sets based on rebuilt positive and negative tools as usual.

## 7.3 Simultaneous Anomaly and Misuse Intrusion Detections

### 7.3.1 Introduction

Nowadays, people run their applications in a complex open computing environment including all sorts of interconnected devices. While this environment permanently changes, people watch their applications, work with one of them, and, in general, also follow details of other applications with attention. Many applications, at the same time, work unnoticeably in the background, and some of them, even by stealth. In order to meet the computer security challenge in human environments, Intrusion Detection Systems (IDS) have to be designed.

To a large extent, acceptable and/or unacceptable patterns in the behaviors of the observed system cannot be designed and/or forecast in advance. This strange situation is smartly described by B. Schneier:

You have to imagine an intelligent and malicious adversary inside your system (the ‘Satan’ of Satan’s computer), constantly trying new ways to subvert it. You have to consider all the ways your system can fail, most of them having nothing to do with the design itself. You have to look at everything backwards, upside down, and sideways. You have to think like an alien. ([4], from the Foreword by B. Schneier)

Computer security has definitely different challenges in corporate information systems and non-professional human computing environments. In the former one there are many approaches for security policy specification. Traditionally, security policies are formulated along the so-called CIA taxonomy which sees security as the combination of three attributes—confidentiality, integrity, and availability [66].

People in non-professional human computing environments are flooded by recommendations how they operate their system and use their applications. In headwords only: strong passwords creation tips and maintenance, virus protection, software downloading and installation, removable media risks, encryption and cryptographic means, system backups, incident handling, e-mail and internet use best practises, etc.

Non-professionals, of course, cannot convert these good pieces of advice into security policies, especially into formal ones. Meanwhile, arising from the human thinking, all non-professional users have *anticipated hypotheses* how an application or the whole system should or should not work [51]. These presupposes may range from informal expected behaviors, their constituents might call expected ‘milestones’, to more formal ones described in user manuals and other development artifacts.

To built up a formal security model for computer systems, first, ones have to understand what has to be protected and why. The answers determine the *security strategy* which is, in turn, expressed by *security policies* [10, 29]. Security policies as a general rule *prescribe* and *proscribe* behaviors of software systems *in advance*, only with more or less knowledge about future applications.

We model a computer system as a semantic system model, a so-called *traced-based model*. A traced-based model describes the behaviors of a computer system as sets of execution traces. We focus *solely* on externally observable execution traces sent out by the observed computing system.

An important note. An information system, among others, consists of different software components of finite number. Each component has an individual behavior, and the global behavior of the whole system is the collection of the individual ones. The components can operate with each other. Their interconnections may be deliberate or *ad hoc*. Notice, however, that in both cases, the mechanism of these interconnections mostly remains concealed from the external observers. In particular, based on only external observations we cannot model these synchronization mechanisms.

According to the trace-based model, it is assumed that security policies specify the prescribed and proscribed behaviors of a computer system via the patterns of *acceptable* and *unacceptable* execution traces, respectively [4, 7]. We also take into account the partial nature of security policies. Typically some policies may only apply to specific hardware appliances, software applications or type of information. For instance, possibly it is enough to enforce the information flow policy on such software processes which handle confidential information.

In order to meet the computer security challenge outlined above, a sort of sophisticated *Intrusion Detection System* (IDS) has to be developed. Intrusion detection techniques can be categorized in different ways. For a survey of intrusion detection methods, see, e.g. [6, 74].

Intrusion detection techniques are categorized into *anomaly* and *misuse* detections. Both techniques use patterns based on different types of data [74, 75]. Anomaly detection, originally proposed by Denning [23], profiles *expected* behaviors to identify abnormal behaviors as anomalies which deviate from the defined profile. Misuse detection profiles patterns of known attacks, i.e., *unexpected* behaviors, to identify abnormal behavior directly.

In our IDS model, the patterns of expected and unexpected execution traces provide positive and negative *reference sets*, while the patterns of acceptable and unacceptable execution traces determined by the security policies serve as positive and negative *tools*.

### 7.3.2 The Intrusion Detection Model

Let us assume that  $A$  denotes a nonempty finite set of *symbols*. A *string* is a finite or infinite sequence of symbols chosen from  $A$ . String containing no symbols is called the *empty string* and is denoted by  $\lambda$ . Let  $A^*$  and  $A^\omega$  denote the set of all finite and infinite strings made up of symbols chosen from  $A$ , respectively. We also use the following notations:  $A^+ = A^* \setminus \{\lambda\}$ ,  $A^\infty = A^* \cup A^\omega$ .

An *execution sequence* or *trace* consists of linearly ordered observable atomic actions concerning the observed computer system [5]. Types of atomic actions are the following:

- Let  $A_{req}$  be a finite nonempty set of externally observable required atomic actions. It is called the *required action set*.
- Let  $A_{uns}$  be a finite nonempty set of insecure atomic actions which may happen during the running time of the observed system. It is called the *unsafe action set*.
- Let  $A_{neu}$  be a finite nonempty set of additional atomic actions which by themselves may not influence the safety of the observed system. It is called the *neutral action set*.

Let us assume that  $A_{req}$ ,  $A_{uns}$  and  $A_{neu}$  are pairwise disjoint. Let  $\mathcal{A} = A_{req} \cup A_{uns} \cup A_{neu}$  which is called the *system action set*.

An *execution trace*  $\sigma \in \mathcal{A}^\infty$  is a finite or infinite sequence of not necessarily different system atomic actions.

**7.1 Definition.** By a *computer system* we mean an  $(\mathcal{A}, \Sigma)$  pair, where  $\Sigma(\neq \emptyset) \subseteq \mathcal{A}^\omega$ .

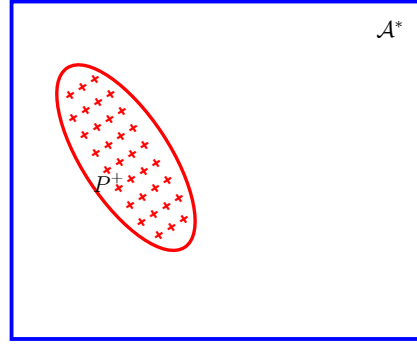
If the computer system terminates, we as usual model it as an infinite execution trace by infinitely stuttering the empty action  $\lambda$ .

$\mathcal{A}^*$  is the set of all possible finite observable execution traces generated by the computer system.

**7.2 Definition.** By a *system observation* we mean an  $(\mathcal{A}, \Sigma^{obs})$  pair, where  $\Sigma^{obs}(\neq \emptyset) \subseteq \mathcal{A}^*$ . A subset  $S \subseteq \Sigma^{obs}$  is called the *snapshot*.

Let  $\mathcal{A}^*$  be the universe of discourse in our IDS model.

Let  $P^+ \subseteq (A_{req} \cup A_{neu})^* \subsetneq \mathcal{A}^*$  denote the set of *expected* execution traces which describes the expected behavior of the running system (see Figure 7.17).



$\mathcal{A}^*$  is the universe of the IDS model: the set of all possible finite observable execution traces.

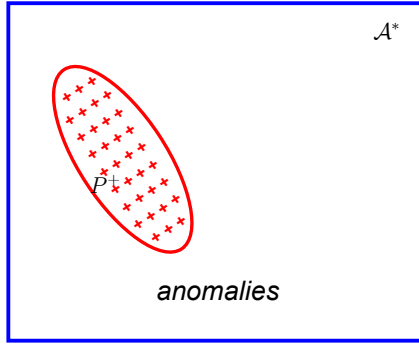
$P^+ \subseteq (A_{req} \cup A_{neu})^* \subsetneq \mathcal{A}^*$  profiles the expected behavior of the computer system.

Figure 7.17: Initialization of the IDS model

$\mathcal{A}^* \setminus P^+$  can be seen as the abnormal behavior of the system which deviates from the previously defined expected profile. Its elements are called *anomalies* [23] (see Figure 7.18).

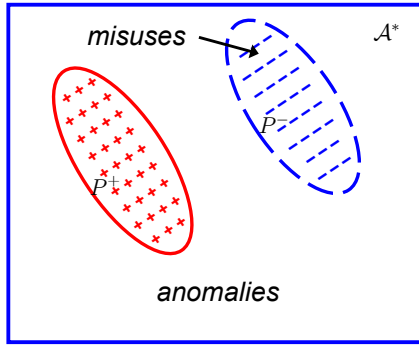
According to our current available knowledge, however, only a subset  $P^- \subseteq \mathcal{A}^* \setminus P^+$  can really be modelled as the *unexpected* behavior of the system. Its elements are usually called *misuses*. Of course, the unexpected behavior  $P^-$  has its own right to be profiled (see Figure 7.19).

Expected and unexpected behaviors serve as positive and negative reference sets in our IDS model.



$\mathcal{A}^* \setminus P^+$  can be seen as the abnormal behavior of the system which deviates from the previously defined expected profile  $P^+$ . Its elements are called *anomalies*.

Figure 7.18: Anomalies



$P^- \subseteq \mathcal{A}^* \setminus P^+$  is the unexpected behavior of the system. Its elements are called *misuses*.  $P^+$  and  $P^-$  are considered as positive and negative reference sets in our IDS model.

Figure 7.19: Misuses

Security strategy is modelled as a family of sets  $\mathfrak{S} \subseteq 2^{\mathcal{A}^*}$ .

Security strategy is expressed by security policies of finite number. Prescriptions and proscriptions of security policies can also be represented by families of sets of execution traces denoted by  $\mathfrak{S}^+$  and  $\mathfrak{S}^-$  respectively. Let

$$\begin{aligned}\mathfrak{S}^+ &= \{S_i^+ \mid S_i^+ \subseteq \mathcal{A}^*, i = 1, \dots, n^+\} \subseteq 2^{\mathcal{A}^*}, \\ \mathfrak{S}^- &= \{S_i^- \mid S_i^- \subseteq \mathcal{A}^*, i = 1, \dots, n^-\} \subseteq 2^{\mathcal{A}^*},\end{aligned}$$

where  $\mathfrak{S} = \mathfrak{S}^+ \cup \mathfrak{S}^-$ . It is assumed that  $\bigcup \mathfrak{S}^+ \cap \bigcup \mathfrak{S}^- = \emptyset$ , i.e. an execution trace cannot model a prescribed and proscribed behavior at the same time. Note that  $\bigcup \mathfrak{S}^+$  (resp.,  $\bigcup \mathfrak{S}^-$ ) may contains execution traces with unsafe (resp., required and neutral) atomic actions.

Members of  $\mathfrak{S}^+$  are called *acceptable* behaviors, whereas members of  $\mathfrak{S}^-$  are called *unacceptable* behaviors. They serve as positive and nega-

tive tools in our IDS model. Acceptable/unacceptable behaviors and expected/unexpected behaviors *mutually justify each other*.

