Diss. ETH No 15433

PARTIAL ORDER THEORY IN THE ASSESSMENT OF ENVIRONMENTAL CHEMICALS: FORMAL ASPECTS OF A PRECAUTIONARY PRE-SELECTION PROCEDURE

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZÜRICH
for the degree of
Doctor of Natural Sciences

presented by
OLIVIER SCHUCHT
Dipl. natw. ETH
born January 6, 1974
from Zürich (ZH)

accepted on the recommendation of
Prof. Dr. Ulrich Müller-Herold, examiner
Prof. Dr. Konrad Hungerbühler, co-examiner
Prof. Dr. Nils-Eric Sahlin, co-examiner

2004

Abstract

Man-made chemicals have been shown over recent decades to exhibit unwanted effects on a global scale. The possible occurrence of such global effects have turned out to be difficult to model with traditional assessment procedures. Alternative, non-traditional procedures have been proposed in recent years as a response to this challenge. The present thesis proposes to develop such an alternative assessment procedure, namely one that is based on the concept of exposure (the physical occurrence of a substance in the environment), rather than being based exclusively on known effects (which are inherently difficult to predict). This approach meets a restricted definition of the precautionary principle. Exposure of a chemical is presently modeled with two scenarios, while ensuring that the assessment procedure herein developed can allow for additional scenarios to be taken into consideration at a later stage.

The validity of non-traditional assessments that aim at assessing chemical substances beyond known adverse effects is often doubted both inside and outside of the scientific community. In order to respond to these reservations the present thesis establishes a formal setting. As this details each step in the definition of the procedure, this permits a more systematic discussion. It also facilitates the study of the procedure for formal contradictions and helps to contain the use of political input. This setting also allows to develop tools which can be used to decide if the procedure is adequately defined, and consequently to isolate the critical areas in the definition of the procedure if such should emerge.

The formal setting is found to be very similar to what is encountered in classical problems of decision making (i.e. in decision theory), such as election problems. Importantly, the present work shows that well-understood results from these related areas can be used due to this kinship. Such results are used in the thesis, and a number of illuminative insights are gained from their application: a number of derivations are made based on these results, which include the fact that the formal setting presently chosen requires the assessment procedure to be calibrated with the help of a reference. The calibration is conducted with the use of information on existing regulatory status of chemicals. In order for the calibration to be performed, a number of mostly technical assumptions are made.

It is shown that the assessment procedure is not adequately defined by the formal setting and the (technical) assumptions, if only existing regulatory status is taken as a reference. This demonstrates the importance of the formal setting: the discussion of the procedure within this setting permits to show that the procedure is not suitable (i.e. not adequately defined), a result which would else be difficult to arrive at. The formal frame of the assessment and certain assumptions therein made must be revisited as a consequence of this result. This clearly weakens the procedure from a formal point of view. The necessary adjustments, however, do not appear to be far-fetched and it appears that the definition of an assessment procedure that is based on the concept of exposure (and hence embodies the precautionary principle) can be arrived at by choosing an appropriate formal setting and by using existing regulatory status of reference chemicals for the calibration of the assessment.

The results for calibration are critically dependent on stability, and hence cannot be used in a carefree manner: a sensitivity and robustness analysis fortunately shows no major problem with respect to stability.

Despite the fact that a suitable procedure can be described, a number of critical areas are isolated, which should draw more attention. These areas concern the technical definition of the threat scenarios, as well as the number of classes that the assessment procedure is required to produce (e.g. ”innocuous”, ”alarming” etc.).

In order to demonstrate the resulting assessment procedure, a set of test chemicals is assessed with the procedure in the last part of the present work. It turns out that on one hand a number of brominated and chlorinated chemicals exhibit large environmental threat, whereas on the other hand a number of chemicals that are discussed for global regulation (such as silicones and carbamates) are innocuous with respect to global threat (under consideration of the two scenarios herein used).

From a more methodological point of view, the present thesis proposes a novel, simple graphical tool to test the consistency of a procedure with respect to the requirements made to it, and introduces special types of partially ordered sets (weak orders and interval orders) to the assessment of environmental chemicals.

Zusammenfassung

Seit einigen Jahrzehnten sind im Zusammenhang mit Umweltchemikalien stehende globale Schäden bekannt. Die Modellierung des Auftretens solcher Schäden hat sich im Rahmen herkömmlicher Bewertungsmethoden als äusserst schwierig erwiesen. Entsprechend sind in den letzten Jahren verschiedene Methoden mit dem Ziel entwickelt worden, solchen Schwierigkeiten entgegenzutreten. Die vorliegende Arbeit stellt eine solche Methode vor. Im Gegensatz zu den gängigen Methoden, die sich auf bekannte Schäden in Verbindung mit Exposition (der Gegenwart einer chemischen Substanz) beziehen, beruht diese Methode ausschliesslich auf dem Konzept der Umwelt-Exposition. Ausgehend von einer kürzlich veröffentlichten Studie der Euro-päischen Union zur möglichen Anwendung des Vorsorgeprinzips, wird dieser Ansatz in der Arbeit weiterentwickelt. Der Ansatz kann als eine stark eingeschränkte Interpretation des Vorsorgeprinzips verstanden werden. Die Modellierung der Exposition erfolgt mittels zweier Gefährdungsszenarien, wobei die Methode speziell für die Mitberücksichtigung weiterer Szenarien, welche zusätzliche Dimensionen des Exposition erfassen, angelegt ist.

Die Zulässigkeit von nicht-traditionellen Methoden bei der Klassifizierung von Umweltchemikalien hat in den letzten Jahren zu grossen Kontroversen geführt. Um diese zu entschärfen, führt die vorliegende Arbeit einen speziellen, formalen Rahmen ein. Dieser erlaubt eine systematische Diskussion der Methode dadurch, dass die einzelnen Schritte in der Definition der Methode genau beschrieben werden können. Somit kann die Methode einerseits auf Widersprüche untersucht werden, andererseits aber auch vor dem Einfliessen politischer Inhalte soweit wie möglich bewahrt werden. Der formale Rahmen ermöglicht ebenfalls die Ent-wicklung von Kriterien, anhand welcher die Anwendbarkeit der Methode beurteilt werden kann.

Formal erweist sich die Methode als verwandt mit klassischen Problemen der Entscheidungstheorie (wie z.B. der Definition von Wahlverfahren oder Wohlfahrtsfunktionen). Diese Verwandschaft erschliesst der vorliegenden Arbeit eine Reihe von formalen Hilfsmitteln, welche im Kontext der Bewertung von Umweltchemikalien zu wichtigen Einsichten führen: darunter die Erkenntnis, dass das Bewertungsmodell auf eine technische, besondere Art und Weise geeicht werden muss. Die Eichung wiederium setzt eine Anzahl technischer Annahmen voraus. In der vorliegenden

Arbeit wird sie bezüglich des vorherrschenden globalen regulatorischen Status von Referenzchemikalien durchgeführt.

Bei der Kalibrierung stellt sich heraus, dass im Fall der Anwendung der eben erwähnten Referenz, der formale Rahmen zusammen mit den technischen Annahmen das Bewertungsmodell nicht ausreichend definieren. Dieses Resultat illustriert die Bedeutung formaler Aspekte, ohne welche dieser Mangel nur schwierig auszumachen sein würde. Das Resultat schränkt die Zulässigkeit des Bewertungsmodells zwar klar ein. Diese Schwierigkeit kann aber mit wenigen Änderungen der Annahmen, sowie des Hinzuziehens zweier zusätzlicher Annahmen entkräftet werden. Es zeigt sich also, dass ein expositions-bezogenes Bewertungsmodell definiert werden kann, das zentralen Forderungen der Rationalität sowie der Anwendbarkeit genügt. Dasselbe gilt somit auch für die Operationalisierung des Vorsorgeprinzips in der Bewertung von Umweltchemikalien. Eine Sensitivitätsanalyse, sowie eine Untersuchung der Robust-heit zeigen im übrigen, dass diese Resultate weitgehend stabil sind.

Die notwendigen Änderungen am Bewertungsmodell stellen allerdings nur *eine* Möglichkeit dar. Die kritschen Punkte, die sich im Rahmen der Analyse als problematisch erweisen und weitere Aufmerksamkeit verlangen, betreffen einerseits die Forderung nach drei Klassen für das Resultat der Bewertung und andererseits eine Annahme zur technischen Definition der Gefährdungsszenarien. Diese Punkte sollten in Zukunft angemessenere Aufmerksamkeit verdienen.

In einem letzten Teil der Arbeit wird das Bewertungsmodell auf eine Liste ausgesuchter Testchemikalien angewendet. Dabei zeigt sich, dass eine Reihe bromierter und chlorierter Substanzen ein grosses Potenzial zur globalen Gefährdung aufweisen, während andere Umweltchemikalien (Silikone und Karbamate), die ebenfalls gelegentlich als Kandidaten für eine globale Regulierung diskutiert werden, kein solches Potenzial aufweisen.

In methodologischer Hinsicht führt die Arbeit ein graphisches Hilfsmittel ein, um die Widersprüchlichkeit von Forderungen zu diskutieren bzw. um eine geeignete Aggregationsfuntion zu finden. Die Arbeit führt ebenfalls zwei spezielle Teilordnungen, die schwache Ordnung und die Intervalordnung, erstmalig in die Bewertung von Umweltchemikalien ein.

Contents

Chapter 1

Introduction

The formal dicussion of decision problems in the context of rational choice originated in the 18^{th} century over election problems [1], and has ever since experienced a formidable growth in applications. Such applications have to a large part concentrated on economics and related social sciences [2].[*] But many of the tools developed and results that have been derived for the purpose have turned out to be quite general in their validity, and the formal discussion of decision problems within the setting of decision theory quickly spread to other areas. In recent decades this development greatly accelerated, and decision theory has expanded into the natural sciences. This thesis uses the setting of decision theory to discuss a particular problem in the environmental sciences, namely the assessment of environmental chemicals (*xenobiotics* or *man-made* chemicals).

1.1 Assessment of Environmental Chemicals

Environmental chemicals have been known for a few decades to exhibit adverse effects the extent of which has been disastrous on a global scale. The most prominent among these chemicals are the chlorinated hydrocarbons, with their depleting effect on the ozone layer, as well as the insecticide DDT, which has resulted in widespread contamination and thinning of egg-shells of sea-eagles [4].

Disastrous events on a global scale have proved to be difficult to avoid with traditional methods of assessment. Indeed, assessments of the possibility of such

---

[*] For a recent and entertaining introduction to general decision problems and their relation to elections see [3].

events occurring (which more generally include technological, epidemiological, climatic events etc.) have been described as featuring special characteristics such as the occurrence of new types of uncertainties surrounding the nature of the events themselves. By virtue of such particularities, risks related to global disastrous events can be shown to be of a new kind that is distinct from most risks known to date [5]. As a consequence, attempts have been made to establish new decision rules, which could avoid such events to take place in the future and thereby offer a tool to address such risks. Among the more prominent of these rules is the so-called "precautionary principle".

Attempts to establish new chemical assessments based on such rules (e.g. in national legislations, see Chapter 2) have typically been greeted with scepticism and confusion. This highlights the importance of transparent approaches to establishing such rules, as such approaches should only permit discrediting of such rules on clear and clearly articulated grounds. Seen from the decision makers eyes, it is of considerable importance to place the procedure in a setting which enables the decision maker to differentiate between procedures that are suitable and such that are not. The latter case can occur e.g. when the assessment procedure is *inconsistent*. As an example may serve a procedure that claims to benefit human society while concomitantly prohibiting most medically active substances from being used, thereby running against its primary aim. This is an arguably trivial example of inconsistency. Inconsistencies can, however, appear in a much more subtle way[†], and hence a careful examination of the properties of a procedure should not be shunned. But inconsistencies are only (extreme) examples among the many problems that can be encountered with decision rules. At the least then and in a very basic way, a formal setting is required in order to enable a systematic and transparent discussion of an assessment procedure, for the analysis of whatever problems the various stakeholders deem necessary to discuss.

From a general point of view, the assessment of chemicals is a decision problem. It has also been shown that such a chemical assessment has partial orders as its most basic mathematical structure. A number of attempts have been made in recent years to discuss the assessment of chemicals using this latter, more formal structure (see [6]–[10]), or at least by adopting the more general mathematical formalism of decision theory [11]. To the best knowledge of the author, no assessment procedure that allows to deal with global chemical risk, however, has to date been discussed

---

[†] This is known at least since Arrow's proof of the Impossibility Theorem [87]

that uses the concept of partial orders within the frame of decision theory. The current thesis hence proposes to establish such a procedure, to discuss it in the according formal setting and thereby to understand the *formal limitations* of the procedure, if such exist.

1.2 Objectives of Thesis

The main objective of this thesis is to explore the possibility of defining an assessment procedure for environmental chemicals that addresses the limits of current practice, i.e. the uncertainty surrounding adverse effects of chemicals. To this end, the assessment uses the concept of exposure, in line with previous work from the author's group. The use of this concept in the assessment of chemicals has "precautionary" qualities.[‡] The current work does not, however, claim to discuss fully the precautionary principle, or even to find the most suitable among its various definitions. Nevertheless, should a satisfactory assessment based on exposure result from this study, it could also serve as a manifestation of the existence of suitable and well-understood applications of the precautionary principle in the assessment of chemicals.

As we have already pointed out, in order to develop and discuss such an assessment in a formal way, the problem must be placed in a more general context. An important objective of the work hence consists in the elaboration of the appropriate formal frame, within which the assessment procedure can be discussed. The approach should also, if possible, make use of formal results from other decision problems which, although they may appear to be very different in nature, are structurally related.[§] (This applies in particular to results related to consistency of decision rules.) Once the formal setting has been found, the key step of selecting a (mathematical) decision rule should be performed. This step should not be performed prior to knowledge about the formal setting, as the latter may open the possibility of using decision rules that would otherwise not have been considered.

The thesis should also offer a set of tools that allow to pronounce a judgement on

---

[‡] Indeed the concept of exposure was further developed by the author's group during a two-year project funded by the European Union on the application of the precautionary principle in the assessment of chemicals [12]. This project produced a preliminary version, from which the assessment procedure developed herein was later advanced. The procedure is therefore designed to embody a particular definition of the precautionary principle.

[§] such as is the case for the problem of election procedures whether a particular assessment procedure is suitable or not, relating to the question of consistency and beyond, to other types of problems which may surface. This naturally entails a more precise understanding of what is meant by *suitable*[¶]. Without prematurely delving into this question (which may appear somewhat abstract at this point), it may suffice to say that the procedure is suitable if within the formal frame that is established for the procedure, the procedure is consistent and no further assumptions have to be made in order for the procedure to be *applicable*. Conversely, an assessment procedure can be said not to be suitable if it is not applicable without making assumptions in addition to the ones made in the formal frame of the procedure.[‖] Note that the formal frame may necessitate a few assumptions to be made. These should, however, be technical and use as little as possible political input.