The sets in  $\mathfrak{S}^+$  (resp., in  $\mathfrak{S}^-$ ) are not necessarily pairwise disjoint and, in general,  $\bigcup \mathfrak{S}^+$  (resp.,  $\bigcup \mathfrak{S}^-$ ) does not cover  $\mathcal{A}^*$ . That is,  $\mathfrak{S}^+$  and  $\mathfrak{S}^-$  are base systems over the universe  $\mathcal{A}^*$ . In other words,  $\langle \mathcal{A}^*, \mathfrak{D}_{\mathfrak{S}^+}, \mathfrak{C}_{\mathfrak{S}^+}^b, \mathfrak{C}_{\mathfrak{S}^+}^\# \rangle$  and  $\langle \mathcal{A}^*, \mathfrak{D}_{\mathfrak{S}^-}, \mathfrak{C}_{\mathfrak{S}^-}^b, \mathfrak{C}_{\mathfrak{S}^-}^\# \rangle$  form a  $\mathfrak{S}^+$ -semi-strong and a  $\mathfrak{S}^-$ -semi-strong approximation space respectively (see Figures 7.20 and 7.21).

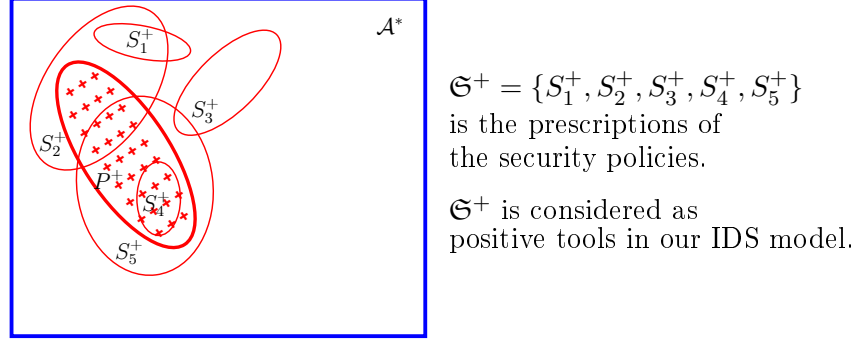


Figure 7.20: Approximation space  $\langle \mathcal{A}^*, \mathfrak{D}_{\mathfrak{S}^+}, \mathfrak{C}_{\mathfrak{S}^+}^b, \mathfrak{C}_{\mathfrak{S}^+}^\# \rangle$  with the positive reference set  $P^+$ .

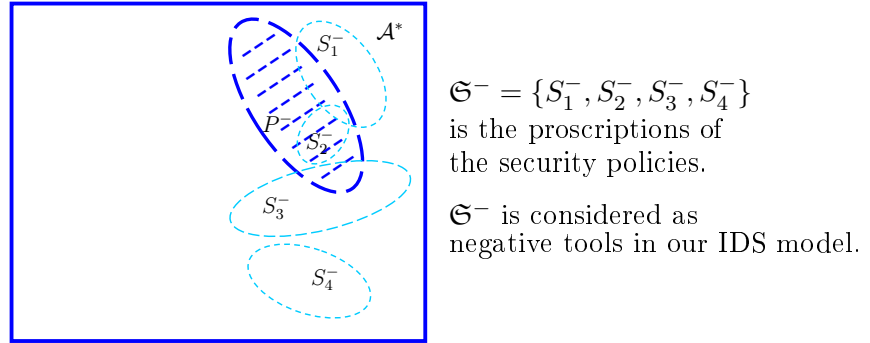
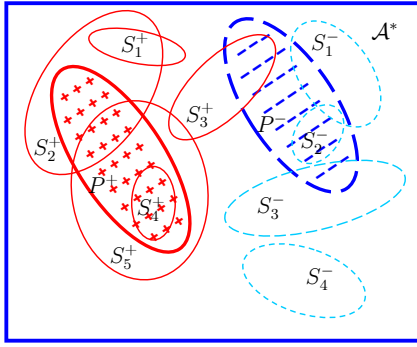


Figure 7.21: Approximation space  $\langle \mathcal{A}^*, \mathfrak{D}_{\mathfrak{S}^-}, \mathfrak{C}_{\mathfrak{S}^-}^b, \mathfrak{C}_{\mathfrak{S}^-}^\# \rangle$  with the negative reference set  $P^-$ .

The complete IDS model is depicted in Figure 7.22.

Using the running example, we show how this model can be applied.

Step 1. *Mutual justifying the reference sets and tools*



Approximation spaces:

$$\langle \mathcal{A}^*, \mathcal{D}_{\mathfrak{S}^+}, \mathfrak{C}_{\mathfrak{S}^+}^b, \mathfrak{C}_{\mathfrak{S}^+}^\# \rangle,$$

$\langle \mathcal{A}^*, \mathcal{D}_{\mathfrak{S}^-}, \mathfrak{C}_{\mathfrak{S}^-}^b, \mathfrak{C}_{\mathfrak{S}^-}^\# \rangle$ , and positive, negative reference sets:  $P^+, P^-$ .

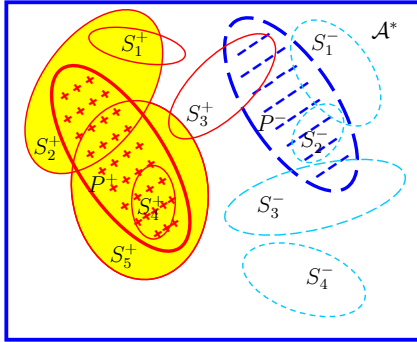
Acceptable, unacceptable traces:

$$\mathfrak{S}^+ = \{S_1^+, \dots, S_5^+\},$$

$\mathfrak{S}^- = \{S_1^-, \dots, S_4^-\}$ , and expected, unexpected traces:  $P^+, P^-$ .

Figure 7.22: The complete IDS model.

(1)  $P^+$  is  $\mathfrak{S}^-$ -consistent and  $\mathfrak{S}^+$ -complete (see Figure 7.23).



$P^+$  is  $\mathfrak{S}^-$ -consistent:  $\mathfrak{C}_{\mathfrak{S}^-}^\#(P^+) = \emptyset$ .

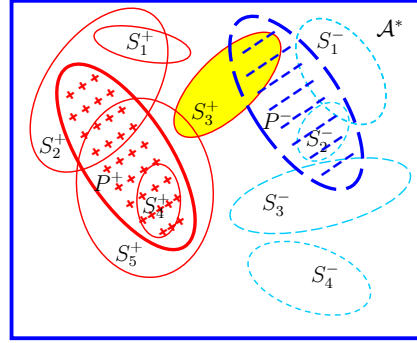
$P^+$  is  $\mathfrak{S}^+$ -complete:

$$P^+ \subseteq \mathfrak{C}_{\mathfrak{S}^+}^\#(P^+) = S_2^+ \cup S_4^+ \cup S_5^+.$$

Figure 7.23:  $P^+$  is  $\mathfrak{S}^-$ -consistent and  $\mathfrak{S}^+$ -complete.

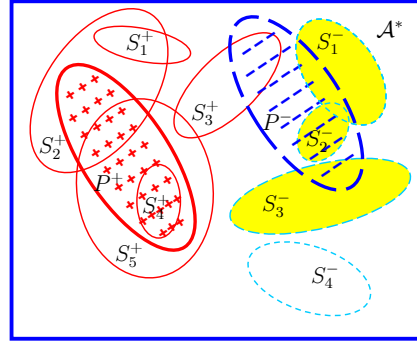


(2)  $P^-$  is  $\mathfrak{S}^+$ -inconsistent (see Figure 7.24) and  $P^-$  is  $\mathfrak{S}^-$ -incomplete (see Figure 7.25).



$P^-$  is  $\mathfrak{S}^+$ -inconsistent:  
 $\mathfrak{C}_{\mathfrak{S}^+}^\#(A^-) = S_3^+ \neq \emptyset$ .

Figure 7.24:  $P^-$  is  $\mathfrak{S}^+$ -inconsistent.



$P^-$  is  $\mathfrak{S}^-$ -incomplete:  
 $P^- \not\subseteq \mathfrak{C}_{\mathfrak{S}^-}^\#(P^-) = S_1^- \cup S_2^- \cup S_3^-$ .

Figure 7.25:  $P^-$  is  $\mathfrak{S}^-$ -incomplete.

Step 2. Rebuilding positive and negative tools

- (1) Since  $P^+$  was  $\mathfrak{S}^-$ -consistent, there was nothing to be done.
- (2) Since  $P^+$  was  $\mathfrak{S}^+$ -complete,  $P^+$  was removed from the framework (see Figure 7.26).

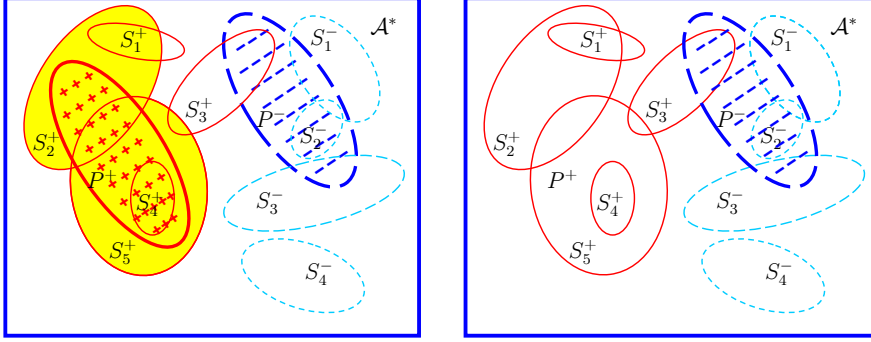


Figure 7.26:  $P^+$  is removed from the framework.

- (3)  $P^-$  was  $\mathfrak{S}^+$ -inconsistent, we *decided* that the positive tool  $S_3^+$  was reasonable (see Figure 7.27).

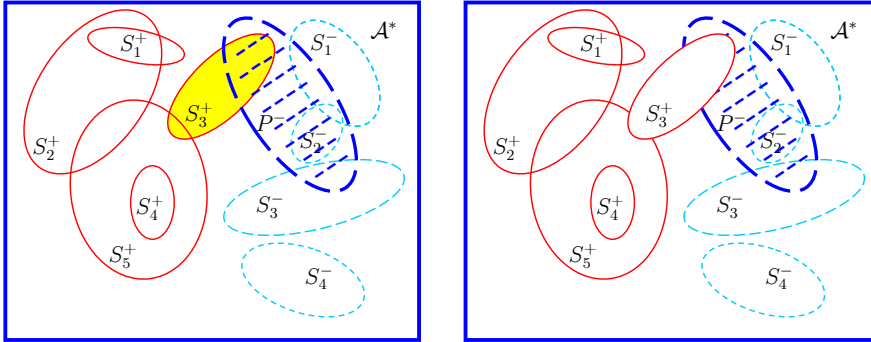


Figure 7.27: We *decide* that the positive tool  $S_3^+$  is reasonable.

(4)  $P^-$  was  $\mathfrak{S}^-$ -incomplete, we *decided* that we augmented negative tools with  $S_5^-$ ,  $S_6^-$  patterned upon one or more elements of the uncovered subset of  $P^-$ . Then  $P^-$  was removed from the framework (see Figure 7.28).

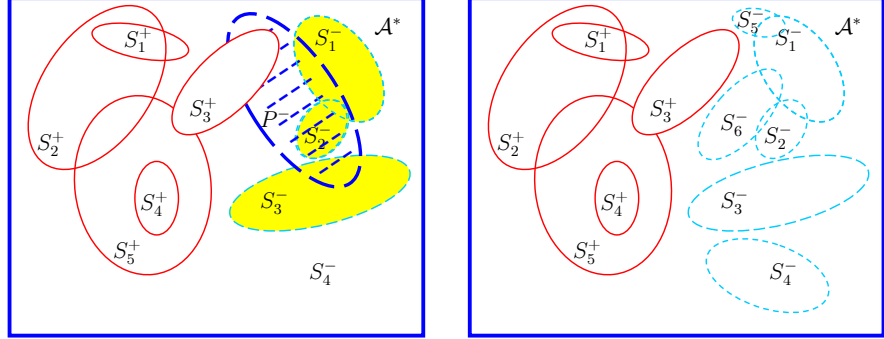
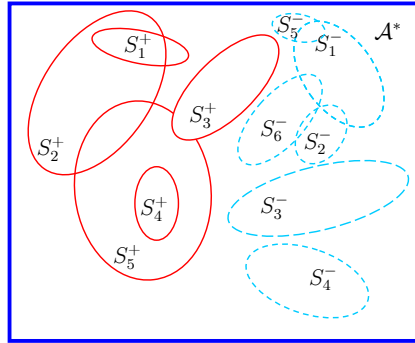


Figure 7.28: New negative tools  $S_5^-$ ,  $S_6^-$ .

By the end of Steps (1) and (2), we obtained the rebuilt positive tools  $\mathfrak{S}_r^+ = \mathfrak{S}^+$ , and the rebuilt negative tools  $\mathfrak{S}_r^- = \mathfrak{S}^- \cup \{S_5^-, S_6^-\}$  (see Figure 7.29).



The rebuilt positive tools:

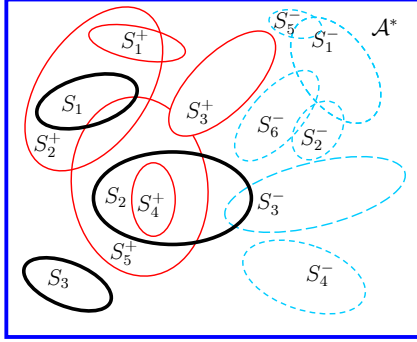
$$\mathfrak{S}_r^+ = \mathfrak{S}^+$$

The rebuilt negative tools:

$$\mathfrak{S}_r^- = \mathfrak{S}^- \cup \{S_5^-, S_6^-\}.$$

Figure 7.29: The rebuilt tools.

Step 3. We apply the rebuilt tools to justify snapshots of the system as follows.



Sample snapshots:  
 $S_1, S_2, S_3 \subseteq \Sigma^{obs}$ .

Figure 7.30: Sample snapshots.

A possible analysis based on the sample snapshots  $S_1, S_2, S_3 \subseteq \Sigma^{obs}$  is the following (Figure 7.30):

- $\mathfrak{E}_{\mathfrak{S}_r^+}^b(S_2)$  contains all prescriptions of the security policies being actually in force which in full pertain to the snapshot  $S_2$ .  
 Since  $\mathfrak{E}_{\mathfrak{S}_r^+}^b(S_2) = S_4^+$ , thus  $S_4^+$  is the only prescription which in full belongs to the snapshot  $S_2$ .
- $\mathfrak{E}_{\mathfrak{S}_r^+}^\#(S_2)$  contains all prescriptions of the security policies being actually in force which *possibly* pertain to the snapshot  $S_2$ .  
 Since  $\mathfrak{E}_{\mathfrak{S}_r^+}^\#(S_2) = S_4^+ \cup S_5^+$ , thus only  $S_4^+$ ,  $S_5^+$  are the prescriptions which on the whole or in part belong to the snapshot  $S_2$ .
- Acceptable execution traces in  $\mathfrak{E}_{\mathfrak{S}_r^+}^\#(S_2) \setminus \mathfrak{E}_{\mathfrak{S}_r^+}^b(S_2) = S_5^+ \setminus S_4^+$  are *abstained* because they cannot be uniquely classified either as belonging to  $S_2$  or as not belonging to  $S_2$  with respect to the prescriptions of the security policies.
- $S_2 \not\subseteq \mathfrak{E}_{\mathfrak{S}_r^+}^\#(S_2)$  and so the execution traces in the subset  $S_2 \setminus \mathfrak{E}_{\mathfrak{S}_r^+}^\#(S_2)$  of  $S_2$  are anomalous. Moreover, since  $\mathfrak{E}_{\mathfrak{S}_r^-}^\#(S_2 \setminus \mathfrak{E}_{\mathfrak{S}_r^+}^\#(S_2)) = S_3^-$ , the execution traces in  $S_3^- \cap (S_2 \setminus \mathfrak{E}_{\mathfrak{S}_r^+}^\#(S_2))$  are actually unacceptable.

A similar analysis can be made in the case of the snapshot  $S_1$ .

notice that the snapshot  $S_3$  cannot be justified at all with the prescriptions and proscriptions of the security policies being actually in force

# Summary

The present thesis can be divided into three main parts. After the first two introductory chapters, Chapter 3–6 contains our theoretical results, whereas Chapter 7 contains their different applications. More precisely, this dissertation consists of the following parts.

The introduction (Chapter 1) contains the historical and philosophical background, the brief summary of our approach and, finally, the thesis overview and our main results.

Chapter 2 summarizes the basic concepts and notations used throughout the thesis.

Chapter 3 defines two general approximations frameworks, a large-scaled initial one, called the Initial Approximation Framework, and a finer-scaled one, called the General Set Theoretic Approximation Framework. These frameworks allow us to treat the common features of the classic rough set theory and its possible generalizations uniformly. This chapter also introduces the notion of the generalized approximation framework.

*The first version of the notion of the generalized approximation space appeared in my paper [14]. I published in its present form in [20] (joint work with Tamás Mihálydeák).*

Chapter 4 contains the basic concepts and properties of the classic rough set theory relying on the General Set Theoretic Approximation Framework developed in the previous chapter. We cite only notions and statements which are required in our subsequent work. We partly restate these well-known facts in the language of our approximation framework and provide new point-free proofs for a few of them.

*Point-free proofs of a few properties of the rough set theory reconstructed in the generalized approximation space can be found in my papers [13], [14] and [16].*

Chapter 5 presents a special approximation framework which is based on the partial covering of the universe and fully integrated into the General Set Theoretic Approximation Framework. Our lower and upper  $\mathfrak{B}$ -approximations are the straightforward point-free generalizations of lower and upper  $\varepsilon$ -approximations relaying on  $\varepsilon$ -elementary sets. This chapter presents the basic results concerning the base system and the lower and upper  $\mathfrak{B}$ -approximations.

After some introductory remarks (Section 5.1), Section 5.2 defines the most fundamental concepts of our approach, the base system  $\mathfrak{B}$  and the family of  $\mathfrak{B}$ -definable subsets.

Section 5.3 introduces a constrained version of the base system, called the single-layered base system, with help of which we prove some properties of our approximation framework which in a sense are similar to the properties of classic rough set theory.

Section 5.4 defines the lower and upper approximations based on partial covering of the universe. First, we prove that these approximations fit into the General Set Theoretic Approximation Framework. We also show that—unlike Pawlak’s rough set theory—the  $\mathfrak{B}$ -definable property is generally not equivalent to the equality of lower and upper  $\mathfrak{B}$ -approximations.

Section 5.5 discusses the  $\mathfrak{B}$ -representations of the  $\mathfrak{B}$ -definable sets. If a subset  $D$  is  $\mathfrak{B}$ -definable, there may exist two or more families of  $\mathfrak{B}$ -sets such that their unions are equal to  $D$ . It is said that  $D$  is  $\mathfrak{B}$ -representable, if there exists exactly one such family of  $\mathfrak{B}$ -sets. We prove that all  $\mathfrak{B}$ -definable subset of the universe are  $\mathfrak{B}$ -representable if and only if the base system  $\mathfrak{B}$  is single-layered. In this case, we also give the explicit  $\mathfrak{B}$ -representations of  $\mathfrak{B}$ -definable subsets, especially, the lower and upper  $\mathfrak{B}$ -approximations.

Section 5.6 is about an especial important notion of the approximation of sets, namely, the exactness. In Pawlak’s approximation spaces the notions of ‘ $\varepsilon$ -crisp’ (i.e. the exactness) and ‘ $\varepsilon$ -definable’ are synonymous to each other. A set is  $\mathfrak{B}$ -crisp, i.e. exact, if its lower and upper  $\mathfrak{B}$ -approximations are equal. However, a  $\mathfrak{B}$ -definable subset is not necessarily  $\mathfrak{B}$ -crisp. Consequently, the notions of ‘definable’ and ‘crisp’ (exactness) are not synonymous to each other in our special approximation framework.

In Section 5.7, we give a possible interpretation of our approach.

*I first defined the notion of the base system and its special variants in [14]. The first definition of the lower and upper approximations based on partial covering appeared here too. A more detailed description of those notions and further results can be found in my paper [12].*

In Pawlak's rough set theory, it is a well known fact that the upper and lower approximations<sup>1</sup> form a Galois connection on the power set of the base set. The partial ordering on the power set is given by the set theoretic inclusion relation. With regard to the informatics significance of the Galois connection, it is important to clarify what conditions have to be satisfied by the upper and lower  $\mathfrak{B}$ -approximations so that they form a Galois connection on  $(2^U, \subseteq)$ . In Chapter 6, we investigate this problem.

In Section 6.1 we prove one of the main results of the thesis: the upper and lower  $\mathfrak{B}$ -approximations form a Galois connection on  $(2^U, \subseteq)$  if and only if the base system  $\mathfrak{B}$  is a partition of  $U$ .

In Section 6.2, we deal with partial lower and upper  $\mathfrak{B}$ -approximations.

The empty set may be the lower  $\mathfrak{B}$ -approximation of certain nonempty subsets provided that all singletons are not  $\mathfrak{B}$ -definable. Excluding to allow that the empty set to be the lower  $\mathfrak{B}$ -approximation of a nonempty subset, we obtain the partial variant of the lower  $\mathfrak{B}$ -approximation. A natural total extension of the partial lower  $\mathfrak{B}$ -approximation, of course, is the original lower  $\mathfrak{B}$ -approximation itself. We show that this extension is unique under some evident conditions.

The empty set may be the upper  $\mathfrak{B}$ -approximation of certain nonempty subsets provided the base system does not cover the universe. Excluding these cases we obtain the partial variant of the upper  $\mathfrak{B}$ -approximation. Of course, the lower approximation remains a total map in the future too. The question naturally arises whether the Galois connection generalize so that the partial upper  $\mathfrak{B}$ -approximation and lower  $\mathfrak{B}$ -approximation may form a Galois connection in a sense. Moreover, if the answer is yes, what conditions have to be fulfilled by them so that they form a Galois connection of this special type. The notion of the *partial* Galois connection borrowing from Miné is suitable for our purpose. We prove that the partial upper  $\mathfrak{B}$ -approximation and the lower  $\mathfrak{B}$ -approximation form a partial Galois connection if and only if the  $\mathfrak{B}$ -set are pairwise disjoint.