Once the settings and tools have been developed, the procedure should be critically discussed in a formal and systematic way. It may then turn out that the procedure is not suitable within the formal frame. In this case the discussion should be conducted in a way as to give an indication to what extent assumptions are implicitly required by the formalism to be modified (or even added), in order to arrive at a procedure that is suitable.

Note that besides the formal aspects of the work, the actual definition of the precautionary approach and how this definition is arrived at, are key points for the present task of defining a precautionary assessment procedure.

1.3 Structure and Main Points of Investigation

Much of the present work is a continuation of research on proxy-measures for persistence and spatial range in the Ecological Risk Prevention program started by Ulrich Müller-Herold in 1993. These exposure-based indicators (*threat indicators*) have been proposed as measures for the large-scale threat potential of organic chemicals, independent of toxicity, and reflect a particular interpretation of the precautionary principle.

This first of two theoretical blocks in this work details the choice of an appropriate version of the precautionary principle (*choice of approach*): it concerns the

---

[¶] or appropriate, meaningful

[‖] The procedure as developed in the course of the document, will be called *formal precautionary procedure*.

traditional assessment of chemicals, its limits, and the possible application of the precautionary principle to the assessment of environmental chemicals. The block will also give a brief overview of different definitions of the principle and of the controversies related to it. Then a brief summary will be given on the derivation of threat indicators from the concept of exposure, which has been performed in recent years and is well-documented in the literature [26]. As an additional step, threat scenarios will be derived from threat indicators and introduced as the basic unit on which the assessment is conducted.[**]

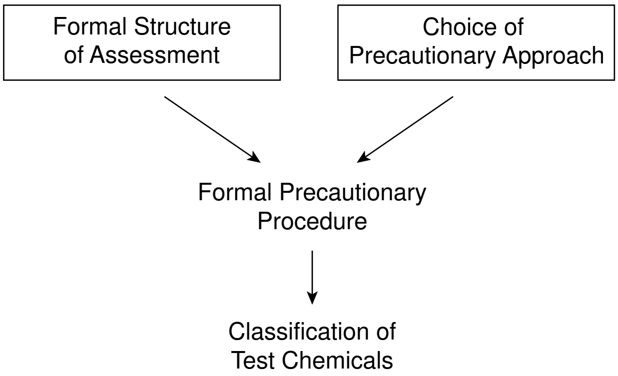

Figure 1.1: Overall structure consists in the combination of the formal structure of the assessment problem with a particular choice of precautionary approach, which leads to a formal precautionary (assessment) procedure.

In a separate, second block the formal setting is chosen for the assessment and a *formal assessment structure* is established. This part introduces the mathematical formalism and notation, which will be used throughout the text (Chapter 3). The following chapter (Chapter 4) starts with a brief overview on how the description of a set of chemicals can be simplified in view of an assessment. It then introduces a number of requirements which apply to the current assessment problem. It furthermore derives consequences and corrolaries from the requirements and thus establishes a *classification scheme*, which has to be respected by the classification we propose. Among the implications, it is shown that the formal structure requires an input in the form of some reference.[††] This chapter also illustrates the close kinship of the assessment of environmental chemicals with seamingly unrelated problems such as

---

[**] Threat scenarios have been described previously, too [11], but not quite systematically against the background of exposure.

[††] This is shown to be a necessity in order to conclusively fix the assessment procedure.

that of fair voting and other decision theoretical problems. The block concludes with the introduction of a linearisation function (i.e. the actual mathematical decision rule) that is compatible with the formal assessment structure.

The two theoretical blocks are *combined* in Chapter 6: The *choice of precautionary approach* is combined with the *formal structure of assessment*, resulting in a *formal precautionary procedure*. At this point the choice of approach is checked for compatibility with the requirements made in the formal part. A main result of the consolidation will consist in the observation that the assessment procedure requires a number of additional assumptions to be made, which are related to a calibration problem. Such assumptions are introduced in Chapters 6 and 7 and join an assumption made earlier in Chapter 2 (concerning the definition of threat scenarios).

Before moving to the actual classification and ensuing discussion, an algorithm will be introduced in Chapters 7 and 8 that performs the calibration (which is shown in Chapter 4 to result as an implication from the formal precautionary procedure). Also, a tool will be developed on the basis of the calibration, in order to adjudicate upon the suitability of the procedure. The discussion of the actual result of the calibration will be conducted in Chapter 9. Subsequent to such discussion and possible adjustments (in order to make the procedure suitable), the assessment procedure will be fully defined. At this point, in order to illustrate how environmental chemicals can be classified with respect to environmental threat (even in the case where adverse effects are only insufficiently understood) an assessment of a set of test chemicals will be performed (Chapter 10). The thesis closes with concluding remarks in Chapter 11.

The organisation of the work will help to address three main problems which reflect the objectives of this thesis set forth in Section 1.2.

Problem 1 *What is an adequate formal structure and what is an adequate (mathematical) decision rule for an exposure-based environmental assessment of man-made (xenobiotic) chemicals?*

Problem 2 *Based on the formal precautionary procedure,*[‡‡] *can a suitable assessment procedure be extracted solely from existing regulations?*

Problem 3 *What does the exposure-based assessment procedure yield when applied to a set of appropriate test chemicals?*

[‡‡] pro memoria; the formal procedure reflects, among other things, the particular definition of the precautionary principle made in the present work.

The main problems 1-3 will be gradually narrowed down in sense and answered in the course of the thesis. Problem 1 is addressed in Chapter 3, 4 and 5 and a summary of the discussion is given at the end of the Chapter 4 for the first part of the problem and at the end of Chapter 5 for the second part. Problem 2 is partly answered in Chapter 5 (namely to establish that such a procedure is, in principle, consistent) and later and in more detail in Chapter 9 with the help of criteria developed in Chapter 7.[*] Finally, Problem 3 is answered in Chapter 10, where the assessment of a set of test chemicals will be discussed.

---

[*] In order for the reader to understand the (rather technical) second problem, we must anticipate that the definition of the precautionary principle is undertaken with respect to possible global damage. (In turn, possible global damage is taken into consideration via threat scenarios.) Hence Problem 2 really deals with the combination of the requirements and assumptions made by the formal structure of the assessment, together with other assumptions made through the precautionary approach chosen in the present work.

Chapter 2

Chemical Risk Assessment and the Precautionary Principle

The present chapter aims first at giving a brief overview of existing risk assessment of chemicals inside the EU as paradigmatic for traditional assessments. In a second part, the chapter looks at the limits of traditional assessments and introduces the precautionary principle as a possible remedy to these limitations. A short discussion of various definitions of the precautionary principle as well of inherent weaknesses of the said principle, is followed by the introduction of the concept of environmental threat potential. It is shown that this concept embodies the principle in a restricted way, and that the concept rests on threat scenarios.

2.1 Chemical Risk Assessment

The first European efforts in establishing regulation for chemicals dates back to 1967, when Council Directive 67/548/*EEC* [13] was introduced. Numerous amendments were subsequently made to the directive. Importantly, the seventh amendment to the Directive introduced in 1992 made risk assessment of chemicals an inherent part of the application and admission of new chemicals [14]. It is therein mentioned that a request must be submitted for any new chemical introduced into the market inside a EU (formely EEC) country, and that the competent authorities are required to conduct a risk assessement. In 1993, the European Commission introduced a directive that detailed procedures for the risk assessment of newly introduced chemicals [15]. The directive was supplemented in 1994 by a Commission Regulation, which concerned chemicals that were already on the market [16, 17].[*]

By obvious necessity, a distinction between "existing" and "new" chemicals was introduced at that time. Thereby were defined as "existing" such chemicals that had been in use before a key date set out in the sixth amendment of Directive $67/548/EEC$, whereas chemicals introduced after that date were defined as "new".[†] The partition thereby introduced put approximately $100'000$ chemicals in the category "existing" (which are listed in the central database [20]). The group of "new" chemicals that were introduced between 1981 and 1993 consists of around 3000 chemicals. (In addition to that there are $300-400$ chemicals introduced each year.)

In order to facilitate the actual risk assessment required, the Commission also released a document on technical implementation of risk assessment (the Technical Guidance Document; TGD [21]).

Therein, the commission explicitly lists four criteria for the risk assessment of chemicals: hazard identification, effect assessment, exposure assessment and risk characterisation. These criteria had first been introduced in 1983 by the U.S. National Academy of Science.

2.2 Current Chemical Risk Assessment and Its Limitations

Under the Technical Guidance Document, the manufacturer or importer of a new chemical substance has to provide the authorities with a technical file. Based on this data (the amount of which can vary in accordance to the volume of production and the pattern of release), the competent authorities are obliged to conduct a risk assessment both on human health risks and on environmental risks. The present section focuses only on the risks pertaining to the environment.

Broadly speaking, the Technical Guidance Document establishes the following rule: on one hand, the risk assessment procedure must give an estimate of the concentration in the environment (*exposure assessment*), on the other hand, based on the adverse-effects known for the chemical, the procedure must give an estimate of the concentration necessary in the environment in order for adverse-effects to manifest

---

[*] Recently, the EU Commission has made a new proposal for Chemical Risk Assessment [18]. We will not discuss this proposal at this point, as practical details are not known yet.

[†] The key date was set as the 18 September, 1981 [19].

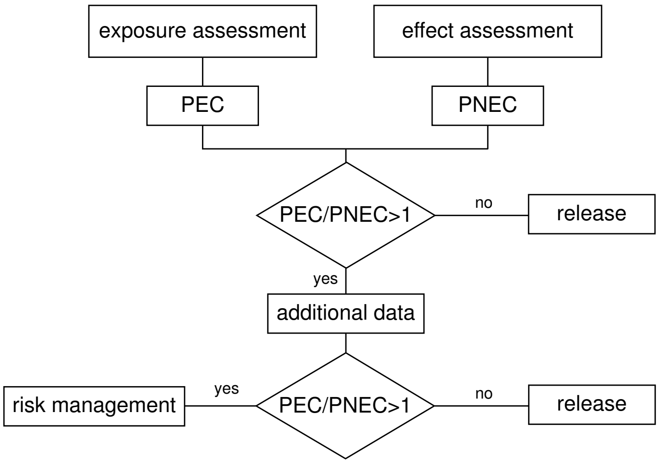

Figure 2.1: Scheme of chemical risk assessment such as is applied inside the EU, to-date.

themselves (*effect assessment*). In order to execute the latter assessment, it is first necessary to establish the adverse-effects themselves, the measurement of which can then be performed at various concentrations (*hazard identification*).[‡] Finally, once the result from the exposure as well as the effect assessments are available, the risk assessment procedure must perform a *risk characterisation*, in order to undertake regulation of the substance or other required steps.

Hazard Identification

Various tests must be performed, in order to identify the substances most important toxic effects (*hazards* or *adverse-effects*). They include several tests for acute toxicity, tests for bioaccumulation and degradation in the environment, screening tests for adsorption/desorption as well as assessment of bacterial inhibition.

Exposure Assessment

The *exposure assessment* yields the *P̲redicted E̲nvironmental C̲oncentration* (*PEC*). This output is modelled based on the amount of substance emitted, the release compartment and pattern of the emission, the degradation rates for the substance, as well as various physico-chemical properties of the substance.

[‡] Although the TGD and other documents speaks of *hazards*, we will use the term *adverse-effects* throughout this thesis, whenever we refer to a toxic effect of the human or more general environmental kind

Effect Assessment

The *effect assessment* yields the *Predicted No-Effect Concentration* ($PNEC$), i.e. the concentration at which no adverse-effect is expected to occur in the environment. The $PNEC$ is based on the *acute effect concentration* that is derived form *toxicity* data[§] from laboratory experiments. The $PNEC$ is simply the result of the division of the *acute effect concentration* by an arbitrary *assessment factor* of 1000.

Risk characterisation

The *risk characterisation* is based on a comparison of the indicators PEC and $PNEC$: By dividing the former by the latter, the *risk quotient* $PEC/PNEC$ decides on which actions are to be taken.

1. If $PEC/PNEC < 1$, the concentration is expected to be lower than the concentration at which no adverse-effects can be measured, and the substance is authorised.

2. If $PEC/PNEC > 1$, the substance can occur in a concentration at which it may exhibit adverse-effects: the substance must be further investigated, i.e. new data on emission pattern is generated, and additional toxicity tests are performed. The procedure is then repeated, if with a lower assessment factor due to reduced uncertainty in the estimate of PEC and $PNEC$.

Limits of Current Risk Assessment

The most serious of shortcomings of *existing* chemical risk assessment is that of the enormous number of chemicals which have been on the market before 1981. Of around $100'000$ existing chemicals, at least $30'000$ are estimated to be produced in volumes exceeding $1t$ per year [22]. Of these chemicals, about 140 have been placed on a priority list, from which in turn only 21 have been completely assessed and have obtained final regulation to date [23]. Furthermore, the existing risk assessment is not only slow, but also costly: it is estimated that the cost for the assessment for a new substance is in the range of $100k - 10Mio$ Euros.[¶]

The EU Commission has identified the shortcomings in a White Paper published in 2001, whereby it identifies the main problem as the fact that the authorities carry the burden of assessment, instead of industry. Accordingly, the main change

---

[§] Note that toxicity is used here in the widest possible sense, i.e. representing any adverse-effect that can occur in the environment

[¶] Cost are particularly high for chemicals that are released in large quantities.

proposed in a recent proposal for the regulation of chemical substances, consists in a re-positioning of the burden of assessment on industrial stake-holders [24]. This proposition has been critised by various interest groups, as well as by observers as simplistic.

The current risk assessment of both existing and new chemicals has also been attacked on grounds of more fundamental problems related to effect assessment [25, 26, 27]. The effect assessment is therein described as fundamentally flawed, due to an *intrinsic overcomplexity* of the environment. Overcomplexity, in turn, leads to difficulties or even to the impossibility of predicting adverse-effects. Hence an assessment based on effect may be of limited use, since the actual adverse-effects of a chemical may turn out to be of an unknown kind. In opposition to these considerations, risk assessment as it is practiced today departs form the assumption that all possible adverse effects can be known and in fact are known at the time of the assessment. Coupled with a few more technical limitations of effect assessment, it is presented as questionable whether meaningful $PNEC$-values can be derived that encompass the full extent of *potential damage* that may emanate from a chemical substance. (Indeed, the assessment factor mentioned earlier bears testimony to this fact.) In short, facing the problem of global scale adverse-effects of chemicals, traditional chemical risk assessment in its present form appears to be of limited value, as such effects may well not be known at the time of the assessment [28].