*The results concerning regular Galois connections were first published in my paper [17]. I proved the statements for partial Galois connections provided that the base system is single-layered in [14]. I generalized the statements concerning both regular and partial Galois connections for arbitrary approximation spaces based on partial covering in [12].*

---

<sup>1</sup>Since Galois connections are not necessarily symmetric, the order of the maps is important.

In Chapter 7, to demonstrate the effectiveness of our approach we present three real life examples.

The first application show the relationship of our approach with natural computing via a biological application. In particular, we show how our approach helps us to understand some behavioral features of the natural vegetation heritage of Hungary. This presentation is based on the so-called MÉTA program which is a recognition and evaluation system of the state of the natural and semi-natural vegetation heritage of Hungary.

The second application presents a general tool-based approximation framework. We observe a class of objects and, as usual, we suppose that there are some well-defined features which an object possesses or not. In practice, two relevant groups of objects can be separated. A group whose elements really possess some features in question and another group whose elements do not substantially possess the same features. In the proposed tool-based approximation framework, two approximation spaces are defined. Any proportion of the observed objects can *simultaneously* be approximated in the two approximation spaces.

The third example applies the tool-based approximation framework to model the Intrusion Detection Systems (IDS) in computer security. In accordance with this framework, safe and unsafe traces can be detected simultaneously.

*The model related to the MÉTA program appeared in my paper [12]. I published models for Intrusion Detection Systems in [18] and [15]. The former refers to mainly professional, whereas the latter to non-professional computer environment. Both generalized models are described in [19] (joint work with Tamás Mihálydeák).*



# Összefoglaló

A disszertáció három átfogóbb részre bontható. Az első bevezető fejezetek után a 3–6. fejezetek az elméleti eredményeket, míg a 7. a különböző alkalmazásokra mutat példát. Részletesebben a következő részekből áll.

Az 1. fejezet disszertációm témájának történeti és filozófiai háttérét mutatja be, majd röviden összefoglalom megközelítésem lényegét. Végül egy átfogó áttekintést adok a dolgozatról és ismertetem főbb eredményeim.

A 2. fejezet technikai jellegű rész. Tartalmazza mindazon fogalmakat és jelöléseket, amelyeket a disszertációban használok.

A 3. fejezetben két általános közelítő keretrendszert definiálok. Az első (kezdeti közelítő keretrendszer) csak néhány fontos feltételt rögzít. A második (általános halmazelméleti közelítő keretrendszer) az előzőre épül és jóval finomabb felbontású. A két keretrendszer segítségével egységes szempontok alapján tanulmányozhatók mind a klasszikus közelítő halmazelmélet (rough set theory), mind annak különböző lehetséges általánosításai. Erre számos példa van a szakirodalomban, az általam javasolt megoldás új.

Bevezetem az alaprendszer és a jól definiált halmaz fogalmát. A jól definiált halmazokról legáltalánosabban csak annyit tételezünk fel, hogy az üres halmaz, továbbá az alaprendszer elemei jól definiáltak. Meghatározom az alsó-felső közelítésekre vonatkozó minimum követelményrendszert. Ezek segítségével definiálok az általánosított közelítő tér fogalmát. Bemutatom, hogy az általánosított közelítő térben lényeges megállapítások tehetők függetlenül a jól definiált halmazok, illetve az alsó-felső közelítések konkrét képzési módjától.

*Az általánosított közelítő tér fogalmának első változata a [14] cikkemben jelent meg. Jelenlegi formájában [20]-ban publikáltam (közös munka Mihálydeák Tamással).*

A 4. fejezet lényegében a klasszikus közelítő halmazelmélet (rough set theory) rekonstruálása az általános halmazelméleti közelítő keretrendszerben. Elsősorban azokat az eredményeket mutatom be, amelyekre a későbbiekben szükség lesz. Ezeket részben átfogalmazom a keretrendszer terminológiájának megfelelően. A szakirodalom a Pawlak-féle közelítés tulajdonságait kizárólag pontszerűen (halmazok elemeivel) bizonyítja. Több állításra nem-pontszerű bizonyítást adok.

*Az általánosított közelítő térben rekonstruált közelítő halmazelmélet néhány tulajdonságának nem pontszerű bizonyítását a [13], [14] és [16] cikkeim tartalmazzák.*

Az 5. fejezet egy konkrét közelítő teret ismeret, amely teljes egészében integrálható az általános halmazelméleti közelítő keretrendszerbe. A  $\mathfrak{B}$  alaprendszer az univerzum egy parciális lefedése. A jól definiált halmazok, az ún.  $\mathfrak{B}$ -definiálható halmazok az alaprendszer tetszőleges uniói. Az alsó és felső  $\mathfrak{B}$ -közelítések a Pawlak-féle alsó és felső  $\varepsilon$ -közelítések közvetlen általánosításai. A fejezet az alaprendszerre és az alsó-felső  $\mathfrak{B}$ -közelítésekre vonatkozó legalapvetőbb eredményeket tartalmazza.

Néhány bevezető megjegyzés után (5.1 alfejezet), az 5.2 alfejezet a kiinduló alapfogalmakat, a  $\mathfrak{B}$  alaprendszert és a  $\mathfrak{B}$ -definiálható halmazokat tárgyalja.

Az 5.3 alfejezet az alaprendszer egy bizonyos szempontból korlátozott fogalmát vezeti be. Ez az egyszeres rétegű alaprendszer, amelynek segítségével a parciális lefedésen alapuló közelítő térben a klasszikus közelítő halmazelmélet (rough set theory) tulajdonságaira emlékeztető állítások bizonyíthatók.

Az 5.4 alfejezet definiálja a parciális lefedésen alapuló alsó-felső  $\mathfrak{B}$ -közelítések fogalmát. Először bebizonyítom, hogy a fogalom illeszkedik az általános halmazelméleti közelítő keretrendszerbe. Megmutatom továbbá, hogy – ellentétben a Pawlak-féle térrel – a  $\mathfrak{B}$ -definiálható tulajdonság általában nem ekvivalens az alsó-felső  $\mathfrak{B}$ -közelítések egyenlőségével.

Az 5.5 alfejezet a  $\mathfrak{B}$ -definiálható halmazok  $\mathfrak{B}$ -reprezentálhatóságát vizsgálja. Ha  $D$  egy  $\mathfrak{B}$ -definiálható halmaz, akkor az alaprendszer elemeiből álló több olyan halmazrendszer is lehet, amelyek uniója  $D$ . Ha csak egy ilyen van, akkor azt mondjuk, hogy  $D$   $\mathfrak{B}$ -reprezentálható. Bebizonyítom, hogy egy  $\mathfrak{B}$ -definiálható halmaz pontosan akkor  $\mathfrak{B}$ -reprezentálható, ha a  $\mathfrak{B}$  alaprendszer egyszeres rétegű. Ez utóbbi esetben explicite megadom a  $\mathfrak{B}$ -definiálható, ezen belül az alsó-felső  $\mathfrak{B}$ -közelítések  $\mathfrak{B}$ -reprezentációját.

Az 5.6 alfejezet a halmazközelítés kiemelten fontos fogalmáról, az egzaktságról szól. Egy halmaz pontos (egzakt), ha alsó és felső közelítése megegyezik. A Pawlak-féle közelítő térben az „ $\varepsilon$ -pontos” (egzakt) és „ $\varepsilon$ -definiálható” fogalma egybeesik. A parciális lefedésen alapuló közelítő térben azonban egy  $\mathfrak{B}$ -definiálható halmaz nem szükségképpen  $\mathfrak{B}$ -pontos is. Következésképpen a „pontosság” (egzaktság) és „definiálhatóság” fogalma nem egymás szinonimái a parciális lefedésen alapuló közelítő terekben.

Az 5.7 alfejezet a parciális lefedésen alapuló közelítő tér egy lehetséges interpretációját fejti ki.

*Az alaprendszer, illetve az egyszeres rétegű alaprendszer fenti fogalmát először [14]-ben definiáltam. Szintén itt jelent meg elsőként az alsó-felső közelítések definíciója parciális lefedésen alapuló közelítő terekben. A fogalmak lényegesen bővebb kifejtését és további eredményeket a [12] cikkem tartalmazza.*

A közelítő halmazelméletben (rough set theory) jól ismert tény, hogy a Pawlak-féle felső és alsó közelítések<sup>2</sup> Galois kapcsolatot alkotnak az alaphalmaz hatványhalmazán. A hatványhalmaz részenrendezése a halmazelméleti tartalmazás. A Galois kapcsolatok informatikai jelentőségére tekintettel fontos annak tisztázása, hogy a felső-alsó  $\mathfrak{B}$ -közelítések milyen feltételek mellett alkotnak Galois kapcsolatot  $(2^U, \subseteq)$ -n. A 6. fejezetben ezt a kérdést vizsgálom.

A 6.1 alfejezetben bizonyítom disszertációm egyik fő eredményét: a felső és alsó  $\mathfrak{B}$ -közelítések pontosan akkor alkotnak Galois kapcsolatot, ha a  $\mathfrak{B}$  alaprendszer az univerzum egy partíciója.

A 6.2 alfejezetben az alsó-felső  $\mathfrak{B}$ -közelítő leképezések bizonyos leszűkítéseit vizsgálom. Ha nem minden egyelemű halmaz  $\mathfrak{B}$ -definiálható, akkor az üres halmaz lehet nemüres halmaz alsó közelítése. Ha kizárjuk ezt a lehetőséget, akkor az alsó  $\mathfrak{B}$ -közelítés leszűkítését, egy parciális leképezést kapunk. Ennek a parciális leképezésnek viszont természetes teljes kiterjesztése az eredeti alsó  $\mathfrak{B}$ -közelítés. Bebizonyítom, hogy ez kiterjesztés egyértelmű bizonyos kézenfekvő feltételek fennállása mellett

Ha az alaprendszer nem fedi le az univerzumot, akkor az üres halmaz lehet nemüres halmazok felső  $\mathfrak{B}$ -közelítése. Kizárva ezt az lehetőséget az eredeti felső  $\mathfrak{B}$ -közelítés leszűkítéséhez, egy parciális leképezéshez jutunk. Mindez nem érinti az alsó közelítés definícióját, az továbbra is teljes. Kérdés, hogy a parciális felső és a teljes alsó leképezés alkothat-e Galois kapcsolatot

<sup>2</sup>Mivel a Galois kapcsolat nem szimmetrikus a leképezésekre nézve, ezért fontos a közelítések sorrendje.

valamilyen értelemben? A válaszhoz mindenképp a Galois kapcsolat megfelelően módosított fogalma szükséges. Az irodalomban megtalálható ez a fogalom, az ún. parciális Galois kapcsolat. Az eredeti kérdés tisztázása dolgozatom másik fő eredménye: a parciális lefedésben definiált parciális felső és teljes alsó közelítések pontosan akkor alkotnak parciális Galois kapcsolatot, ha az alaprendszer tagjai páronként diszjunktak.

*A reguláris Galois kapcsolatokra vonatkozó első eredményeim [17]-ben jelentek meg. Az alaprendszer egyszeres rétegű tulajdonságának feltételezése mellett [14]-ben bizonyítottam a parciális Galois kapcsolatokkal kapcsolatos állításokat. [12]-ben mind a reguláris, mind a parciális Galois kapcsolatokra vonatkozó tételeket általánosítottam, és azokat tetszőleges parciális lefedésen alapuló közelítő terekre bizonyítottam.*

A 7. fejezetben példákat mutatok a parciális lefedésen alapuló halmazközelítés lehetséges alkalmazásaira.

A 7.1 alfejezetben a MÉTA programhoz kapcsolódó modell egy biológiai példán keresztül mutatja be a parciális lefedésen alapuló halmazközelítés egy lehetséges alkalmazását a természetes elvű számítástechnikában (natural computing). A MÉTA program (Magyarország Élőhelyeinek Térképi Adatbázisa) egy 2003 és 2006 között lezajlott, Magyarország teljes területére kiterjedő felmérés volt az ország nagyléptékű aktuális élőhely-térképének elkészítésére. Modellemben a Pawlak-féle elmélet és a parciális lefedésen alapuló közelítés együttes alkalmazásával elemezhető az ország élőhelyeit veszélyeztető tényezők előfordulása, terjedése.

A 7.1 alfejezet egy általános eszköz-alapú keretrendszert mutat be. Objektumok halmazát vizsgáljuk, és feltesszük, hogy adott az objektumok két könnyen meghatározható, de elkülönült csoportja (ún. referencia halmazok). Az egyik rendelkezik bizonyos tulajdonságokkal, a másik pedig jellemző módon nem. Mindkét halmaz két különböző – szintén elkülönült – halmazrendszerrel közelíthető (eszközök). A referencia halmazok és az eszközök kölcsönös viszonyának elemzése révén két különböző parciális lefedésen alapuló tér alakítható ki. Ezt követően bármely objektumhalmaz egyidejűleg elemezhető mindkét parciális közelítő térben.

A 7.2 alfejezetben az általános eszköz-alapú modellt alkalmazom számítógépes rendszerekbe történő behatolást detektáló modell felállítására. Végeredményben a számítógépes rendszer kívülről megfigyelhető véges nyomsorozatainak halmazán két parciális közelítő tér alakítható ki, a biztonságos, illetve a nem biztonságos nyomsorozatok halmazának közelítésére. A nyomsorozatok halmazai egyidejűleg értékelhető ki mindkét térben.

*A MÉTA programhoz kapcsolódó modell a [12] cikkemben jelent meg. A behatolás detektáló modelleket [18]-ben és [15]-ben publikáltam. Az előbbi elsősorban professzionális, az utóbbi pedig nem professzionális környezetben működő számítógépes rendszerekre vonatkozik. Mindkettő általánosított modelljét a [19] cikk írja le (közös munka Mihálydeák Tamással).*



# Bibliography

- [1] MÉTA programme - Vegetation Heritage of Hungary. Available at <http://www.novenyzetiterkep.hu/?q=en/english/node/70>.
- [2] *Proceedings of the International Multiconference on Computer Science and Information Technology, IMCSIT 2009, Mragowo, Poland, 12-14 October 2009* (2009), Polskie Towarzystwo Informatyczne - IEEE Computer Society Press.
- [3] Glucose tests. In *Lab Tests Online. Glucose*, Produced by American Association for Clinical Chemistry. [Online] Available at <http://labtestsonline.org/understanding/analytes/glucose/tab/glance>, June 2011.
- [4] ANDERSON, R. J. *Security Engineering: A Guide to Building Dependable Distributed Systems*, 2nd ed. Wiley, 2008.
- [5] BAUER, L., LIGATTI, J., AND WALKER, D. More enforceable security policies. In *Foundations of Computer Security* (Copenhagen, Denmark, July 2002).
- [6] BEGHADAD, R. Modelling and solving the intrusion detection problem in computer networks. *Computers & Security* 23, 8 (2004), 687–696.
- [7] BISHOP, M. *Computer Security: Art and Science*. Addison Wesley, 2002.
- [8] BÖLÖNI, J., KUN, A., AND MOLNÁR, Z. *Élőhely-ismereti Útmutató (Habitat guide) (in Hungarian)*. MTA Ökológiai és Botanikai Kutatóintézete (Institute of Ecology and Botany of the Hungarian Academy of Sciences), Vácrátót, 2003.

- [9] BÖLÖNI, J., MOLNÁR, Z., ILLYÉS, E., AND KUN, A. A new habitat classification and manual for standardized habitat mapping. *Annali di Botanica nouva series* 7 (2007), 105–126.
- [10] CAELLI, W., LONGLEY, D., AND SHAIN, M. *Information security handbook*. Stockton Press, New York, 1991.
- [11] CIUCCI, D. Approximation algebra and framework. *Fundamenta Informaticae* 94 (2009), 147–161.
- [12] CSAJBÓK, Z. Approximation of sets based on partial covering. *Theoretical Computer Science* 412, 42 (2011), 5820 – 5833. Rough Sets and Fuzzy Sets in Natural Computing, DOI: 10.1016/j.tcs.2011.05.037 (IF=0.838).
- [13] CSAJBÓK, Z. Rudiments of partial approximative set theory (in hungarian). In *A Nemzetközi III. Nyíregyházi Doktorandusz (PHD/DLA) Konferencia Kiadványa* (2009), pp. 117–120.
- [14] CSAJBÓK, Z. Partial approximative set theory: A generalization of the rough set theory. In *Proceedings of the International Conference of Soft Computing and Pattern Recognition (SoCPaR 2010), December 7-10, 2010., Cergy Pontoise / Paris, France* (2010), T. Martin, A. K. Muda, A. Abraham, H. Prade, A. Laurent, D. Laurent, and V. Sans, Eds., IEEE, pp. 51–56. DOI:10.1109/SOCPAR.2010.5686424. Also published in IEEE Xplore Digital Library.
- [15] CSAJBÓK, Z. A security model for personal information security management based on partial approximative set theory. In *Proceedings of the International Multiconference on Computer Science and Information Technology (IMCSIT 2010), October 18–20, 2010., Wista, Poland* (Katowice, Poland – Los Alamitos, USA, 2010), M. Ganzha and M. Paprzycki, Eds., vol. 5, Polskie Towarzystwo Informatyczne – IEEE Computer Society Press, pp. 839–845. Also published in IEEE Xplore Digital Library.
- [16] CSAJBÓK, Z. On the partial approximation of sets. *Acta Medicinæ et Sociologica* I, 2 (2011), 143–152.
- [17] CSAJBÓK, Z. Partial approximative set theory: A view from Galois connections. In *Proceedings of the 8th International Conference on Applied Informatics (ICAI 2010), January 27–30, 2010, Eger, Hungary* (2011).



- [18] CSAJBÓK, Z. Simultaneous anomaly and misuse intrusion detections based on partial approximative set theory. In *Proceedings of the 19th International Euromicro Conference on Parallel, Distributed and Network-based Processing (PDP 2011), 9-11 February 2011, Ayia Napa, Cyprus* (Los Alamitos, California, 2011), Y. Cotronis, M. Danelutto, and G. A. Papadopoulos, Eds., IEEE Computer Society Press, pp. 651–655. DOI:10.1109/PDP.2011.47. Also published in IEEE Xplore Digital Library.
- [19] CSAJBÓK, Z., AND MIHÁLYDEÁK, T. A general tool-based approximation framework based on partial approximation of sets. In *13th International Conference on Rough Sets, Fuzzy Sets and Granular Computing (RSFDGrC-2011), June 25-27, 2011, Moscow, Russia* (2011), S. O. Kuznetsov and et al., Eds., vol. 6743 of *Lecture Notes in Artificial Intelligence*, Springer-Verlag Berlin Heidelberg, pp. 52–59.
- [20] CSAJBÓK, Z., AND MIHÁLYDEÁK, T. On the general set theoretical framework of set approximation. pp. 12–15. Proceedings of RST 2011, 14-16 September 2011, Milan, Italy (2011).
- [21] DAVEY, B. A., AND PRIESTLY, H. A. *Introduction to Lattices and Order*. Cambridge University Press, Cambridge, 2002.
- [22] DENECKE, K., ERNÉ, M., AND WISMATH, S., Eds. *Galois Connections and Applications*. Kluwer Academic Publishers, 2004.
- [23] DENNING, D. E. An intrusion-detection model. *IEEE Transactions on Software Engineering SE-13*, 2 (February 1987), 222–232.
- [24] DUBOIS, D., ESTEVA, F., GODO, L., AND PRADE, H. An information-based discussion of vagueness. In *Handbook of Categorization in Cognitive Science*, C. Lefebvre and H. Cohen, Eds. Elsevier, 2005, pp. 892–913. Part 8. Machine Category learning.
- [25] DÜNTSCH, I., AND GEDIGA, G. Approximation operators in qualitative data analysis. In *Theory and Applications of Relational Structures as Knowledge Instruments* (2003), H. C. M. de Swart, E. Orłowska, G. Schmidt, and M. Roubens, Eds., vol. 2929 of *LNCS*, Springer, pp. 214–230.
- [26] FEKETE, G., MOLNÁR, Z., AND HOTVÁTH, F., Eds. *A magyarországi élőhelyek leírása és hatáározókönyve. A Nemzeti Élőhely-*

- osztályozási Rendszer (Guide and description of the Hungarian habitats. The National Habitat Classification System) (in Hungarian)*. Természettudományi Múzeum, Budapest, 1997.
- [27] FREGE, G. *Grundgesetzen der Arithmetik, begriffsschriftlich abgeleitet*, vol. 2. Verlag von Hermann Pohle, Jena, 1903.
  - [28] GIERZ, G., HOFMANN, K., KEIMEL, K., LAWSON, J., MISLOVE, M., AND SCOTT, D. *Continuous Lattices and Domains*. No. 93 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2003.
  - [29] GOGUEN, J. A., AND MESEGUER, J. Security policies and security models. In *Proceedings of the IEEE Symposium on Research in Security and Privacy* (1982), IEEE Computer Society Press, pp. 11–20.
  - [30] HORVÁTH, F., MOLNÁR, Z., BÖLÖNI, J., PATAKI, Z., POLGÁR, L., RÉVÉSZ, A., OLÁH, K., KRASSER, D., AND ILLYÉS, E. Fact sheet of the MÉTA database 1.2. *Acta Botanica Hungarica 50(Suppl.)* (2008), 11–34.
  - [31] JÄRVINEN, J. Properties of rough approximations. *Journal of Advanced Computational Intelligence and Intelligent Informatics 9*, 5 (2005), 502–505.
  - [32] JÄRVINEN, J. Lattice theory for rough sets. In *Transactions on Rough Sets VI*, J. F. Peters, A. Skowron, I. Düntsch, J. W. Grzymała-Busse, E. Orłowska, and L. Polkowski, Eds., vol. 4374 of *LNCS*. Springer-Verlag, 2007, pp. 400–498.
  - [33] JÄRVINEN, J. Lattice theory for rough sets. In *Transactions on Rough Sets VI: Commemorating the Life and Work of Zdzisław Pawlak, Part I*, J. F. Peters, A. Skowron, I. Düntsch, J. W. Grzymała-Busse, E. Orłowska, and L. Polkowski, Eds., vol. 4374 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, Heidelberg, 2007, pp. 400–498.
  - [34] JÄRVINEN, J. Pawlak’s information systems in terms of Galois connections and functional dependencies. *Fundamenta Informaticae 75*, 1–4 (2007), 315–330.
  - [35] JÄRVINEN, J., KONDO, M., AND KORTELAINEEN, J. Modal-like operators in boolean lattices, Galois connections and fixed points. *Fundamenta Informaticae 76*, 1–2 (2007), 129–145.