In the face of the limits of current chemical risk assessment, the *precautionary principle* has been advocated for a number of years as an additional tool in assessing the risk related to chemical. Indeed the principle has found entry to the regulation of chemicals in a number of European countries, such as in Nordic countries [29] and in Germany [30]. On a pan-European, as well as on a global scale, the principle has been discussed more recently as a possible approach in the risk assessment of chemical risks. E.g. the EU Commission's White Paper on The Precautionary Principle, although it does not mention chemical risk assessment explicitly, defines the applicability of the precautionary principle in a way as to make the risk assessment of a chemical substance subject to it [31]. More recently, in the EU Commission's Proposal for a new regime of chemical regulation, the principle does explicitly appear [24].

Before proceeding with a more detailed discussion of the precautionary principle in the form in which it is used in the present work, we will first look at the principle from a more general point of view and introduce the adversaries of the principle as well as a general definition of the principle. Importantly, we will also look at fundamental limitations which have been identified for the principle.

2.3 The Precautionary Principle

History

The *precautionary principle* originates from German legislation, where it is first mentioned as *Vorsorgeprinzip* in a legal text in 1976. The first mentioning of the principle in the anglo-saxon world ocurred in 1981 [30, 32, 33].

During the 1980's, the promotion of the principle received significant momentum, primarily in conjunction with an increasing awareness of *scientifc uncertainty* [34]. This development was triggered by a series of adverse conditions (mainly related to human health) that were of a new, unknown type (e.g. HIV). In face of traditional risk assessment, which dealt deterministically with known conditions, these suddenly emerging phenomena posed an important challenge (for a historical review on precautionary principle and emerging biological risks see [35]). Concomitantly, a number of adverse effects with respect to the environment — such as the depletion of the ozone layer [36], the egg-shell thinning of sea-eagles due to DDT [4] and the degradation of forest wetland [37] — triggered similar considerations [34].

It was realised at that time that, scientifically speaking, there was an imporant distinction between uncertainty of the probabilistic type (*Unsicherheit*) and uncertainty related to the unknown (*Ahnungslosigkeit*). The difference between the two had been pointed out early in other fields, most notably in the 1920's by Frank Knight, albeit in an economic context [38].[‖]

The first mentioning of the precautionary principle in a an international treaty occurred with the declaration of the Paris Convention in 1992, in which the precautionary principle is explicitly cited [39, 40]). The principle also finds mentioning in the same year in the Rio Declaration [41].

More recently, and in the context of risk assessment of chemicals, the precautionary principle has been cited repeatedly in controversies surrounding chemical substances, such as persistent organic pollutants (POPs), endocrine disruptors (ER) and phta-

---

[‖] This distinction is introduced in [38], particularly at the end of chapter *VII*: *The Meaning of Risk and Uncertainty.*

lates.[**] In fact, the principle has been introduced into several national legislations in the context of the regulation of chemicals. To date, however, the principle has not been implemented, nor explicitly mentioned in any international treaty dealing specifically with environmental chemicals.

Generally speaking, the precautionary principle has gained considerable prominence in environmental sciences and more generally in risk management, during the last two decades or so. As much as the principle has been welcomed from many decision makers and interest groups as a useful decision tool, it has, however, also been greeted with scepticism if not stiff resistance from various interest groups.

Adversaries of the Precautionary Principle
The precautionary principle has been attacked from various fronts, including from industrial groups, but also from researchers and political scientists. The main criticism has often revolved around the lack of specific definition of the principle. Indeed in one publication [43], the author mentions more than 14 different definitions given in a single survey (see also [44] and more generally [45]). Furthermore the principle has been attacked as a thin veil for protectionism [43, 46], a cause of innovation paralysis [47], as well as on grounds that it is ideological [48] or that the application of the principle in most cases reflects little more than a mask for inaction by competent authorities or a mask for political pressure of other sorts [49].

Perhaps most gravely, the principle has been repeatedly attacked for being unscientific [48, 50, 51, 52, 53]. The principle, it is argued, permits to take decision (including regulatory action) based on little more than anectodal evidence. Indeed, if the principle is interpreted in the sense that no new technologies (including chemical substances) can be used before absolute proof of safety, then no new technology could meet the condition, as one author pointed out [43].

Recent work has shown that much of the controversy surrounding the principle is due to the wide range of definitions existing for the principle. Indeed, the adversaries of the principle tend to attack the principle in its most extreme form, and recent contributions by advocators of the principle have focused on detailed analysis and more stringent definitions of the principle [45].

[**] There exists no comprehensive review article on the application of the precautionary principle to the assessment of chemicals, to the knowledge of the author. Instances of application are advocated and/or critically discussed in references [42]-[54].

Definitions

As we have pointed out earlier, the precautionary principle has its root in the German *Vorsorgeprinzip* (see also [54], cited in [43]). Numerous attempts have been made since to define the principle, and one of the more prominent of which is the definition of the principle as adopted in during the Wingspread Conference:

> When an activity raises threats to the environment or human health, precautionary measures should be taken, even if some cause-and-effect relationships are not fully established scientifically. (cited in [55])

Although it is not within the scope of the present thesis to conduct an extensive survey on definitions of the principle, we deem it useful to introduce a general frame, within which many definitions can be discussed. The frame was proposed by Sandin ([56], where besides the many definition discussed, further literature is indicated). Therein, 4 dimensions (or components) are identified, which are shown to be shared by most definitions of the principle (for a more involved discussion of these dimensions see also [26]):

1. threat dimension
2. uncertainty dimension
3. action dimension
4. command dimension

These dimensions being found in some variation in most definitions of the principle, Sandin derives a general definition of the principle, which reads as follows [56]:

> If there is (1) a threat, which is (2) uncertain, then (3) some kind of action (4) is mandatory.

Depending on how the four dimensions are chosen, Sandin notes that the definition of the principle can range from weak versions to self-defeating stronger versions. For the purpose of the present work, we keep hold of the fact that in this frame of analysis, any definition we choose must address the two descriptive dimensions *threat* and *uncertainty*, as well as the more normative dimensions *action* and *command*.

By *threat*, is meant some undesired state of the world. (We will in the following speak of *environmental threats* in this context, and when not specified, *threat* will have this more restricted meaning.) *Uncertainty* has a wide range of different meanings, from milder forms of uncertainty about probabilities or incomplete proof of supposed cause-and-effect relations (cf. *Unsicherheit*, above) to the most extreme form of *ignorance of ignorance* where the kind of possible unwanted effects itself is unknown (cf. *Ahnungslosigkeit* above). As we have already seen, this latter possibility applies to environmental chemicals, for which the complete spectrum of possible adverse effects is but insufficiently known.

Note that the *action* and *command* dimension can be linked to each other. In the present case they are simply reflected in the regulation of a chemical substance, in case the substance poses a threat/uncertainty significant enough. Since the present thesis proposes to find a pre-screening of chemicals, the action/command dimension presently consists in the decision whether to ban a chemical based on its significant environmental threat, or not.

Therefore, more weight will be placed on the *threat* and *uncertainty* dimensions in the following. Both dimensions are addressed and reflected in the concept of *environmental threat*[††].), as described in the following section.

2.4 Environmental Threat

Scheringer in [26] stressed that it is appropriate to section environmental contaminations into three stages: *emission*, *influences* (exposure, immission) and *consequences* (effects). Hence *influences* are a prerequisite for effects. As a consequence, as much as *effects* can constitute a threat,[‡‡] so can the "presence of anthropogenic agents in itself" [26, 57]; threats can exist and can be described without exact knowledge of adverse effects. A chemical assessment embodying this paradigm is termed an *exposure-based* assessment (see Figure 2.2), in contrast to more traditional *effect-based* assessments (see Figure 2.1).

The exposure-based assessment then relates to the possible occurrence of a chemical in the environment, whereas the effect-based assessment concerns the actual adverse-effects that are observed or modelled for a chemical. Hence by definition the choice of indicators, based on which an effect-based assessment can be conducted, poses no problem; it suffices to take the said (adverse) effects. The situation is more involved in the case of an exposure-based assessment.

[††] Alternatively, the term *endangerment* is found in the literature as a synonym [68]

[‡‡] i.e. an unwanted state of the environment

Exposure-based Assessment: Threat Indicators

Since the 1970's a number of indicators have been discussed as possible *indicators for exposure-based assessments.* Besides many simple physico-chemical properties (degradation constants etc.) that have been discussed as such indicators, the *persistence* of a chemical substance was established early on as a necessary (if not sufficient) dimension for environmental adverse effects [58, 59, 60, 61], thus establishing it as one such possible indicator. Among other indicators proposed in the frame of exposure-based risk assessment, is the *bioaccumulation potential* (especially after the emergence of adverse-effects caused by chlorinated hydrocarbons), and several measures of *mobility.*

Persistence

The rationale for the use of e.g. persistence in chemical assessment was first introduced by Scheringer, Berg and Müller-Herold [62]: The simple logical fact that the presence of a chemical is a necessary condition for adverse effects gives persistence a special importance for exposure-based assessment: if eventually it turns out, possibly a long time after release, that an apparently innocuous *persistent* chemical has adverse (biological) effects, it may become extremely difficult (costly) if not impossible (inestimably costly) to eliminate the chemical from the environment.

Spatial Range

In recent years, a new measure has been introduced as a threat indicator for exposure-based assessments, which bears considerable resemblance to the persistence: the *spatial range* of a substance. The concept was introduced in 1994, together with a model for calculation of the parameter [57, 63, 64, 65, 66]. A number of alternative concepts have been introduced in the meantime (for references see [23] p.19) and, at present, the concept of spatial range appears to be recognised as an important parameter for assessment of chemical substances (e.g. for plant protection products, see [67] p. 26). The model proposed by Scheringer yields both a measure for the persistence and a measure for the spatial range, based on chemical exposure patterns.*

---

*The present work will use the notation introduced by Müller-Herold, i.e. the symbol ρ for the spatial range. Scheringer himself uses the symbol R for the spatial range. The (technical) definition of persistence and the spatial range used in the present work are due to Müller-Herold and will be referenced later in Section 2.6.

Bioaccumulation Potential
The tendency to bioaccumulate has long been known to aggravate the consequences of adverse-effects, due to the difficulty of managing effects caused by an agent that is wide-spread and the emission of which can lie deep in the past.

Summary: Indicators
Summing up this brief overview on possible indicators for an exposure-based assessment, it is noted that a number of such indicators are presently known and well understood. Consequently, these indicators should lend themselves in one form or another to an *assessment procedure based on exposure.* The prerequisite for this last step is to ensure that indicators as they are used have a relation to normative criteria (i.e. criteria for decision), which, as is shown in [26], they appear to have.

Exposure-based Assessment: Procedures
The next step consists in defining a procedure. Indeed, Müller-Herold (in general terms) and Scheringer et al. (in more detail) have proposed such procedures [68, 69]. The incentive as well as the relevance of this early work is summed up succinctly in the following declaration by Kuhnreuther et al.:

> Risk denotes the combination of probability times magnitude of adverse effects. Conventional management strategies are designed to reduce either one of the two components. Müller-Herold argues that the precautionary principle in environmental policy making requires an approach that implies management steps at an earlier stage. His target is what he calls "endangerment"[†]. Controlling endangerment means controlling the scope and range of the potential for damage. He develops a taxonomy of endangerment that comprises two main factors: spatial extension and persistence over time. These two factors determine the degree of endangerment regardless of their strength in destructive potential. (from [70])

This taxonomy was further advanced and developed in [69] as a *pre-screening of chemicals.* In this procedure a pre-assessment is introduced (depicted in Figure 2.2) two pre-assessments (or pre-screenings) are performed in parallel: one related to exposure, the other to effect. In accordance to that there are two sets of indicators.

[†] For *endangerment* read *environmental threat potential* (Footnote by the author of the present thesis)

The first set consists of the exposure-based indicators persistence and spatial range. If only reduction of exposure is sought, then the application of these indicators is sufficient. (The use of the indicators in this sense also implies that the two indicators are accepted as sufficient *prioritisation criteria* (as noted in [23], p. 22. See also [27, 71]).)

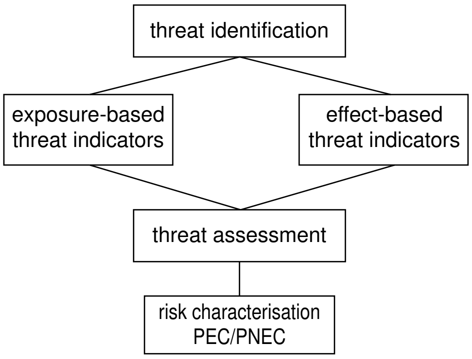

Figure 2.2: The scheme as proposed by Scheringer [69]. Upstream of the traditional chemical risk assessment are two additional assessment steps that are conducted in parallel: the exposure-based assessment that is based on exposure-based threat indicators and the effect-based assessment that is based on effect-based threat indicators.

In the same scheme in Figure 2.2, a second set of criteria is introduced, containing the bioaccumulation potential, as well as a measure for toxicitiy. (Again, this implies the assumption that the two indicators may serve as prioritisation criteria.[‡])

The exposure-based, as well as the effect-based indicators are treated independently by Scheringer. Each of the two classify chemicals according to whether they "pass" or "don't pass". In this way, the indicators are operationalised as *filters*: each filter takes values from two indicators (persistence and spatial range in one case, persistence and bioaccumulation potential in the other) and aggregates them to a single value, based on which the filter "retains" the chemical or not ([26], p. 206).

The approach described is taken one step further in the following, i.e. is compacted into a theory of threat scenario, where the scenarios constitute the basic unit of description of threat. To the extent that we are interested in a purely exposure-based assessment, it is sufficient to discuss one set of indicators, namely the exposure-based

---

[‡] As pointed out in [26] (on p. 205), the bioaccumulation potential has the distinct property of being suitable for use in both exposure-based and effect-based assessements.

indicators. The effect-based indicators used by Scheringer and also shown in the Figure 2.2 will be disregarded.

As a brief note on a more technical subject it should be pointed out that the discussion surrounding the concept of persistence and (to a lesser extent) of spatial range has been intense, and many models have been proposed. We will not dwell upon these at present, but point out to the literature cited in [26] and [67].

2.5 Threat Scenarios

In the exposure-based assessment introduced in the previous section, the two indicators persistence and spatial range are studied for a chemical. Based on the values of these two indicators, the chemical is classified. In the case of Figure 2.2, the assessment consists in either declaring a chemical as removed (on grounds of posing a threat), or releasing the chemical into further evaluation (i.e. into traditional risk assessment). Since both indicators are used for this evaluation, one indicator alone does not suffice to perform this classification, and both indicators must be considered in order to decide upon the fate of the chemical.

In the case of persistence and spatial range, this method has a clear-cut explanation: one has to take into account that longevity alone does not lead to possibly unmanageable situations (cf. concrete, bitumen, plastics, etc.). It is only in combination with other factors such as mobility that persistence is a significant indicator for large-scale chemical threat.