- [36] JÄRVINEN, J., RADELECZKI, S., AND VERES, L. Rough sets determined by quasiorders. *Order* 26, 4 (2009), 337–355.
- [37] KARI, L., AND ROZENBERG, G. The many facets of natural computing. *Communications of the ACM* 51, 10 (2008), 72–83.
- [38] KEEFE, R. *Theories of Vagueness*. Cambridge Studies in Philosophy. Cambridge University Press, Cambridge, UK, 2000.
- [39] KEEFE, R., AND SMITH, P. Introduction: Theories of vagueness. In *Vagueness: A Reader* [40], pp. 1–57.
- [40] KEEFE, R., AND SMITH, P., Eds. *Vagueness: A Reader*. MIT Press, Cambridge, MA, 1996.
- [41] LAVRAČ, N., AND DŽEROSKI, S. *Inductive Logic Programming: Techniques and Applications*. Ellis Horwood, New York, 1994.
- [42] LIN, T. Y. Approximation theories: Granular computing vs rough sets. In *Rough Sets and Current Trends in Computing, Proceedings of RSCTC 2008*, (2008), C.-C. Chan, J. W. Grzymala-Busse, and W. Ziarko, Eds., vol. 5306 of *LNCS*, Springer, pp. 520–529.
- [43] LIN, T. Y. Granular computing: Practices, theories, and future directions. In *Encyclopedia of Complexity and Systems Science*, R. A. Meyers, Ed. Springer, 2009, pp. 4339–4355.
- [44] LIN, T. Y. Update and illustration on granular computing: Practices, theory and future directions. In *2010 IEEE International Conference on Granular Computing, GrC 2010, San Jose, California, USA, 14-16 August 2010* (2010), X. Hu, T. Y. Lin, V. V. Raghavan, J. W. Grzymala-Busse, Q. Liu, and A. Z. Broder, Eds., IEEE Computer Society, pp. 32–33.
- [45] MIHÁLYDEÁK, T. On tarskian models of general type-theoretical languages. In *Proceedings of the 7th Panhellenic Logic Symposium* (Patras, 2009), C. Drossos, P. Peppas, and C. Tsinakis, Eds., Patras University Press, pp. 127–131.
- [46] MINÉ, A. *Weakly Relational Numerical Abstract Domains*. PhD thesis, École Polytechnique, Palaiseau, France, December 2004.

- [47] MOLNÁR, Z., BARTHA, S., SEREGÉLYES, T., ILLYÉS, E., BOTTA-DUKÁT, Z., TÍMÁR, G., HORVÁTH, F., RÉVÉSZ, A., KUN, A., BÖLÖNI, J., BIRÓ, M., BODONCZI, L., DEÁK, A. J., FOGARASI, P., HORVÁTH, A., ISÉPY, I., KARAS, L., KECSKÉS, F., MOLNÁR, C., ORTMANN-NÉ AJKAI, A., AND RÉV, S. A grid-based, satellite-image supported multi-attributed vegetation mapping method (MÉTA). *Folia Geobotanica* 42 (2007), 225–247.
- [48] MOLNÁR, Z., BIRÓ, M., AND BÖLÖNI, J. Appendix - English names of the Á-NÉR habitat types. *Acta Botanica Hungarica* 50(Supl.) (2008), 249–255.
- [49] MOLNÁR, Z., BÖLÖNI, J., AND HORVÁTH, F. Threatening factors encountered: Actual endangerment of the hungarian (semi-)natural habitats. *Acta Botanica Hungarica* 50(Supl.) (2008), 199–217.
- [50] ODIFREDDI, P. *Classical Recursion Theory. The Theory of Functions and Sets of Natural Numbers*. No. 125 in Studies in Logic and the Foundations of Mathematics. Elsevier, 1989.
- [51] PAGLIANI, P., AND CHAKRABORTY, M. *A Geometry of Approximation: Rough Set Theory Logic, Algebra and Topology of Conceptual Patterns (Trends in Logic)*. Springer Publishing Company, Incorporated, 2008.
- [52] PAWLAK, Z. Some issues on rough sets. In *Transactions on Rough Sets I*, J. F. Peters, A. Skowron, J. W. Grzymala-Busse, B. Kostek, R. W. Swiniarski, and M. S. Szczuka, Eds., vol. 3100 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, pp. 1–58.
- [53] PAWLAK, Z. Information systems theoretical foundations. *Information Systems* 6, 3 (1981), 205–218.
- [54] PAWLAK, Z. Rough sets. *International Journal of Information and Computer Science* 11, 5 (1982), 341–356.
- [55] PAWLAK, Z. *Rough Sets: Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publishers, Dordrecht, 1991.
- [56] PAWLAK, Z. Vagueness - a rough set view. In *Structures in Logic and Computer Science* (1997), J. Mycielski, G. Rozenberg, and A. Salomaa, Eds., vol. 1261 of *Lecture Notes in Computer Science*, Springer, pp. 106–117.

- [57] PAWLAK, Z., POLKOWSKI, L., AND SKOWRON, A. Rough sets: An approach to vagueness. In *Encyclopedia of Database Technologies and Applications*, L. C. Rivero, J. Doorn, and V. Ferraggine, Eds. Idea Group Inc., Hershey, PA, 2005, pp. 575–580.
- [58] PAWLAK, Z., AND SKOWRON, A. Rudiments of rough sets. *Information Sciences* 177, 1 (2007), 3–27.
- [59] PEIRCE, C. S. Vague. In *Dictionary of Philosophy and Psychology* (New York, 1902), J. M. Baldwin, Ed., MacMillan, p. 748.
- [60] POMYKAŁA, J. A. Some remarks on approximation. *Demonstratio Mathematica*, 24, 1-2 (1991), 95–104.
- [61] PRIEST, G. A site for sorites. In *Liars and Heaps: New Essays on Paradox*, J. C. Beall, Ed. Oxford University Press Inc., New York, 2003, pp. 9–23.
- [62] REVETT, K., GORUNESCU, F., AND SALEM, A.-B. M. Feature selection in parkinson’s disease: A rough sets approach. In *IMCSIT* [2], pp. 425–428.
- [63] RUSSELL, B. Vagueness. *Australasian Journal of Philosophy and Psychology* 1 (1923), 84–92.
- [64] SALEM, A.-B. M., REVETT, K., AND EL-DAHSHAN, E.-S. A. Machine learning in electrocardiogram diagnosis. In *IMCSIT* [2], pp. 429–433.
- [65] SALOMAA, A. *Computation and automata*, vol. 25 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, New York, NY, USA, 1985.
- [66] SENIOR OFFICIALS GROUP-INFORMATION SYSTEMS SECURITY. *Information Technology Security Evaluation Criteria (ITSEC)*. Department of Trade and Industry, 1991.
- [67] SKOWRON, A. On topology in information system. *Bulletin of the Polish Academy of Sciences, Mathematics* 36 (1988), 477–479.
- [68] SKOWRON, A. Vague concepts: A rough-set approach. In *Proceedings of EUROFUSE 2004*, B. De Baets, R. De Caluwe, G. De Tré, J. Fodor, J. Kacprzyk, and S. Zadrozny, Eds. Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2004, pp. 480–493.

- [69] SKOWRON, A., ŚWINIARSKI, R., AND SYNAK, P. Approximation spaces and information granulation. In *Transactions on Rough Sets III: Journal Subline* (2005), J. F. Peters and A. Skowron, Eds., vol. 3400 of *LNCIS*, Springer, Heidelberg, pp. 175–189.
- [70] SORENSEN, R. Vagueness. In *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition) (E. N. Zalta, Ed.) [Online] Available at <http://plato.stanford.edu/archives/fall2008/entries/vagueness/>.
- [71] STADLER, P. F., AND STADLER, B. M. R. Genotype phenotype maps. *Biological Theory* 3 (2006), 268–279.
- [72] TAKÁCS, G., AND MOLNÁR, Z., Eds. *National Biodiversity Monitoring System XI. Habitat mapping*. MTA Ökológiai és Botanikai Kutatóintézet (Institute of Ecology and Botany of the Hungarian Academy of Sciences) (Vácrátót) and Környezetvédelmi és Vízügyi Minisztérium (Ministry of Environment and Water) (Budapest), Vácrátót, 2009.
- [73] VARZI, A. C. Cut-offs and their neighbors. In *Liars and Heaps: New Essays on Paradox*, J. C. Beall, Ed. Oxford University Press Inc., New York, 2003, pp. 24–38.
- [74] WANG, G., LONG, C., AND YU, W. Rough set based solutions for network security. In *Monitoring, Security, and Rescue Techniques in Multiagent Systems*, Advances in Soft Computing. Springer-Verlag, Berlin, Germany, 2005, pp. 455–465.
- [75] WANG, W., GUAN, X., ZHANG, X., AND YANG, L. Profiling program behavior for anomaly intrusion detection based on the transition and frequency property of computer audit data. *Computers & Security*, 7, 539–550.
- [76] XU, F., YAO, Y., AND MIAO, D. Rough set approximations in formal concept analysis and knowledge spaces. In *ISMIS* (2008), A. An, S. Matwin, Z. W. Ras, and D. Slezak, Eds., vol. 4994 of *Lecture Notes in Computer Science*, Springer, pp. 319–328.
- [77] YAO, Y., AND CHEN, Y. Rough set approximations in formal concept analysis. In *Transactions on Rough Sets V* (2006), J. F. Peters and A. Skowron, Eds., vol. 4100 of *Lecture Notes in Computer Science*, Springer, pp. 285–305.

- [78] YAO, Y. Y. On generalizing Pawlak approximation operators. In *Rough Sets and Current Trends in Computing* (1998), L. Polkowski and A. Skowron, Eds., vol. 1424 of *Lecture Notes in Computer Science*, Springer, pp. 298–307.
- [79] YAO, Y. Y. Information granulation and rough set approximation. *International Journal of Intelligent Systems* 16, 1 (2001), 87–104.
- [80] YAO, Y. Y. On generalizing rough set theory. In *Proceedings of the 9th International Conference Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing (RSFDGrC 2003)* (2003), LNAI 2639, Springer-Verlag, pp. 44–51.
- [81] YAO, Y. Y., AND LIN, T. Y. Generalization of rough sets using modal logics. *Intelligent Automation and Soft Computing, An International Journal* 2 (1996), 103–120.
- [82] ZADEH, L. A. Fuzzy sets. *Information and Control* 8, 3 (1965), 338–353.
- [83] ZADEH, L. A. Granular computing and rough set theory. In *Rough Sets and Intelligent Systems Paradigms, Proceedings of RSEISP 2007* (2007), M. Kryszkiewicz, J. F. Peters, H. Rybinski, and A. Skowron, Eds., vol. 4585 of *Lecture Notes in Computer Science*, Springer, pp. 1–4.
- [84] ZAKOWSKI, W. Approximations in the space  $(U, \Pi)$ . *Demonstratio Mathematica*, 16, 3 (1983), 761–769.
- [85] ZHU, P. Covering rough sets based on neighborhoods: An approach without using neighborhoods. *International Journal of Approximate Reasoning* 52, 3 (March 2011), 461–472.
- [86] ZHU, W. Topological approaches to covering rough sets. *Information Sciences* 177, 6 (2007), 1499–1508.
- [87] ZHU, W. Relationship between generalized rough sets based on binary relation and covering. *Inf. Sci.* 179, 3 (2009), 210–225.





## Appendix A

### List of Papers of the Author

- [1] CSAJBÓK, Z.: *Approximation of sets based on partial covering*, Theoretical Computer Science, Rough Sets and Fuzzy Sets in Natural Computing Special Issue, Volume 412, Issue 42, pp. 5820 - 5833, 2011. (Impact Factor = 0.838)

ISSN 0304-3975, DOI: 10.1016/j.tcs.2011.05.037

Also published in ScienceDirect

Available at <http://www.sciencedirect.com/science/article/pii/S0304397511004555>

- [2] CSAJBÓK, Z., MIHÁLYDEÁK, T.: *A General Tool-Based Approximation Framework Based on Partial Approximation of Sets*, in S. O. Kuznetsov, D. Slezak, D. H. Hepting, B. G. Mirkin, Eds.: Proceedings of 13th International Conference on Rough Sets, Fuzzy Sets and Granular Computing (RSFDGrC-2011), June 25-27, 2011, Moscow, Russia, Lecture Notes in Computer Science Series, Lecture Notes in Artificial Intelligence Subseries, Volume 6743, Springer-Verlag Berlin Heidelberg, pp. 52-59, 2011.

ISBN: 978-3-642-21880-4, ISSN 0302-9743

DOI: 10.1007/978-3-642-21881-1\_10

Indexed by ACM Digital Library

Available at

<http://www.springerlink.com/content/61r6g15v51831614/>

- [3] CSAJBÓK, Z.: *Partial Approximative Set Theory: A Generalization of the Rough Set Theory*, in: T. Martin, A. K. Muda, A. Abraham, H. Prade, A. Laurent, D. Laurent, V. Sans, Eds.: Proceedings of the International Conference of Soft Computing and Pattern Recognition (SoCPaR 2010), December 7-10, 2010., Cergy Pontoise / Paris, France, IEEE Press, pp. 51–56, 2010.

ISBN 978-1-4244-7897-2, DOI:10.1109/SOCPAR.2010.5686424

Indexed by IET Inspec

Also published in IEEE Xplore Digital Library

Available at [http://ieeexplore.ieee.org/xpl/freeabs\\_all.jsp?arnumber=5686424#](http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5686424#).

- [4] CSAJBÓK, Z.: *Simultaneous Anomaly and Misuse Intrusion Detections Based on Partial Approximative Set Theory*, in: Y. Cotronis, M. Danelutto, G. A. Papadopoulos, Eds.: Proceedings of the 19th International Euromicro Conference on Parallel, Distributed and Network-based Processing (PDP 2011), 9-11 February 2011, Ayia Napa, Cyprus, IEEE Computer Society Press, Los Alamitos, California, pp. 651–655, 2011.

ISBN 978-1-4244-7895-8, ISSN 1066-6192, DOI:10.1109/PDP.2011.47

Indexed by IET Inspec

Also published in IEEE Xplore Digital Library

Available at [http://ieeexplore.ieee.org/xpl/freeabs\\_all.jsp?arnumber=5739062](http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5739062).

- [5] CSAJBÓK, Z.: *A Security Model for Personal Information Security Management Based on Partial Approximative Set Theory*, in: M. Ganzha, M. Paprzycki, Eds.: Proceedings of the International Multiconference on Computer Science and Information Technology (IMC-SIT 2010), October 18-20, 2010., Wisła, Poland, vol. 5, Polish Information Processing Society, Katowice, Poland – IEEE Computer Society Press, Los Alamitos, USA, pp. 839–845, 2010.

ISBN 978-1-4244-6432-6, ISSN 2157-5525, DOI:10.1109/PDP.2011.47

Indexed by IET Inspec, DBLP

Also published in IEEE Xplore Digital Library

Available at [http://ieeexplore.ieee.org/xpl/freeabs\\_all.jsp?arnumber=5739062](http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5739062).

- [6] CSAJBÓK, Z., MIHÁLYDEÁK, T.: *On the general set theoretical framework of set approximation*, Proceedings of the 3rd Rough Set Theory Workshops 2011, 14-16 September 2011, Milan, Italy, pp. 12-15, 2011.
- [7] MIHÁLYDEÁK, T. CSAJBÓK, Z.: *Partial logic to investigate the logical behavior of different systems of approximations of sets*, Proceedings of the 3rd Rough Set Theory Workshops 2011, 14-16 September 2011, Milan, Italy, pp. 47-51, 2011.
- [8] CSAJBÓK, Z.: *Partial Approximative Set Theory: A View from Galois Connections*, in A. Egri-Nagy, E. Kovács, G. Kovásznai, G. Kusper, T. Tómacs, Eds.: Proceedings of the 8th International Conference on Applied Informatics (ICAI 2010), January 27–30, 2010, Eger, Hungary, UNIDEB Faculty of Informatics – Eszterházy Károly College, Debrecen, Eger, pp. 53–60, 2011.  
ISBN 978-963-9894-72-3
- [9] CSAJBÓK, Z.: *On the Partial Approximation of Sets*, Acta Medicinae et Sociologica, Debreceni Egyetem, Orvos- és Egészségtudományi Centrum, Egészségügyi Kar, Nyíregyháza, pp. 143-152, 2011.  
ISSN 2062 0284
- [10] CSAJBÓK, Z.: *A parciális közelítő halmazelmélet alapelvei (Rudiments of Partial Approximative Set Theory)* (in Hungarian), in: K. Vincze, Ed.: A Nemzetközi III. Nyíregyházi Doktorandusz (PHD/DLA) Konferencia Kiadványa, 2009. november 20., Nyíregyháza, Szent Atanáz Görög Katolikus Hittudományi Főiskola, Nyíregyháza, Hungary, pp. 117–120, 2009.  
ISBN 978-963-87809-6-6.



## Appendix B

### Additional Publications

- [1] Csajbók, Z.: *On the mappings of elliptic curves defined over  $\mathbb{Q}$  into  $[0, 1)^2$* , Acta Mathematica Academiae Paedagogicae Nyíregyháziensis, **23**(2), pp. 115–123, 2007

Reviewed by Zentralblatt für Mathematik

ISSN 1786-0091

Available at [http://www.emis.de/journals/AMAPN/vol23\\_2/amapn23\\_12.pdf](http://www.emis.de/journals/AMAPN/vol23_2/amapn23_12.pdf).

- [2] Csajbók, Z.: *On the Privacy and Formal Representations of Medical Guidelines*, in: The 23rd microCAD International Scientific Conference, 19-20 March 2009, Vol. R: Health Science, Innovation and Technology Transfer Centre, University of Miskolc, Miskolc, Hungary, pp. 95–98, 2009.

ISBN 978-963-661-883-4.

- [3] CSAJBÓK, Z.: *Orvosi protokollok információtechnológiai reprezentációja (Information Technology Representation of Medical Protocols)* (in Hungarian), in: Kerekes, B., Ed.: *Proceedings of the 2nd International Doctoral (PHD/DLA) Conference, 21 November 2008, Nyíregyháza, Hungary*, Bessenyei György Könyvkiadó, Nyíregyháza, Hungary, pp. 217–221, 2009.

ISBN 978-963-9909-19-9.

- [4] CSAJBÓK, Z., TAKÁCS, P., KRISTÓF, Z., BODNÁR, K.: *eLearning rendszerek biztonsági jellemzőinek vizsgálata (Examinations of Safety*

*Features of eLearning Systems*) (in Hungarian) in: Kerekes, B., Ed.: Proceedings of the 2nd International Doctoral (PHD/DLA) Conference, 21 November 2008, Nyíregyháza, Hungary, Bessenyei György Könyvkiadó, Nyíregyháza, Hungary, pp. 255–259, 2009.

ISBN 978-963-9909-19-9.

- [5] CSAJBÓK, Z.: *Azonosságon alapuló titkosítás kórházi információs rendszerekben (Identity Based Encryption in Hospital Information Systems)* (in Hungarian), in: Pethő, A., Herdon, M., Eds.: Proceedings of the Informatics in the Hungarian Higher Education 2008 Conference, 27-29, August 2008 Debrecen, Hungary, Faculty of Informatics, University of Debrecen, Debrecen, Hungary, 2008.

ISBN 978-963-473-129-0.

Available at <http://www.agr.unideb.hu/if2008/kiadvany/papers/D81.pdf>.

- [6] CSAJBÓK, Z.: *Elliptikus görbék és elliptikus oszthatósági sorozatok (Elliptic Curves and Elliptic Divisibility Sequences)* (in Hungarian), in: Proceedings of the 1st International Doctoral (PHD/DLA) Conference, 7 December 2007, Nyíregyháza, Hungary, Nyíregyháza, Hungary, 2007. 6 p.

- [7] CSAJBÓK, Z.: *A számelmélet néhány algebrai eszköze (On some Algebraic Tools of Number Theory (in Hungarian))*, in: Fábián, G., Ed.: Bulletin of Faculty of Health College, University of Debrecen, Vol. 3, Faculty of Health College, University of Debrecen, Nyíregyháza, Hungary, pp. 77–90, 2005.

- [8] CSAJBÓK, Z.: *On the Mappings of Elliptic Curves*, in: Pethő, A., Herdon, M., Eds.: Proceedings of the Informatics in the Hungarian Higher Education 2005 Conference, 24-26 August 2005, Debrecen, Hungary, Faculty of Informatics, University of Debrecen, Debrecen, Hungary, 2005, 5 p.

ISBN 963 472 909 6.

Available at <http://agrinf.agr.unideb.hu/if2005/kiadvany/papers/E13.pdf>.

- [9] Csajbók, Z.: *Kriptográfia a gazdaságinformatikai képzésben – Mit és hogyan? (The Cryptography in the Higher Education of the Business*

*Information Technology—What and How*) (in Hungarian), Journal of The Scientific and Educational Forum for Business Information Systems (GIKOF), **3**(4), pp. 49–56, 2004.

ISSN 1588-9130.

Available at <http://www.gikof.hu/GIKOF-3evf-4szam.pdf>.