The combination of indicators as encountered so far is narrative: two indicators are studied concomitantly because there is an understanding that a combination of some value for one indicator with some other value for the second indicator represents a particular quality of the chemical, e.g. that is poses a threat. Such a concomitant study of indicators is termed a *threat scenario*. Scenarios are, however, not sufficient for the classification problem, since the classification must decide what exact combination of values of indicators correspond to a particular quality or state, an answer which the threat scenario cannot give. Hence the indicators must be combined to a filter, which – in the present case – decides what combinations of values allow a chemical to pass the filter and which not. The filter then is *a kind of mapping* of a threat scenario and the filter represents the quantification of the threat scenario that is required in order to enable classification.

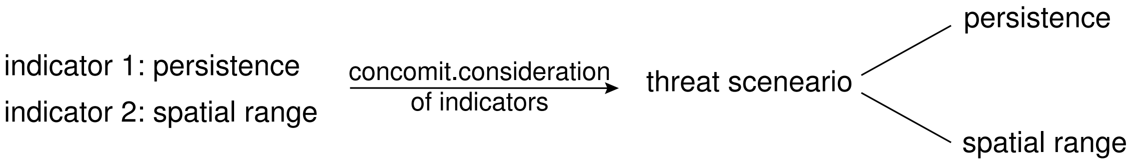

Figure 2.3: The assessment is depicted as based on two indicators (1 and 2). In case the two indicators are concomitantly considered, the basic unit changes to the threat scenario, which is constituted of two indicators (which now may be more correctly called parameters).

In this sense, it is the threat scenarios and not the indicators by themselves that enable the assessment to operate, and we may say that the scenarios (and representations thereof) are the *basic units* of description in the assessment (see Figure 2.3). Put differently, the scenarios generate the *characterisation* which enables the assessment procedure to be performed. Based on such characterisation, classes such as "posing significant environmental threat" or "not posing significant environmental threat" are defined by a corresponding filter. As such, the *filter* described by Scheringer are the *quantification of the threat scenarios*, and in this way the narrative threat scenarios are translated into a quantified representation.

Definition 1 *Filters are the (linear) quantified representations of threat scenarios.*

As is noted elsewhere in [23] (p. 23), the setting of cut-off values that is required for this classification defines the assessment procedure and can be the source of considerable disagreements among interest groups. The setting of such values constitutes the important step in the quantification, that of defining classes, for which different actions are then taken. We will come back to this very important point in Chapter 6 when we discuss the implications resulting from the *formal precautionary procedure*.

Also, it is important to note that the four indicators used in Scheringer's scheme could well be organised into other sets of scenarios, e.g. four scenarios (each defined completely by a single indicator), if it was believed that each indicator *on its own* reflected environmental threat. (Alternatively, if somewhat difficult to mentally picture, all indicators could be combined to a single scenario etc.).

2.6 Precautionary Approach: Exposure-based Threat Scenarios

Since the precautionary principle contains a threat dimension, a particular definition of the precautionary principle may be fixed by choosing particular threat scenarios. In this sense, the precautionary aspect enters *via the choice of threat scenarios* in the present approach. As we have just pointed out, a possibility to do so consists in choosing such threat scenarios that are constituted of exposure-based indicators. We have encountered three such indicators: τ, ρ and K_{ow}. The scenarios used herein (*Pandora* and *Bioaccumulation*) will precisely be constituted by these three indicators.

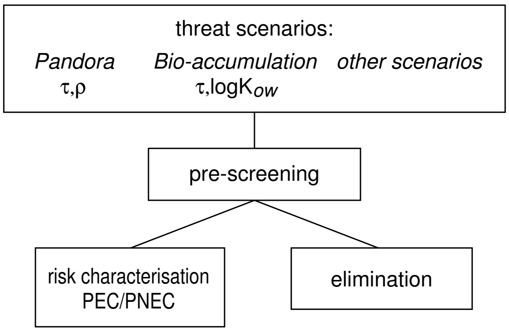

Figure 2.4: Choice of Precautionary Approach: An exposure-based assessment scheme, founded on the threat scenarios *Pandora* and *Bioaccumulation*.

Pandora

The Pandora scenario was named after the Greek myth of Pandoras box containing all evils and complaints [72]. In the myth, when the box was opened, all of its contents were unleashed upon the earth, causing irreversible harm. The corresponding threats in the context of environmental chemicals are characterised by uncertainty relating to probability of occurrence, type and extent of damage, and high persistency.[§]

The scenario of possible irreversible damage of very large reach has been proposed to be reflected by the two threat indicators *persistence* and *spatial range* [26], and these two will presently be used. The corresponding indicators are τ and ρ and are

---

[§] The enduring ubiquity of persistent organic pollutants (POPs) is regarded as the epitome of the Pandora scenario.

computed according to [66] and [73], respectively. (Concerning the symbol ρ see also the footnote on page 17.)

Bioaccumulation
Bioaccumulation [74] is a phenomenon combining bioconcentration and biomagnification. Bioconcentration relates to the partition of a chemical between an organism and a surrounding inorganic medium (e.g. leaves/air, fish/water). Biomagnification denotes the heterotrophic enhancement of concentration in subsequent elements of the food chain (grass/cow, cow/man). In heterotrophic aquatic species bioconcentration and biomagnification occur in parallel. The measure for persistence is τ (as for *Pandora*) and the *octanol-water* partition coefficient K_{ow} is used for bioaccumulation.[¶]

The threat scenario related to bioaccumulation is given through the possibility of substances to have adverse effects on living organisms even if their concentration in e.g. the oceans or the atmosphere is low: Bioaccumulating and persistent chemicals can built up high concentration in living organisms over time, although the concentration may initially have been low. In case of occurrence of adverse effects, the extent of damage may be as unmanageable as in the case of large spatial range and large persistence.

Since fat tissue is the relevant storage medium in an organism and since the partition of a chemical between water and organismic fat tissue is modeled through its octanol-water partition coefficient K_{ow}, this coefficient is one of the relevant indicators for bioaccumulation. In analogy to the Pandora case, the bioaccumulation filter is based on two indicators: a combination of high K_{ow} values and increased global characteristic persistence τ. (In order to bioaccumulate, a chemical has to survive a minimal period of time before degradation.)

Assumption 1 *The two threat scenarios Pandora and Bioaccumulation can be used for a precautionary assessment of chemicals. The two scenarios are each (best) defined by two indicators: for Pandora the indicators τ and ρ and for Bioaccumulation the indicators τ and K_{ow}.*

The precautionary approach defined by the choice of the two scenarios can be stated as follows: if a chemical exhibits either concomitant high values for persistence and

---

[¶] Details as to the source of Data and methods of computation for the indicators, are given in Chapter 7.

spatial range or for persistence and bioaccumulation potential, then it should be regulated, regardless of possible adverse-effects. The rationale can then be drawn as avoiding occurrence of events that are of such an extent as to make reparation almost impossible. It should not go unnoticed that such decision rule is akin to the *maximum probable loss rule.*[‖]

Other Scenarios
Another scenario frequently discussed is the *cold-trap*. Also, the Pandora scenario can be extended to include *transformation products* [67]. It is important therefore to ensure that the set of scenarios that are considered can be expanded, if additional scenarios are found that meet criteria of relevance.

Observation 1 *At present only two threat scenarios have been chosen, and such choice does not constitute a willful limitation. Rather, additional scenario could be considered if deemed relevant, in addition to Pandora and Bioaccumulation.*

2.7 Formal Limits of Precautionary Principles

Besides dismissive voices raised from various interest groups, the precautionary principle has been critically analysed in a more thorough manner by researchers in recent years. An important result with respect to the present work was derived by Peterson. In [76] Peterson proves that in case a number of *desiderata* are accepted (which are reflected in many definitions of so-called averse-decision rules), then there is no decision rule that satisfies all of these desiderata. These rules includes the precautionary principle, which is defined in the sense of the maximum probable loss rule (which is, as we just stated, is akin to the definition chosen in the present work).

This is of primary importance to the present work, as a formal precautionary procedure that is inconsistent is one that can be discarded from the outset. The logical inconsistency of the desiderata, however, is proven for definitions of the principle in which the *probability* of catastrophic events takes a prominent role. It seems that the inconsistency of the principle is limited to such probabilistic definitions of the principle.

[‖] For further references on this rule see [75].

This stems from the fact that the desiderata therein used are not meaningful in case the precautionary principle does not include probabilistic considerations and that in the absence of such definitions, the principle is not logically invalidated [77].[**]

Formal limits of other definitions need not be presently considered, as the choice of scenarios fix the type of definition quite narrowly. For a formal discussion of other definitions of the precautionary principle, see [11], among others.

[**] Peterson himself uses another method to circumvent the problem posed by the impossibility he has proven; by introducing two different types of decision rules: rigid and non-rigid rules [75, 76].

Chapter 3

Formal Structure of Assessments

The discussion of the formal structure of the assessment procedure starts with finding the appropriate theoretical setting, inside of which the discussion can be conducted. The setting presently chosen formalises the assessment problem as that of *mapping* a partially ordered set to a linear ordering. The mapping can also be viewed as a *function* which yields the assessment of a chemical and is often called a *decision rule.* The assessment procedure is in this sense captured and represented by an *assessment function* which we denominate by the symbol Π. To formalise the procedure as a mapping is canonical in decision theory, and this method has been used in a wide variety of decision problems (see [87] for the case of elections, but also any summary introduction to welfare functions). Since decision theory generally deals with the *classification* of objects, the term is used synonymously with assessment in the following.[*]

Partially ordered sets are a concept from the Theory of Ordered Sets, and have been employed before for the purpose of classifying chemicals, most notably in [6, 7, 8, 9]. It is also encountered in many other areas of (Social) Decision Theory ([2]).

In order to arrive at the classification function (or decision rule etc.) for a decision problem a useful method consists in applying certain requirements (or desiderata, restrictions, conditions etc.) to the classification function. As subsequent application of requirements define a function increasingly precisely, requirements can first also be used to define a more general frame of the procedure, subsequently to be narrowed down increasingly. Hence, before moving on to the more technical aspects

---

[*] An assessment is a form of classification and as such the two terms will be used synonymously, despite the fact that assessment contains a value dimension, which the term classification does to a lesser degree.

mentioned in the last two paragraphs, a quick overview will first be given in the next section, in a summary and schematic way, on how requirements can be used to define the general frame of a procedure. These requirements to the assessment procedure will also include implications derived from the original requirements.

3.1 Overview

The assessment of chemicals is formalised in a series of basic *requirements*. These, as will be seen later, are partly general (in the sense that they are common in decision making) and partly specific to the assessment of chemicals.

The rationale of this approach consists in fully describing the classification with the help of requirements. Then, if the requirements are shown to be consistent, the same can be said of the classification procedure.

Structure of Classification
The requirements fix the *structure* of the classification procedure. I.e. the requirements define the *class* of functions that are compatible with all requirements made. (Note that it can not be excluded that no such class exists, in which case the classification procedure is inconsistent.) In order to facilitate the discussion of such classes, we introduce the following notation: the class containing all functions compatible with a set R of requirements is denoted C^R. Then, to state that the requirements are not be consistent, can be expressed as $C^R = \emptyset$.

As a first and most basic requirement we first (and for illustration purposes) introduce that of *linearity*: The outcome of the procedure has to be such that all chemicals that are analysed can be compared with each other.[†] The class of functions which transform the original characterisation into such a linear order[‡] is C^L. Assuming that this "linearity" requirement is compatible with the set R of requirements, we have $C^R \subseteq C^L$.

Optimisation
After the requirements have helped to restrict C^L to C^R (i.e. we have only functions that respect all requirements concomitantly) the class of functions still has to be percolated further, as an actual function must be chosen out of the C^R (which is a *set* of functions). Now imagine that the functions in C^R are parametrised.

[†] As will be seen shortly, this will make the outcome of the procedure a linear ordering.

[‡] see below

The actual function then can be characterised through specification of parameter values. This step is referred to as *calibration.* If calibration is performed in view of optimality with respect to a reference,[§] it is referred to as *optimisation.* (The choice to perform calibration through optimality with respect to a reference will be stated more formally in the form of Assumption 2 in Chapter 6.)

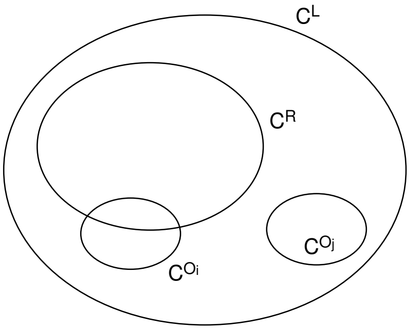

Figure 3.1: Diagram of class C^L of functions that yield linear orderings. Therein contained is the class C^R that is defined by a set R of requirements. One type of calibration (represented by the class C^{O_i}) is compatible with the requirements, the other (C^{O_j}) is not.

Once the class C^R of admissible functions is described, the calibration selects one such function, namely the classification procedure Π. This function (or subclass of functions) both reflects optimality with respect to a reference and — if there are different ways optimisation can be performed — the *type* of optimisation chosen for the classification procedure. It is important to note that there is no apriori guarantee that any type of optimisation should be consistent with the requirements. In other words, it may be that we start to look for an aggregation function while having a type of optimisation in mind that is not compatible with the requirements.

In order to evaluate this important point we define another set of classes, which relate to the type of optimisation used: assuming there are a number of different such types of optimisation, then the ith way an optimisation can be performed defines a class of functions C^{O_i}. Then if C^{O_i} and C^R overlap it follows that the ith optimisation procedure is compatible with the requirements made. On the other hand we can think of a type of optimisation j that gives rise to a class of functions C^{O_j} that, albeit a subset of C^L — i.e. it is a class of optimised linear functions — is not a subset of C^R, i.e. is not consistent with the requirement.

[§] In the present work the reference will be given by a set of reference chemicals.

In Chapter 4 we will see that a commonly used type of optimisation indeed has the property that it is not consistent with the basic requirements used in the present work.

Choice of Reference
With respect to optimisation a few additional comments may be appropriate. Whereas a classification scheme itself can be assessed by means of rationality (or other) requirements, the optimisation of a once given classification scheme cannot be justified by rationality arguments alone. This is primarily due to the fact that the choice of reference, with respect to which the optimisation is performed, is itself not given in any "natural" way. Hence, there is an important element of arbitrariness in the optimisation process. As a consequence and as optimisation is the last step in the definition of the interface between a classification scheme and the world of objects to be classified, it is essential for the question of appropriateness that calibrating is done with some kind of "common sense".[¶]

Calibration
Accordingly, for objects to be classified as "unwanted" we choose reference objects branded as unwanted in the broadest conceivable, international context; for objects to be classified as inconspicuous, we choose examples where general acceptance is stable, e.g. by high economic value for as many stake-holders as possible (for the choice of reference see Assumption 4 in Chapter 7).