- [10] CSAJBÓK, Z.: *Elliptikus görbék (Elliptic Curves)* (in Hungarian), in: M. Galó, Vass, Lajosné, Eds.: Szabolcs-Szatmár-Bereg Megyei Tudományos Közalapítvány füzetek 18., Szabolcs-Szatmár-Bereg Megyei Tudományos Közalapítvány Kuratóriuma, Nyíregyháza, Hungary, pp. 273–275, 2002.  
ISBN 963 206 192 6, ISSN 1215-7686.
- [11] Csajbók, Z.: *SDL Modelling of Cryptographic Systems*, A Magyar Tudományos Akadémia Szabolcs-Szatmár-Bereg Megyei Tudományos Testülete Tudományos Ülésének Előadásai. Nyíregyháza, 2002. szeptember 28-29., Vol. 1, pp. 263-268.
- [12] CSAJBÓK, Z.: *Objektum-orientált analízis oktatása a DOTE Nyíregyházi Egészségügyi Főiskolai Karon (Object-Oriented Analysis Training at the DOTE Health College of Nyíregyháza)* (in Hungarian), in: Csirik, J., Herdon, M., Eds.: Proceedings of the Informatics in the Hungarian Higher Education 1999 Conference, 27-29 August 1999, Debrecen, Hungary, Vol. II, Federation of Debrecen Universities, Debrecen, Hungary, pp. 608–613, 1999.  
ISBN 963 03 8320 9.
- [13] CSAJBÓK, Z., POZSGAI, S.: *Az 1%-os törvény adatvédelmi elemzése (The Data Protection Analysis of the 1% Act)* (in Hungarian), Magyar Közigazgatás, **47**(10), 1997, pp. 608–614,  
ISSN 0865-736X.
- [14] CSAJBÓK, Z., POZSGAI, S.: *Az 1%-os törvény adatvédelmi elemzése (The Data Protection Analysis of the 1% Act)* (in Hungarian), in: Cselényi, I., Ed.: A Magyar Tudományos Akadémia Szabolcs-Szatmár-Bereg Megyei Tudományos Testülete 1997. évi Tudományos Ülésén elhangzott előadások összefoglalói, Kiadja: Az MTA Szabolcs-Szatmár-Bereg megyei Tudományos Testülete megbízásából a Rím Kiadó, Nyíregyháza, Hungary, p. 137, 1997.

ISBN 963 7868 26 7.

- [15] CSAJBÓK, Z.: *Az Internet-jelenség: információs társadalom és az adatvédelem (The Internet Phenomenon: The Information Society and the Data Protection)* (in Hungarian), in: Proceedings of the Hungarian Information Security 1996 Conference (HISEC'96), 13-14 June 1996, Budapest, Hungary, Budapest, Hungary, pp. 86–89, 1996.

- [16] CSAJBÓK, Z.: *Számítógépes rendszerek veszélyeztetettsége (Threats to Computer Systems II.)*, (in Hungarian), *DETEKTOR plusz*, **3**(2), pp. 50–52, 1996.

ISSN 1217-9175.

- [17] CSAJBÓK, Z.: *Számítógépes rendszerek veszélyeztetettsége (Threats to Computer Systems)*, (in Hungarian), *DETEKTOR plusz*, **3**(1), pp. 50–51, 1996.

ISSN 1217-9175.

- [18] CSAJBÓK, Z.: *Bevezetés a számítástechnikába, I. Szoftver (Introduction to Information Technology, Vol. II, Software)*, (in Hungarian), Information Technology Educational Notebooks, Vol. 36, ÉGSZI-Scola Foundation, Budapest, Hungary, 1996. 52 p.

- [19] CSAJBÓK, Z.: *Bevezetés a számítástechnikába, II. Hardver (Introduction to Information Technology, Vol. II, Hardware)*, (in Hungarian), Information Technology Educational Notebooks, Vol. 35, ÉGSZI-Scola Foundation, Budapest, Hungary, 1996. 65 p.

- [20] CSAJBÓK, Z.: *Az Európa Tanács adatvédelmi egyezménye (The Data Protection Treaty of the European Council)* (in Hungarian), *DETEKTOR plusz*, **2**(6), pp. 32–35, 1995.

ISSN 1217-9175.

- [21] CSAJBÓK, Z.: *Bevezetés a számítástechnikába (Introduction to the Information Technology)* (in Hungarian), KOMPKONZULT, Budapest, Hungary, 1995, 107 p.

ISBN 963 7842 29 2

- [22] CSAJBÓK, Z.: *Bevezetés a magyar nyelvű EXCEL 4.0 táblázatkezelő használatába Introduction to Excel 4.0 (in Hungarian)*, Information



Technology Educational Notebooks, Vol. 33, ÉGSZI-Scola Foundation, Budapest, Hungary, 1995. 56 p.

- [23] CSAJBÓK, Z.: *Bevezetés a magyar nyelvű Word for Windows 6.0 szövegszerkesztő használatába (Introduction to Word for Windows 6.0)* (in Hungarian), Information Technology Educational Notebooks, Vol. 32, ÉGSZI-Scola Foundation, Budapest, Hungary, 1994. 72 p.
- [24] CSAJBÓK, Z.: *Számítógép és ellenőrzés (Computers and Audit)* (in Hungarian), in: Papp, G., Ed.: *Proceedings of the Hungarian Information Security 1994 Conference (HISEC'94)*, 8-11 June 1994, Budapest, Hungary, John von Neumann Computer Society, Budapest, Hungary, pp. 99–108, 1994.  
ISBN 963 8431 81 4.
- [25] CSAJBÓK, Z.: *Az Európa Tanács Adatvédelmi Egyezménye és a magyar adatvédelmi szabályozás (The Data Protection Treaty of the European Council and the Hungarian Data Protection Regulation)* (in Hungarian), *APEH HÍR-ADÓ*, **3**(11), pp. 33–36, 1994.  
ISSN 1216-5166.
- [26] CSAJBÓK, Z.: *Számítógépes bűnözés (Cyber Crime)*, (in Hungarian), *APEH HÍR-ADÓ*, **3**(7), pp. 17–19, 1994. ISSN 1216-5166.
- [27] CSAJBÓK, Z.: *Szakértő rendszerek és adóinformatika (Expert Systems and Tax Informatics)*, (in Hungarian), *APEH HÍR-ADÓ*, **3**(2), pp. 26–28, 1994. ISSN 1216-5166.
- [28] CSAJBÓK, Z.: *Adóinformatika és adatvédelem (Tax Informatics and Data Protection)*, (in Hungarian), in: Papp, G., Ed.: *Proceedings of the Hungarian Information Security 1993 Conference (HISEC'93)*, 7-9 December 1993, Budapest, Hungary, John von Neumann Computer Society, Budapest, Hungary, pp. 189–198, 1993.  
ISBN 963 8431 806
- [29] CSAJBÓK, Z.: *Relációs adatbázisok frame-alapú modellje (A Frame Based Model of Relational Databases)*, (in Hungarian), in: *Conference of Relational Databases*, 25-27 January 1993, Budapest, Hungary, FAIR Information Systems Ltd. – John von Neumann Computer Society, Budapest, Hungary, pp. 19–26, 1993.

- 
- [30] CSAJBÓK, Z.: *Számítástechnikai alapismeretek (Introduction to Information Technology)*, (in Hungarian), Information Technology Educational Notebooks, Vol. 19, ÉGSZI-Scola Foundation, Budapest, Hungary, 1992. 56 p.
- [31] CSAJBÓK, Z.: *A számítástechnika története (The History of the Information Technology)* (in Hungarian), Information Technology Educational Notebooks, Vol. 21, ÉGSZI-Scola Foundation, Budapest, Hungary, 1991. 56 p.
- [32] CSAJBÓK, Z., JÁNOSA, A., FRIEDRICH P., LAPIS, P.: *Számítástechnika az adóellenőrzésben (Information Technology in Tax Audit)* (in Hungarian), College of Finance and Accountancy, Budapest, Hungary, 1991. 191 p.

## Appendix C

### List of talks of the author

- [5] CSAJBÓK, Z., MIHÁLYDEÁK, T.: *On the general set theoretical framework of set approximation*,  
The 3rd Rough Set Theory Workshops 2011, 14-16 September 2011, Milan, Italy, pp. 12-15, 2011.
- [6] MIHÁLYDEÁK, T. CSAJBÓK, Z.: *Partial logic to investigate the logical behavior of different systems of approximations of sets*,  
The 3rd Rough Set Theory Workshops 2011, 14-16 September 2011, Milan, Italy
- [1] CSAJBÓK, Z., MIHÁLYDEÁK, T.: *A General Tool-Based Approximation Framework Based on Partial Approximation of Sets*  
13th International Conference on Rough Sets, Fuzzy Sets and Granular Computing (RSFDGrC-2011), June 25-27, 2011, Moscow, Russia
- [3] CSAJBÓK, Z.: *Simultaneous Anomaly and Misuse Intrusion Detectors Based on Partial Approximative Set Theory*  
The 19th International Euromicro Conference on Parallel, Distributed and Network-based Processing (PDP 2011), 9-11 February 2011, Ayia Napa, Cyprus
- [2] CSAJBÓK, Z.: *Partial Approximative Set Theory: A Generalization of the Rough Set Theory*  
International Conference of Soft Computing and Pattern Recognition (SoCPaR 2010), December 7-10, 2010., Cergy Pontoise / Paris, France

- 
- [4] CSAJBÓK, Z.: *A Security Model for Personal Information Security Management Based on Partial Approximative Set Theory*  
International Multiconference on Computer Science and Information Technology (IMCSIT 2010), October 18-20, 2010., Wisła, Poland
- [7] CSAJBÓK, Z.: *Partial Approximative Set Theory: A View from Galois Connections*  
The 8th International Conference on Applied Informatics (ICAI 2010), January 27–30, 2010, Eger, Hungary
- [8] CSAJBÓK, Z.: *A parciális halmazelmélet alapelvei (Rudiments of Partial Approximative Set Theory)* (in Hungarian)  
III. Nemzetközi Doktori (PHD/DLA) Konferencia (The 3rd International Doctoral (PHD/DLA) Conference), 2009. november 20, Nyíregyháza, Hungary

## Appendix D

# Acknowledgements

I would like to thank Tamás Mihálydeák and Attila Pethő for being excellent supervisors. I am grateful their numerous insightful hints and suggestions.

I wish to thank my colleagues, especially József Ködmön and Péter Takács, for their encouragement and assistance.

I would like to thank Faculty of Health, University of Debrecen for the financial support.

Last, but not least I would like to thank, first of all, my wife Margó, my children Dóra and Zoli for their endless love and patience throughout working on my thesis. I also wish to thank my younger sister Marika and my younger brother Csaba for their encouragement and assistance.



## Appendix E

# Köszönetnyilvánítások

Köszönöm Pethő Attilának és Mihálydeák Tamásnak, hogy kiváló témavezetőim volt. Hálás vagyok a számtalan élelisméjű ötletért, iránymutató javaslatért.

Köszönöm munkatársaimnak, mindenekelőtt Ködmön Józsefnek és Takács Péternek a biztatást és segítségüket.

Köszönöm a Debreceni Egyetem Egészségügyi Karának az anyagi támogatást.

Végül, de nem utolsó sorban, köszönöm elsősorban feleségemnek, Margónak, továbbá gyermekeinek, Dórának és Zolinak a végtelen türelmet és szeretetet, amellyel lehetővé tették e disszertáció elkészítését. Köszönöm húgomnak, Marikának és öcsémnek, Csabának a biztatást és segítségüket.