The optimisation is defined as the selection of the optimal function with respect to the reference: The optimal calibration is defined as the function which optimally separates the two sets of examplary "ideal" objects through a given classification procedure. (The *criteria* with the help of which the optimality of separation is assessed is crucial and influences the optimisation accordingly.) Hence, the classification must yield at least two classes, one corresponding to "regulated" chemicals, the other corresponding to "non-regulated" chemicals. These classes are coined "red" and "green", respectively.

Since the classification procedure should mirror the reference sets, the elements of the

---

[¶] This is easily seen in cases where optimisation is done by experts (which experts have been selected, what is the precise nature of their "best available knowledge" etc.). It is less visible in cases where optimisation is done by comparing the outcome of a classification to some sets of ideal objects or examples. These are selected by the stakeholders of the classification process. In the context of political decisions where stakeholders are of the most various kinds, it is wise to select ideal examples according to up-to-date political practice reflecting the relevant setting of values.

reference sets should be optimally placed in the respective sets "green" and "red". It may be that such a separation is possible. Most likely, however, this is not the case. (This outcome, too, can be influenced by the choice of the criteria mentioned above.) Rather, some chemicals from the reference set of regulated chemicals cannot be placed in "red", whereas some of the non-regulated reference chemicals will not end up as "green" in the classification performed by the optimised function. For such chemicals that cannot be placed in the desired class via optimisation a third class is defined: "orange".

Observation 2 *In order for allow for chemicals from the reference sets to be misplaced, one class is introduced in the ordering of the assessment procedure, in addition to the two classes corresponding to the classes in the reference set.*

3.2 Importance of Formal Approach

It may be asked why it is of importance at all to set up requirements for the classification funtion. As we have already anticipated there can be situations where an optimisation technique can be incompatible with a classification rationale. The proof of such an inconsistency may not be straightforward and hence may require a formal treatment of the problem. Similarly, a formal analysis of the requirements can help to see whether a classification procedure *itself* is consistent, i.e. whether the set C^R actually contains any functions, irrespective of the optimisation method.

In order to better illustrate this last point, we now look at a particular classification procedure *that is inconsistent*, although it may not appear so at first glance.

As an example we take six chemicals $a - f$ which have to be evaluated by a decision maker who considers two different criteria. Although nothing more is known, the decision maker is required to be "unbiased", i.e. each of the two criteria at his disposal to classify the chemicals should be of equal importance to him.
The preference situation of the decision maker is depicted in Figure 3.2. The diagram can be read in the following way: the chemical a is preferred over the chemicals c, d, e, while b is preferred over f. The chemicals a and b are not comparable to each other (i.e. the decision maker is not in a position to make a judgment of preference over a and b). Also d and c are also preferred over e but d is preferred over f, whereas c is not.

Now suppose the following simple way is used to establish a preference among the

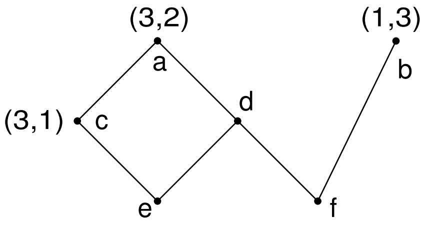

Figure 3.2: Graphical depiction of a decision problem for a decision maker. Each element vertically higher than another to which it is connected by a line is preferred.

elements (from [9]): Whenever an element is preferred to some other element and no element is preferred to it,[‖] then it belongs to an equivalence class of "preferred elements". In the example in Figure 3.2 the "preferred elements" contains the elements a and b.

As an additional information we learn that the decision maker has been asked to use the following scale for each of the two criterion: highly preferred (3), medium preferred (2) and least preferred (1). Hence the decision maker was asked to describe each chemical with the help of tuple ($(1,1)$, $(1,2)$ etc.).

In fact we now also learn that the description of the decision maker for the two chemical c and b was: he put c in the best category for one criterion, but in the worst category for the second criterion (i.e. for c he has the tuple $(3,1)$). For chemical b the situation was the exact opposite: it ranked it worst for the first criterion and best for the second criterion: for b he has the tuple $(1,3)$. (For the chemical a he has the tuple $(3,2)$.)

Now remember that the decision maker was asked to regard the two criteria of equal importance to him. As a consequence, the decision maker *must* be indifferent between the chemicals b and c, since b just has the inverted tuple of c. On the other hand, the class "preferred elements" contains only a and b. But if the decision maker is indifferent between b and c then this excludes the possibility that he prefers b over c, which in turn implies that c is in the class "preferred elements", which it is not. Hence the definition of the "preferred class" contradicts the demand that the decision maker be unbiased.

The example illustrates that requirements can be made which under close scrutiny may turn out to invalidate a methodology for decision making (the definition of

---

[‖] mathematically the element is said to be *maximal*, see Section 3.3.3

"preferred elements"). It also illustrates, conversely, that the choice of methodology may imply an unwanted characteristic (such as that a decision maker can be biased with respect to the criteria).

3.3 Order Theoretical Setting

In order to formalise the procedure we start by putting the procedure in an order theoretical setting. Later a number of tools are introduced that will be subsequently used for visualisation purposes and as operators[**]. Most of the following definitions can be found in classical text books on partial order or lattices [80, 81].

The classification procedure evaluates a set X of chemicals, the *ground-set.* At the beginning of the classification stands the *initial characterisation* of the chemicals: The elements of the set X are characterised by *evaluative properties* e_i, which form the *set of evaluative properties* $E := \{e_1, e_2, \ldots\}$.[††]

An *order relation* P (also called *ordering*) puts the chemicals in pair-wise relations. The order relation P is typically an expression like "is smaller than or equal to" or "is a subset of" and designated with the commonly used mathematical symbol ($P := \leq$) and ($P := \subseteq$), respectively. An order relation (P, $\leq$, $\subseteq$ etc.) can also be *specified as a set* of tuples of elements from X, i.e. as a subset[‡‡] of the Cartesian product $X \times X$.

In the present case the result of the assessment must be a linear ordering as was pointed out in Section 3.1. This result puts all chemicals in a ranking and hence all chemicals will be in a pair-wise relation in that ordering. As such, the result of the classification of chemicals can be represented as an order relation, too[*]. Note that a linear order relation is denoted by L.

The order relation P is a subset of $X \times X$ and nothing more has been said about

---

[**] to perform operations on sets and ordering relations

[††] Chapter 6 will include the crucial results that in the present case of precautionary classification of chemicals, the evaluative properties are given by filters derived from threat scenarios.

[‡‡] In turn, set P of such tuples can define a symbol ($\leq := P$, $\subset := P$). In general P has to be the subset of the set which contains *all possible pairs* of chemicals. This set can be formed by forming the Cartesian product of the set X of chemicals with itself: $P \subseteq X \times X$. Thus any order relation, including the commonly used ($<, \leq, \subset, \subseteq$) can always be defined as subsets of the Cartesian product of the set of elements X with itself.

[*] that order relation expresses a preference of chemical over another with respect to some property.

it up to this point. In order to specify the order relation we wish to use in some context, we can make use of basic definitions which are found in Appendix A. A number of order relations are commonly found in decision theory that are based on these basic definitions.

The choice of ordering can be the source of considerable controversy and confusion, not least because orderings have different designations depending on the author [2, 86, 87]. Presently we may state that certainly the axiom of *transitivity* is fundamentally rational, even though this axiom, too, can be weakened [2]. The choice of using a non-strict (i.e. anti-symmetric) rather than a strict order relation, on the other hand is more debatable and can be regarded as being based on grounds of the author's subjective preference of non-strict relations for the present problem. Since P reflects the initial characterisation, it need not be complete, i.e. it is possible for two chemicals not to be comparable in the order relation.

Definition 2 *The order relation P which reflects the initial characterisation of the chemicals is a partial ordering. I.e. it is a transitive, reflexive and anti-symmetric order relation.*

We now know the basic properties that the ordering relation must have, regardless of how it is defined. We can hence establish the ordering relation in *words* and check if the basic requirements are fulfilled.

Definition 3 *The order relation $P = \preceq$ used in the characterisation is: "has — for all properties $e_1, e_2, \dots$ taken into consideration — smaller values than".*

i.e. if a chemical x has lower (or equal) values *for all evaluative properties e_i* used in the characterisation (e.g. for all indicator used in the assessment of chemicals) than a chemical y (in the sense of the common magnitude relation $\leq$), then x is smaller than (or equal to) y:

$$x \preceq y \iff \quad e_i(x) \leq e_i(y) \;\; \forall e_i \in E. \tag{3.1}$$

It is easily verified that the order relation $\preceq$ is reflexive, anti-symmetric and transitive and thus fulfills the definition of a partial order as required by Definition 2.

Note that if relation 3.1 holds for two chemicals, then the same tuple of chemicals is found in the *linear orders given individually by the evaluative properties* (L_i's).

Since these linear orders also are sets of tuples, the partial order relation can be given alternatively as the intersection of the linear orders given by the evaluative properties:

$$P = \cap_i L_i. \tag{3.2}$$

3.3.1 The Poset

An order relation P together with the set of chemicals X forms a partially ordered set $\mathbf{P} := \langle X, P \rangle$ (or *poset*).[†] To the extent that the poset contains all relationships which are given by the data, the poset of the chemicals can be viewed as the *given* description of the set of chemicals. An example thereof is given by the set of eight chemicals given in Table 3.1.

chemical	e_1: molecular weight (g/mol)	e_2: boiling point (° C, 760mm Hg)
$\alpha1$:= acetone	58.08	56
$\alpha2$:= acetonitrile	41.05	82
β:=benzene	78.1	80
ι:=iodopentafluoroethane	245.9	12
μ:=methylene chloride	84.93	40
$\tau1$:= 1, 1, 2, 2-tetrafluoroethane	102.0	-27
$\tau2$:= tetrahydrofuran	72.11	65
ω:= H_2O	18.02	100

Table 3.1: Data for *molecular weight* and *boiling point* for the set X of chemicals.

The example in Table 3.1 is a set X of 8 chemicals which are characterised by a set E of 2 evaluative properties (the parameters molecular weight and boiling point). The choice of chemicals (commercially available solvents) as well as the choice of the two evaluative properties serve solely the purpose of illustrating a partially ordered set which results from the description of a set of chemicals with the help of two properties.

The elements of the set X can be depicted in a Cartesian plane, where each dimension represents an evaluative property. The relation $\preceq$ now has a ready interpretation in the plane: if a chemical b is to the upper right-side of a chemical a, then $a \preceq b$.

[†] Often a partially ordered set is referred to simply as an *ordered set*.

$$\begin{aligned}
&X = \{\alpha 1, \alpha 2, \beta, \iota, \tau 1, \tau 2, \mu, \omega\}\\
&E = \{e_1, e_2\}\\
&P = \{(\alpha 1, \tau 2), (\tau 2, \beta), (\alpha 1, \beta), (\tau 1, \iota),\\
&\qquad (\alpha 1, \alpha 1), (\alpha 2, \alpha 2), (\mu, \mu), (\omega, \omega),\\
&\qquad (\iota, \iota), (\tau 1, \tau 1), (\tau 2, \tau 2), (\beta, \beta)\}
\end{aligned}$$

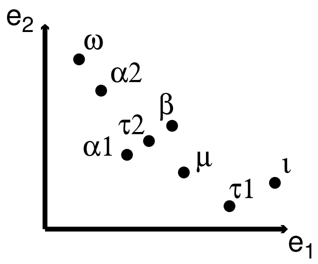

Figure 3.3: *Left*: The set X of chemicals, the set E of evaluative properties and P the order relation containing all tuples related via P. *Right*: Chemicals X displayed in the Cartesian plane spanned by *molecular weight* (e_1) and *boiling point* (e_2)

.

3.3.2 Hasse Diagrams

But whereas the drawing of a poset in a Cartesian plane is easily done for the 1- and 2-dimensional case – by simply placing the elements of the ground-set in the plane – it becomes challenging for three and impossible for higher dimensions. We hence now introduce a method that permits to visualise posets of any dimension. The method is commonly used and described in some of the textbooks already mentioned [80], but is also discussed more specifically in a number of papers in the context of the classification of chemicals [9].

When depicting the poset $\mathbf{P} = \langle X, \preceq \rangle$ more generally, the corresponding graph must represent the tuples of comparable elements in $\preceq$ by connecting lines: all relations which are elements in the order relation P should be visible in the graph. Since the order relation $\preceq$ itself is transitive, some relations arise solely from transitivity. Such relations need not be shown in the graph, since they result obviously from other relations. Similarly, it is not necessary to represent every relation arising from reflexivity. Hence a relation stripped from the tuples which result from transitivity and reflexivity (*covering relation* $\preceq^c$) is sufficient for the purpose of drawing a graph.

$$x \preceq^c y \text{ if } x \preceq z \prec y \Longrightarrow x = z \tag{3.3}$$

The graph depicting the covering relation is called the *Hasse diagram*. Graphically, the diagram is obtained by

1. associating with each chemical from the set X a point in the Euclidean plane

2. drawing a line segment between any two chemicals, for which the relation $\preceq^c$ holds and
3. ensuring that this is done in a way that for $x \preceq^c y$ the point representing x lies lower than the point representing y
4. avoiding that a point z is on a line connecting x and y ($x \preceq^c y$), except if $x \preceq^c z$ and $z \preceq^c y$

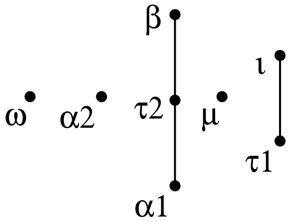

Figure 3.4: Representation of the poset $\mathbf{P} = \langle X, P\rangle$ as a Hasse diagram. Elements ω, $\alpha 2$ and μ are isolated; $\alpha 1 \preceq^c \tau 2 \preceq^c \beta$, $\tau 1 \preceq^c \iota$.

In the Hasse diagram of the poset $\mathbf{P}$ in Figure 3.4 the chemicals $\alpha 1$ and $\tau 2$, and the chemicals $\tau 2$ and β are connected with a line segment, since the covering relation holds for these tuples, whereas the chemicals $\alpha 1$ and β are not connected, since the relation arises not from the covering relation but from transitivity. Note also that the orientation of the diagram in the y-plane indicates $\alpha 1 \preceq^c \tau 2 \preceq^c \beta$ and $\tau 1 \preceq^c \iota$, which both are correct. The vertical position of these chains of comparable chemicals relative to each other as well as the vertical position of the isolated chemicals are irrelevant.

3.3.3 Useful Operators

Composition
The *composition* of two binary order relations R and S on X is defined as

$$xRSy \iff xRz \text{ and } zSy \text{ for some } z \in X \tag{3.4}$$

Maximum, Minimum, Maximal, Minimal
En element of a set S is *maximal* if there is no other element larger in the set S. The element is the *maximum* if it is not only maximal, but also is larger than any other element in S. *Minimal* and *minimum* are defined analogously.

Upsets and Downsets
All chemicals that are larger than the chemical x (in the sense of $\preceq$) form the upset of x, $\uparrow x$. Conversely, all elements that are smaller than x form the downset $\downarrow x$

$$\uparrow x := \{z : x \preceq z\}, \;\; \downarrow x := \{z : z \preceq x\} \tag{3.5}$$

In Figure 3.4 we have $\uparrow \alpha 1 = \{\alpha 1, \tau 2, \beta\}$. Upsets and downsets can be defined also for sets, e.g. $\uparrow \{x, y\}$ contains all elements larger than both x and y: in Figure 3.4 we have $\uparrow \{\alpha_1, \tau_2\} = \{\tau_2, \beta\}$.

Supremum and Infimum
In case the upset of an element x ($\uparrow x$) has a minimum, then this minimum is called the *supremum* of x (symbol: $\wedge x$). In case the downset ($\downarrow x$) has a maximum, it is called the $infimum$ (symbol: $\vee x$). Supremum and infimum can also be defined for subsets instead of elements.

3.3.4 Special Partial Orderings

With respect to order relation that are derived from other relations (i.e. for which $R \subseteq P$), we adopt the following notation: If a relation is derived from another relation, then this provenance is recorded with the according suffix. Hence a relation R that is defined with respect to a relation P is written R_P. Often there is no ambiguity and the suffix is dropped.

An equivalence relation is defined as a relation which is *symmetric*, *reflexive* and *transitive*. For a given relation P distinct relations can possess these properties.

If two chemicals have the exact same characterisation in terms of numerical or other values produced by the evaluative properties then the *identity relation* $\simeq$ ($\simeq_{\preceq}$) holds:

$$x \simeq y \iff x \preceq y \text{ and } y \preceq x \tag{3.6}$$

This relation obviously is an equivalence relation. Another important relation which is defined for $\preceq$ is the *incomparability* relation $\parallel$ ($\parallel_{\preceq}$).

$$x \parallel y \iff \text{not}(x \preceq y) \text{ and } \text{not}(y \preceq x) \tag{3.7}$$

It can be easily shown that although $\parallel$ is reflexive, the relation is not an equivalence relation. Take the two chemicals $\alpha 1$ and $\tau 2$ in the example for which $\alpha 1 \preceq \tau 2$ and third one, $\alpha 2$, that is isolated, i.e. is comparable to no element except itself. Then

$$\alpha 1 \parallel \alpha 2 \parallel \tau 2. \tag{3.8}$$

If $\parallel$ was an equivalence relation then from transitivity, it would follow that $\alpha 1 \parallel \tau 2$, which is clearly not the case.

As a foreclosing comment for Chapter 5 note that the symbol $\sim$ is often found to designate the *symmetric complement* of an ordering (written $sc(\preceq)$, hence $\sim := sc(\preceq))$. For the non-strict relation $\preceq$ the symmetric complement is exactly equal to the incomparability relation $\parallel$.[‡]

$$\parallel = \sim \tag{3.9}$$

This equality is of importance for linearisation function, introduced in Chapter 5, where a relation is used which is derived from $\sim$.[§]

As we have just seen, the relation $\parallel$ is not an equivalence relations. We now introduce an ordering relation that holds between incomparable elements, but is also an equivalence relation (i.e. beyond being *symmetric* and $reflexive$ the relation is also $transitive$). The relation looks at *interchangeability*: if the tuples representing two chemicals in a graph can be interchanged without changing the graph itself then the relation $\approx$ holds between the chemicals

$$x \approx y \text{ if } \{z : z \parallel x \text{ or } z \simeq x\} = \{z : z \parallel y \text{ or } z \simeq y\}. \tag{3.10}$$

In other words if $\approx$ holds then the upsets and downsets are identical for the elements x, y. In turn, if upsets and downsets are identical for two chemicals, then they are interchangeable. As is easily verified, the relation $\approx$ is *symmetric*, $reflexive$ and $transitive$ and hence is an equivalence relation.

3.3.5 Partitions and Extensions

Partitions

[‡] The complement of $\preceq$ is $c(\preceq) = \succ \cup \parallel$. The relation $\succ$ is not symmetric, since it is strict. Hence the symmetric part of the complement of $\preceq$ is $s(\succ \cup \parallel) = \parallel$.

[§] Note that the relation 3.9 only holds for the order relation $\preceq$ as it is non-strict. In the case of a strict partial order this relation is $\sim := sc(\prec) = s(\succeq \cup \parallel) = \simeq \cup \parallel$, and thus not only contains incomparable elements but also identical ones, in the sense of $\simeq$.

For any equivalence relation E, X/E is the set of all equivalence classes as defined by E, and X/E is a partition of X. Hence $X/\simeq$ and $X/\approx$ both are partitions of X, but $X/\parallel$ is not.

Partitions of sets are of interest in the present context as the elements of the partitions (i.e. the classes) are linearly ordered (see Theorem 1).

Extensions
A subset of an order relation P is itself an order relation. Conversely, if an order relation P is included in a set Q then Q is called an *extension*. If P is included in a relation L that is a *linear* order relation, then L called a *linear extension*.

Extensions are important in decision theory as they are *order homomorphims* (or *mononone maps*, s. below) of an order relation P. An extension always contains all tuples of comparable elements in the characterisation given by the initial data. Hence no information from the characterisation is "lost" during the mapping of a poset, if the map is an extension. Hence, decision problems are typically confined to finding (linear) extensions of poset, since decision problems should be based on the information in the initial characterisation (see Requirement 1 in Chapter 4).

3.4 From Posets to Linearly Ordered Sets

The classification of chemicals is a decision problem, and the goal of the classification is to obtain classes of chemicals which are ordered in a linear manner. This *linear ordering* L expresses preference of one class over another and of that class over a third class, and so on. Class 1 may contain the most precarious substances, class 2 chemicals of medium precariousness and 3 innocuous chemicals. If such preferences cannot be obtained, then it is difficult to attach meaning to the classification.

In the decision problem of classifying chemicals into classes of regulatory status, suppose two such classes are not comparable to each other. Then no preference can be stated about these two classes and hence the problem ceases to be one of choice.[¶] As a consequence, if the classification is to yield a linear order than *all* chemicals

---

[¶] There is one possible exception [2]: Suppose only a top candidate has to be chosen. Then such a chemical (or an equivalence class of chemicals) can be chosen even if some elements are incomparable: as long as a maximum element exists in the poset it can be chosen, regardless of whether elements smaller than the maximum element are incomparable among each other. That, however, is also a linear order: all elements except for the maximum element are taken together to form a second class, among which the decision maker is indifferent.

have to be made comparable. That means that whatever the characterisation of the chemicals at the beginning of the procedure, all the incomparable chemicals — in the poset **P** that summarises the characterisation — will have to be made comparable. (The linear ordering L is derived from the partial ordering $P =\preceq$ and hence is correctly written as $L_\preceq$ (or L_P).)

Definition 4 *The order relation ($L_\preceq$) reflecting the final classification of the chemicals should be a linear ordering, i.e. a complete partial ordering (transitive, reflexive, and complete).*

(Note that the relation $L_\preceq$ must certainly be an equivalence relation (without proof)). We can now restate the problem of finding a classification of a set of chemicals as the *mapping* (function) Π of a partially ordered set $\mathbf{P} = \langle X, \preceq \rangle$ onto a complete partially ordered set. But if the partial ordering is complete, then it is a linear ordering, and as a result, the complete partially ordered set is a *linearly ordered set* (*loset*) $\mathbf{L} = \langle X, L_\preceq \rangle$. This leads to a more formal expression for the classification function, or decision rule, which we introduced at the beginning of this chapter:

$$\mathbf{P} \xrightarrow{\Pi} \mathbf{L} = \langle X, L_\preceq \rangle. \tag{3.11}$$

The mapping function Π of a poset to a linear ordered set is often called an *aggregation* function. Thereby is meant the aggregation of a series of numerical values for different evaluative properties into a single number, e.g. by forming the sum of such values, or the product or some other appropriate operator.

When viewed against the background of the assessment of chemicals, then the choice of function Π, together with the choice of evaluative properties $e_1, e_2, \ldots$ *amounts to defining the assessment procedure.* The characterisation (based on the properties) can then be represented mathematically by a partially ordered set (*poset*) and the final ranking by a linearly ordered set (*loset*), which results from applying the assessment function to the characterisation. The loset consists of three classes, as stated in Observation 2.

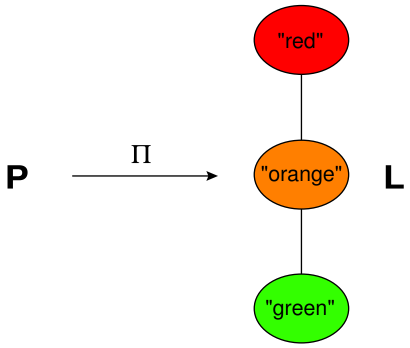

Figure 3.5: Eqn. 3.11 illustrated: the assessment procedure consists of a mapping of a poset **P** to linear ordering **L** by a decision rule Π. The final linear ordering is required to consist of three equivalence classes: "green", "orange" and "red".

Chapter 4

Coarse-graining and Requirements

The mapping function Π has been identified as the aggregation function that yields the linear ordering $L_{\preceq}$ that is required as the result of the assessment. This chapter discusses some general problems related to the function Π, and introduces a number of *requirements* which are made with respect to that function. In addition, some other requirements that are frequently used in decision theory are mentioned at the end, but not discussed at great length. Nevertheless consistency of requirements including the latter is briefly discussed. Besides showing that the requirements made are mutually consistent, a number of *implications* are drawn, which are of relevance for the present problem of finding a suitable assessment procedure. We also present a simple (graphical) method for finding a classification function that respects all requirements made herein. This method relies on the *event-space* of the assessment problem. In Chapter 5, the analysis follows on how the requirements and their implications restrict the possible choice of Π.

Chapter 3 established the poset as the structure that represents the characterisation of a set of chemicals and as such is the domain of the classification funtion. In order to understand the functioning of a classification function it is helpful to analyse its domain and to study if certain domains can be singled out that are easier to handle than others. If such is the case, then the possibility is given to first transform an original domain to a domain that is easier to subject to the classification functions.

4.1 Elaborateness vs. Comparability

If there is only one evaluative property then the poset $\mathbf{P} = \mathbf{L}$ is linear from the very beginning and we need not concern ourselves with the problem of finding a suitable aggregation function Π. The more elaborate the description, however, the more chemicals become incomparable, since it becomes increasingly unlikely that all evaluative properties yield the same relation. Hence the more elaborate the description, the more such pairs of chemicals have to be made comparable. There is therefore a dilemma at the start of the classification of chemicals (and more generally of any decision problem): the dilemma of choosing between detailedness of description (i.e. the number of evaluative properties) and the relative number of incomparable pairs in the poset.

4.2 Simplified Posets

The number of pairs of incomparable elements is, however, also dependent on the number and on the *definition* of the evaluative properties, and manipulation of the properties can lead to simplified posets. As a general rule, the more detailed and exact the characterisation by the evaluative properties, the more pairs of incomparable elements emerge. On the other hand, a loss in detail (less evaluative properties) and/or resolution can lead to significantly simplified posets. The question then arises to what extent the poset can be simplified by manipulating the evaluative properties before the characterisation reflected by the poset becomes meaningless through over-simplification.

One way to simplify the poset, while not sacrificing too much accuracy (and not sacrificing an evaluative property), is to replace continuous values with discrete value ranges. A justification for this approach can be given by supposing that the decision maker is only interested in approximate values for an evaluative property (because it is all she needs for a particular purpose). Using the example of solvents introduced earlier, such a simplification may be obtained by *rounding* of the values for boiling point to the resolution of five degrees may be acceptable to the decision maker. Now suppose the decision maker requires that the solvent should have as low as possible a molecular weight and as low as possible a boiling point. The two solvents *acetonitrile* and *benzene* from the example in Figure 3.1 are then incomparable in **P**. But if the boiling point was rounded to the resolution of five degrees, then the data would read as shown in figure Table 4.2. In this modified characterisation, we obtain *acetonitrile* $\preceq$ *benzene* and hence the two elements are now comparable, whereas in the original poset the two solvents were incomparable.

In many cases high precision is not necessarily required and indeed — while boiling

	bp	bp rounded	MW
acetonitrile	82° C	80° C	41.05 g/mol
benzene	80° C	80° C	78.1 g/mol

point is in most cases a well-defined and well-measurable parameter — in many cases measurements are inexact. Then too strict a definition of the evaluative properties may introduce an incomparability which in reality is not really existent: e.g. if the difference of two measurements was within a interval negligible for the purpose of the classification then the original partial order should be viewed with caution.

In short, the characterisation of the chemicals may in fact be either *too imprecise* (due to measurement error) or indeed *too precise* in the context of a rough classification. In both cases continuous values can be replaced with discrete numbers representing intervals inside of which lie the continuous values for the property.

4.2.1 Coarse-graining

Besides rounding, another possibility consists of setting a fixed number of intervals over a continuous interval. This process is presently referred to as *coarse-graining*. The coarse-graining is demonstrated on the example of the solvents by substituting each of the two evaluative properties *boiling point* and *molecular weight* with the 3 classes (1), (2) and (3).

The simplified poset resulting from the coarse-graining Δ is depicted as a Hasse diagram in the Cartesian plane as depicted in Figure 4.1.[*]

Inspection of Figure 4.1 reveals that many more chemicals are now comparable. Indeed, the coarse-grained poset $\mathbf{P}'$ now no longer contains isolated elements. Additionally, the ordered elements have been reduced to 4 equivalence classes, inside of which lie all the chemicals. Note that the two ordered pairs $(\alpha 1, \tau 2)$ and $(\tau 1, \iota)$ now belong to the same equivalence classes.

If coarse-graining is used to simplify the poset, then the classification is no longer performed by a single linearisation function. Rather, the original poset $\mathbf{P}$ is first mapped by a coarse-graining function Δ to the coarse-grained poset $\mathbf{P}'$ and only then linearised by a linearisation function Λ to $\mathbf{L}$:

---

[*] It should not go unnoticed that the position of the grid on the plane influences the resulting coarse-grained poset. The current placing is arbitrary, and one of the key problems of simplifying a poset will be later shown to consist in the placing of the grid.

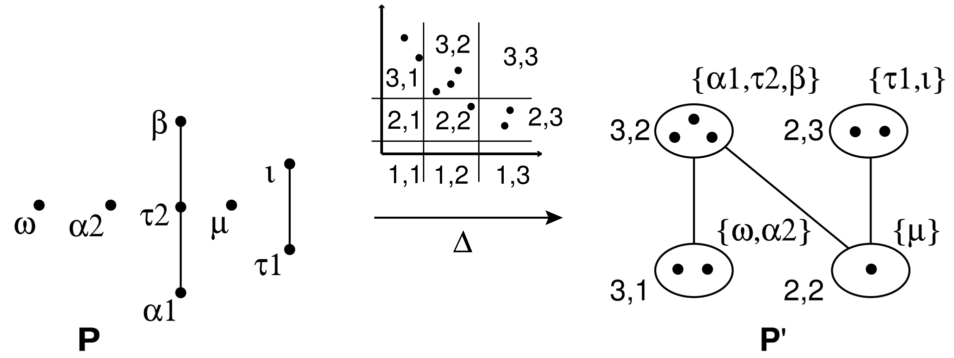

Figure 4.1: Mapping of poset $\mathbf{P}$ to poset $\mathbf{P}'$ with the coarse-graining function Δ. The function places a 3×3-grid over the Cartesian plane shown in Figure 3.3.1. There result nine classes, of which four are populated: $(3,1)$, $(3,2)$, $(2,3)$ and $(2,2)$.

$$\mathbf{P} \overset{\Delta}{\to} \mathbf{P}' \overset{\Lambda}{\to} \mathbf{L} \tag{4.1}$$

Since the goal of the present work consists of establishing a precautionary prescreening, it is assumed that some sort of coarse-graining will be appropriate. Hence all further considerations will be based on discrete characterisations. This will lead to a significant simplification of the poset to such an extent, that graphical representation of the poset within the space of all possible *events* will be possible.[†] Note that the assessment function Π now results from concatenation of two functions, the coarse-graining function Δ and the linearisation function Λ: $\Pi = \Lambda \circ \Delta$.

We retain the notion that posets can be simplified (in the sense that more chemicals are made comparable) by *coarse-graining* of evaluative properties that define it. The order relation of such a simplified poset is a subset of the order relation of the former poset and hence coarse-graining is an order homomorphism and the order relation of the simplified poset is an extension of the original one.

It may be appropriate at this point to specify the notation adopted for tuples: unordered tuples of discrete values are generally written as $\{m, n\}$. The symbol is also used to describe a class that contains all possible ordered tuples of m and n (in this case two such tuples). For ordered tuples, on the other hand we presently adopt the notation (m, n), and sometimes also the abbreviated mn.[‡]

[†] Note that the dissection of the classification function into two functions, one transforming the poset, the other linearising it, has been discussed, if slightly differently by Peterson in [75]. In the referenced work a distinction is made between *effecive* and *transformative* decision rules.

[‡] In logic, the notation $< m, n >$ is often found for ordered tuples. *Ordered* simply means that

4.2.2 Event-spaces

An *event* is defined as a possible combination of values for each of evaluative properties e_i (cf. beginning of Section 3.3). In the general case, these values are cardinal, but they can also be ordinal such as was the case for the values encountered in the coarse-graining in Section 4.2.1, where the ordinal descriptors $1, 2, 3$ were used. If two evaluative properties are considered then an event is a tuple (k, l) where k and l are values (i.e. discrete values). All such events for a given characterisation form a set. The set contains all possible combinations of k and l and hence is the Cartesian product $k \times l$. In coarse-grained posets the number of possible *events* — again for the case of two evaluative properties — is finite and so is the set containing them. The set can be ordered and thereby gives rise to an ordered set:

$$\mathbf{O} = (\{k \times l\}, O), \quad O \subset \{k \times l\} \times \{k \times l\}. \tag{4.2}$$

The poset $\mathbf{O}$ consists of the ground-set $\{k \times l\}$ (which in turn consists of all possible combinations of the discrete values) and is equipped with the order relation O which is a proper subset of $\{k \times l\} \times \{k \times l\}$ (i.e. the set of all possible pair-wise combinations of the tuples of discrete values.)[§]. The order relation O can be defined analogously to $P = \preceq$ in Eqn. 3.1, with the extra characteristic that since *events* rather than *chemicals* are ordered, the order relation can be confined to the strict case: $O = \prec$.

In allusion to Ω in probability theory — which contains all events to which a probability can be attached — we choose to term the poset $\mathbf{O}$ the *event-space*[¶]. Like any poset, the event-space can be depicted as a Hasse diagram. It is not very difficult to imagine that in most cases the numbers of events (say 100) and the connections between them are such that a graphic representation is of little use. But if we take

---

an order relation holds between m and n. Since this includes asymmetric relations, the order of the two elements matter.

[§] In analogy to Eqn. 3.2, the order relation O can be viewed as the intersection of the linear orders representing the evaluative properties. To see how note first that all events are occupied, in the two-dimensional space as well as in the one-dimensional case of each of the evaluative properties. Now take the two tuples $(1, 2)$, $(1, 3)$: a chemical x which occupies $1, 2$ is smaller or equal to another chemical y occupying $1, 3$ (since $1 \leq 1$ and $2 \leq 3 \Rightarrow x \preceq y$.). As a result we find the tuple (x, y) of chemicals in the linear ordering corresponding to the two properties, and hence if we intersect the two also in the partial order, i.e. in O

[¶] In Kolmogoroff's classical formalism of probability theory the event space Ω is given as the set of which the elements are the elementary events and on which all subsets forming σ-algebras are defined [82].

the case of one evaluative property, then the depiction is straightforward and clear. The same holds for two evaluative properties, where we obtain a grid that can be depicted in a two-dimensional plane. In the case of three dimensions the situation is more challenging. But even in that case, the poset can be depicted, as long as the number of intervals is kept low. (For higher dimensions, however, the graphic of a poset becomes so cumbersome as to make is useless.)

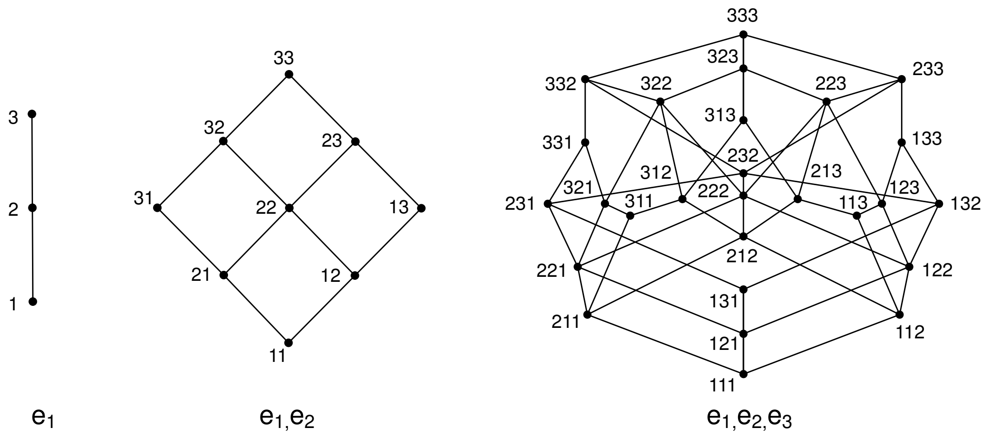

Figure 4.2: Hasse diagrams of event spaces. *Left*: one evaluative property; event-space **3**. *Middle*: two evaluative properties; event-space **3** × **3**. *Right*: three evaluative properties; event-space **3** × **3** × **3**

Hence, if we keep the number of intervals as well as the number of evaluative properties low, we have a tool at hand to study all possible events and the relations among such events concomitantly.

4.3 Event-space as a Tool

In the introduction of Chapter 3 we have defined the set C^R as the class containing all functions that are compatible with a set of requirements R. We also saw that if C^R contains at least *one* function then the requirements can co-exist and hence are consistent. Thus, in order to first show that a set R is consistent and second to find an appropriate function for the classification procedure, functions can be analysed with respect to their compliance with the set R.

4.3.1 Event-space as Test Domain

Now suppose all linear orderings resulting from linearisation by some particular function f were known for all possible initial characterisations. Then even if we didn't know the function itself we could still check if each linear ordering was in compliance with the set R. And, if indeed all such orderings were in compliance with R, then we would be able to deduct that the function f is compatible with all requirements in R (and hence that the set R is consistent). Hence exact knowledge of f appears not to be required for the proof of consistency of R. The event-space is by definition composed of all possible events that can occur for some set of evaluative properties, i.e. every possible characterisation — given a number of evaluative properties — is included in the event-space. Put more formally, every possible characterisation P is a subset of the event-space O.

Conjecture 1 *If the linear ordering of a set U respects a set R of requirements, then any subset of that linear ordering of U also respects the requirements R.*

Now suppose Conjecture 1 holds, then, as a consequence, if the linear ordering of the event-space reflects all requirements in R, it follows that all linear orderings of any characterisation that is a subset of O — resulting from the same function that linearises the event-space in a manner respecting the requirements — also respect the requirements in R.

Consequence 1 *In order to prove that a set R of requirements is consistent it suffices to show that a single linear ordering of the corresponding event-space respects all requirements in R.*

If, furthermore, the event-space is such that a graphical representation is possible in a Hasse diagram then the proof of consistency of the requirements in R can be performed graphically.

4.3.2 Other Geometrical Representations

The use of event-spaces as geometrical tools is somewhat similar to another method used for the discussion of rational choice, and more specifically voting problems: the *representation triangles* introduced by Saari [83, 84].[‖] The method depicts all

---

[‖] The two books are essentially equivalent. For a very summary introduction see also [85].

possible rankings of alternatives by voters[**], and allows to discuss all procedures that are commonly used in elections. The tool is very powerful in demonstrating the weaknesses and strengths of the various election procedures. As such the method allows to demonstrate inconsistencies in a simple and elegant manner.[††]

These demonstrations are, however, conducted on a simple system of three alternatives, for which all rankings are represented in the triangle. In the present case, it appears that we deal with potentially many more such alternatives,[‡‡] and for such higher dimensional cases, an analysis along the method pursued by Saari appears to be much more involved (see the introduction and concluding notes in [83]).

4.4 Formal Constraints on Aggregation Functions

After establishing the necessary formal order theoretical setting, its tools and the evaluative properties therein appearing, we can now concentrate on the proper definition of the classification procedure Π that leads to the required linear order $\mathbf{L}$.

To start with, we can think of many functions of very different types satisfying Eqn. 3.11: all functions forming the set C^L. These functions, albeit all leading to linear orderings may have very different properties, though, and some may be better suited for decision problems than others. Moreover, a *specific* aggregation problem such as the precautionary pre-screening of chemicals may yet have special characteristics which an aggregation procedure in general would not necessarily be asked to require. In the parlance of the diagram in Figure 3.1 C^L must be narrowed down to C^R, thereby reflecting all such requirements as are made.

The requirements made to the function — and hence to the classification problem — can be divided into two categories: the *general* requirements (which can also be called "rationality" requirements) and *special* requirements (which are specifically defined for and meaningful only in the context of the present classification of chemicals).

As mentioned earlier, the class of functions C^R may well be empty, in which case the requirements made for the classification procedure are inconsistent.

[**] to which correspond evaluative properties in the present work

[††] The proof of Arrow's theorem can be conducted by the method in a geometrical manner.

[‡‡] Chemicals are the alternatives which are ordered by each of the "voters".

A reader familiar with Saar's work will note that the event tuples used in the representation in the event-space contain the coordinates of *all* hypothetical alternatives inside the triangle.

4.4.1 General Requirements for Aggregation

Some requirements are elementary in the sense that if these requirements are violated then this leads to logical implications which are generally regarded to be false or at least irrational [2, 90].

Requirement 1 (Order-preservation) *The function* Π *representing the classification procedure is an order homomorphism*

$$\forall a, b \in F, a \preceq b \Rightarrow \Pi(a) \preceq \Pi(b) \tag{4.3}$$

This implies that any function that is a component of Π *is also order-preserving.*

This important requirement states that the final linear ordering L must be a linear extension of P: if in a poset $\mathbf{P}$ two chemicals are comparable in the sense of the order relation (e.g. $x \preceq y$), then this relation should remain valid in the linear order $\mathbf{L}$. Since the order relation can be seen as a set of ordered tuples, order preservation requires that all elements of the order relation P are also elements of the order relation L:

$$\Pi \text{ is order-preserving} \iff P \subseteq L. \tag{4.4}$$

Note that if the poset $\mathbf{P}$ is first mapped to an intermediate poset $\mathbf{Q}$ which then in turn is mapped to a linear order $\mathbf{L}$, then order-preservation requires $P \subseteq Q \subseteq L$.

Requirement 2 (Symmetry) *The classification procedure* Π *is symmetric*[*] *if*

$$\forall x, y \in X, \Pi(e_1(x), \ldots; e_1(y), \ldots) = \Pi(e_{\Psi(1)}(x), \ldots; e_{\Psi(1)}(y), \ldots) \tag{4.5}$$

where $\Psi(i)$ *is a permutation function on the index set* $\{1, 2, \ldots, n\}$.

The permutation function $\Psi(i)$ relates to the sequence, in the order of which the evaluative properties are considered when characterising two chemical substances x and y, or indeed the relation between the two chemicals. The requirement asks for symmetry with regard to the evaluative properties considered. This is important when evaluative properties are to be of equal importance. In case Requirement 2 is not fulfilled, evaluative properties are clearly not all of equal importance (cf. counter-example in Section 3.2). The requirement is often found in the literature as the requirement of *neutrality* or *anonymity* [90].

---

[*] The denomination of this requirement is homonimous with the property of symmetry of an ordering. The coincidence is accidental and the two properties are unrelated.

4.4.2 Special Requirements

In contrast to the general requirements, *special requirements* have to be articulated which reflect the special nature of the decision problem as a precautionary pre-screening of chemicals. To start, we remember an observation made at the beginning of Chapter 3, namely that — in the present context — the linear ordering $L_{\preceq}$ must contain at least 3 classes (Observation 2). There is no reason why more than three classes should be allowed, and hence we may sharpen the observation to the following requirement.

Requirement 3 (three-valuedness) *The final order should yield three distinct classes.*

This Requirement is specific to the present classification problem (cf. Introduction). It could be easily adapted to any number of classes. Presently the three classes are labelled "green", "orange" and "red", as shown in Figure 3.5.

Requirement 4 (Transparence) *If a chemical is classified in the maximal class, the classification should be traceable to the threat scenarios that are the source of this classification.*

This requirement is also specific to the present problem, but maybe less so then the previous Requirement 3: there may be other multi-criteria decision problems where it is of interest to know based on mainly what criterion (or criteria) an element is ranked highest.

Requirement 5 (Expandability1) *If a set of evaluative properties $E = \{\ldots, e_k\}$ has classified a chemical in the maximum class, then this should remain so if the set of evaluative properties is expanded by one to $E = \{\ldots, e_k, e_{k+1}\}$.*

The present work studies the classification of environmental chemicals based on threat scenarios. Two such scenarios are presently used. No claim is made, however, that these two threat scenarios form an ultimate set of scenarios, as we have observed at the end of Chapter 2 (Observation 1). It is on the contrary very likely that additional scenarios will be discussed in the future, as indeed they are already today. In order to ensure that the consideration of additional scenarios does not distort the results obtained before, the procedure should not change the classification of chemicals in the maximum class if additional scenarios are broadly accepted and introduced in the future.

Note that at the same time, the classification of chemicals for a given set of threat scenarios is not immutable: it may be that a threat scenario itself is modified, by using other constituting parameters or another function that quantifies the scenario. Such changes in the *setting* of the scenario are not excluded by Requirement 5 (Expandability) and can lead to changes in the classification of a substance.

A proper ”violation” of Requirement 5 (Expandability) can furthermore occur if one or several of the other requirements are modified, drop out or if new requirements are added. Such changes could (and most certainly would) influence the outcome of the assessment, which would be acceptable.[†]

4.4.3 Consistency of Requirements

Ever since Arrow's classical work on the inconsistency of a set of certain widely accepted conditions for social welfare functions (Impossibility Theorem), the relationship between such conditions has become the focus of much inquiry [2, 87].

Hence an important task of the present work consists in showing that requirements in R are consistent. But it will be important also to state if such consistency is limited to certain special cases; Or put differently; what consistency of the requirements implies. As scetched out in Figure 3.1 fulfillment of the requirements may indeed restrict the fashion in which some key steps in the definition of the procedure are conducted.

As detailed earlier, the consistency of the set R will be shown through the existence of a linear ordering of the corresponding event-space that respects all requirements in R (thus ensuring that C^R is never empty). It will be shown with the help of an example that there exists a function that incorporates all requirements made herein, and hence that C^R is not empty, i.e. that the requirements in R are consistent. We must point out at this point, however, that this analysis is founded Conjecture 1.

4.5 Implications: Calibration

Two or more requirements may be inconsistent, as is the case for Arrow's conditions for social welfare functions [87]. In the parlance of the introduction, C^R is then always void, i.e. there is no function that satisfies all requirements in R concomitantly. As another possilibity, it may also be that although requirements are

---

[†] This very important observation was pointed out to the author by Prof. Nils-Eric Sahlin.

consistent ($C^R \neq \emptyset$), they may nevertheless be inconsistent with some additional aspect of the classification procedure, such as certain methods of optimisation. Note that a calibration is necessary, as we pointed out in general terms at the beginning of the chapter. If a type of optimisation O_j is not compatible with the requirements, then we have $C^R \cap C^{O_j} = \emptyset$ (see Figure 3.1 in Chapter 3). We will now show that one such type of optimisation for the classification procedure is indeed incompatible with the requirements in R. (The proof of Consequence 2 is shown in Appendix E.)

Consequence 2 *Assume the classification function is calibrated with an optimality criterion concomitantly for all evaluative properties with the help of a reference, then if follows that the Requirements 2 (Symmetry) and 5 (Expandability) are not compatible.*

Corrolary 1 *If the Requirements 2 und 5 hold, then it follows that the optimisation of the classification procedure with respect to a reference has to be conducted for each of the two evaluative property e_i (i.e. for each threat scenario) separately.*

Corrolary 2 *It follows from Corrolary 1 and Requirement 3 that each evaluative property must yield a linear ordering of three equivalence classes.*

Corrolary 3 *In the special case that a parameter occurs in the context of two or more threat scenarios, it follows from Corrolary 1 that a unique discretisation of that parameter is generally not possible.*

We are now in a position to re-label the classes of functions introduced in Figure 3.1, and thereby find a graphical representation of the inconsistency or consistency of certain types of optimisation. The illustration is given in Figure 4.3. The necessity for the optimisation to be conducted will be discussed at the end of Chapter 6.

4.6 Additional Requirements

Three requirements do not figure on the list that are often encountered in decision theory: *Non-dictatorship*, *Unrestricted Domain* and *Independence of Irrelevant Alternatives.* These three requirements together with Requirement 1 (Order-preservation) form the classical set of requirements introduced by Arrow in his proof of the Impossibility Theorem [87].[‡]

[‡] Requirement 1 is often called the *Pareto Principle* in the Arrovian context of social welfare functions, including in Arrow's book [87].

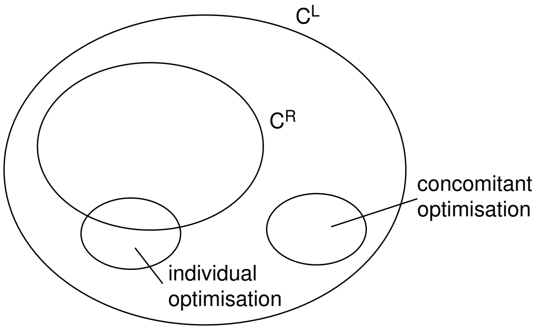

Figure 4.3: The class of functions C^R that reflects all requirements is in C^L (and non-void): the requirements are consistent. The class containing functions resulting from individual calibration overlaps with C^R and hence is consistent with all requiremens in R. For the class of functions that results from concomitant calibration, this is not the case.

Non-dictatorship
To start, a (historically) important and often used requirements is that of *non-dictatorship*. This requirement states that if e.g. individuals vote on a set of alternatives then the opinion of one individual (i.e. his ranking of alternatives) should never prevail over all other opinions. In the present context we have not given prominence to any evaluative property and hence this requirement can be expected to be fulfilled.

Unrestricted Domain
The Requirement of Unrestricted Domain states that the classification procedure should include all possible combinations of linear orderings for the evaluative properties. This condition is readily met here since we have at no point imposed any restriction on these individual orderings and since we discuss the classification procedure precisely on the event-space (i.e. all possible events).

Irrelevant Alternatives
The requirement states that the ordering of two elements should under no circumstance depend on the ordering of a third element. This includes the case where a third element is dropped from the set of elements entirely. This requirement is important in the present approach, since the ground-set can change. Note, however, that the problems is of a special kind here: the requirement can be violated by a change in ordering during the application of the classification problem on some set of chemicals to be evaluated (one type of ground-set). It can, however, also apply

to the *calibration* of the procedure (different kind of ground-set). The linearisation function as it is dictated by the other requirements respects the requirement of Irrelevant Alternatives. The classification procedure as such, however, doesn't really, since it incorporates the calibration step which is sensitive to changes in the reference sets. It seems, however, that the violation of the requirement during the calibration process is inherent to any calibration process and hence is of minor importance.

These three additional requirements are quite general, and based on the description of the requirements, there is no apparent reason why these requirements should not apply to the present problem. This carries the consequence, that the assessment procedure may be invalidated by the Impossibility Theorem, as the theorem shows that the requirements together with Order-Preservation are mutually inconsistent.

The Impossilibity Theorem, however, only holds if we accept that cardinalisation of individual utilities is not meaningful [87]. This assumption can be questioned, has in fact been challenged, and early on "levels of discrimination" (introduced by Goodman and Markovitz [88]) have been applied in order to obtain cardinalisation. In the current problem of classifying chemicals, it appears that such levels of discrimination are valid. If this is assumed[§], then the mechanics of the proof of the Impossibiltiy Theorem vanish and the set of requirements is again consistent. (For a good overview on the topic of interpersonal aggregation in the context of rational choice, see Chapter 7 of [2]. The Chapter also contains references to various other methods of cardinalisation.)

When introducing the linearisation function in Chapter 5 we will show that all requirements that are introduced are also mutually consistent, by simply introducing an ordering of the event-space that fulfills all requirements (and a corresponding function). For the time being, we assume this to be the case.

4.7 Classification Scheme

We can now use the requirements, definitions and consequences to simplify the classification procedure and draw the corresponding *classification scheme*. The definitions and requirements relevant to the scheme are: Definition 3 (Partial Ordering), Definition 4 (Linear Ordering), Corrolary 2 and Requirement 3 (three-valuedness).

[§] This assumption is technical and appears to be so straightforward as not to require further discussion.

As the simplest non-trivial case (and the one concerning us) we study the classification scheme for two threat scenarios. (The corresponding case for three evaluative properties results in the event-space shown on the right in Figure 4.2). The most general case for the scheme is obtained when assuming that all possible tuples are actually occupied by a chemical, i.e. the *poset* is exactly equal to the $event - space$.

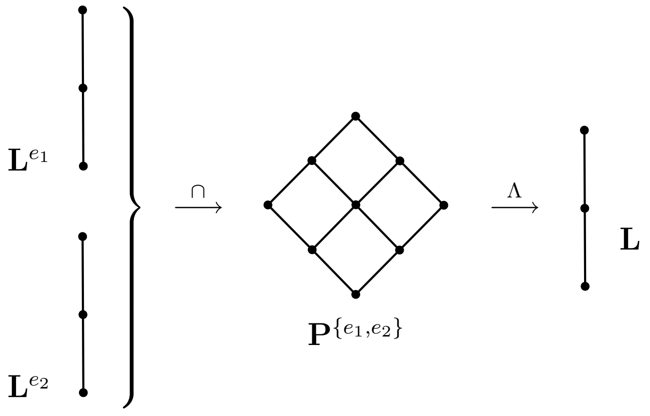

Figure 4.4: Scheme of classification procedure: Two threat scenarios e_1, e_2 resulting in the corresponding linear orders $\mathbf{L}^{e_1}$ and $\mathbf{L}^{e_2}$ are combined to the poset $\mathbf{P}^{\{e_1,e_2\}}$, which in turn is mapped by linearisation function Λ to final linear oder $\mathbf{L}$.

Inspection of Figure 5.1 shows that the formal requirements and consequences therefrom derived have significantly simplified the classification procedure. The classification procedure starts with coarse-grained linear orders $\mathbf{L}^{e_1}$ and $\mathbf{L}^{e_2}$ for each evaluative property. These *ordered sets* then define the poset $\mathbf{P}^{\{e_1,e_2\}}$ (which is relevant for the overall linearisation) via the intersection operator $\cap$. Finally that poset is linearised with the overall linearisation function Λ to the linear order $\mathbf{L}$ which is the overall result of the classification procedure.

The most important feature of the scheme depicted in Figure 5.1 is that the poset $\mathbf{P}^{\{e_1,e_2\}}$ is in fact already discrete: the poset $\mathbf{P}^{\{e_1,e_2\}}$ which is the domain of the linearisation function is already coarse-grained. To appreciate the simplification, it is illustrative to imagine the event-space in case the evaluative properties each were represented by a linear ordering of, say, 10 equivalence classes.

In contrast to the general case which was encountered in the example of solvents, the poset $\mathbf{P} = \langle X, O\rangle$[¶] is itself a discrete *poset* as if it was a mapping from a

---

[¶] The order relation O is the relation defined for the ordering of all events and is used since the depicted *poset* is the event-space: $\mathbf{P} = \mathbf{O} = \langle X, O\rangle$

continuous poset. Rather, since the coarse-graining is performed on the evaluative properties, the un-modified, continuous poset presently never occurs. Note that in view of Eqn. 4.1, the correct denomination for the poset is $\mathbf{P}'^{\{e_1,e_2\}}$ and $\mathbf{P}^{\{e_1,e_2\}}$ is somewhat incorrect as it should be reserved for poset corresponding to the continuous characterisation. Since, however, no continuous poset ever appears as a result of characterisation and for the sake of simplicitiy we keep to $\mathbf{P}^{\{e_1,e_2\}}$.

We may of course try to picture the "original" *poset*. In order to do this, we look at the discrete, linear orderings for each of the properties as real-valued functions that have been coarse-grained to the discrete properties (by placing *cut-off* values). The "original" *poset* would then be given by the "original", real-valued evaluative properties. In this sense, the coarse-graining function Δ (see Eqn. 4.1) consists in mapping each continuous evaluative property to a discrete linear ordering.[‖] With the second step of linearisation by the function Λ, the assessment function Π is then fully defined.

The situation is complicated by the fact that *no such real-valued functions* are immediately at hand for the evaluative properties and hence the method of visualisation of the last paragraph remains hypothetical. In fact, the evaluative properties must be first constructed from two parameters (see Chapter 6.2). The optimisation is thus performed with the choice of function that aggregates the two parameters, directly to the three discrete intervals that define the equivalence classes. This is a form of linearisation, too: But contrary to the linearisation of threat scenarios, linearisation here occurs as a special case since it is used to define the function representing the evaluative properties. In the general case this linearisation would not occur (as the real-valued functions for the evaluative properties could be assumed to be unambiguously established), whereas linearisation of the characterisation (i.e. the poset) always occurs as a central feature of the classification problem.

4.8 Stability Problem

Let us presently assume that all requirements are consistnet. We may then inquire what would happen if one or several of the requirements were changed. The following may just serve as a rough overlook on *type* of modifications of the requirements, and their possible consequences.

[‖] For each evaluative property e_i there would be a separate coarse-graining function Δ_i which would make use of specific cut-off values.

Should one or several of the requirements be weakened, this results in a weaker restriction on the class of admissible functions by the requirements, and it may turn out to be problematic to choose a classification function from C^R without making further assumptions.

If, on the other hand, one or several of the requirements were strengthened, then this could possibly result in the requirements no longer being compatible. Such is the case if it is assumed that the cardinal values for the evaluative properties are not meaningful (in the sense of a strengthening of the requirements), in which case the requirements may no longer be consistent, i.e. there may exist no assessment function that can reflect all requirements concomitantly.

4.9 Summary

The relevant setting of the assessment procedure has been established using the concept of partially ordered sets, and mappings thereof (see Chapter 3). It has also been shown what kind of requirements apply to such mappings, and that such mappings are consistent[**] (if in a technically restricted way). Hence we have identified an adequate formal structure for the assessment problem (cf. Problem 2).

[**] We foreclose the existence of a function that fulfills the requirements, and which is presented in the next chapter.

Chapter 5

Linearisation

The scheme in Figure 5.1 establishes that each evaluative property (threat scenario) is represented by a linear order of three equivalence classes. It also exhibits the fact that the poset (i.e. the characterisation of chemicals with the evaluative properties) has to be *linearised* to a final (linear) classification by a linearisation function Λ. The resulting classification is again a linear order of three equivalence classes. The aim of the present chapter is first to establish whether concomitant fulfillment of the requirements is possible. In the affirmative case, the second part will consist in finding an appropriate linearisation function.

5.1 Target ordering $\mathbf{L}^R$

In order to check for the fulfillment of the requirements in R we will proceed to find a linear ordering that reflects all requirements in R; we will look at possible orderings of the event-space in order to achieve this *target ordering* $\mathbf{L}^R$ (*loset*) that reflects all requirements, without discussing the exact nature of the linearisation function Λ for the time being. (If the ordering is asked to fulfill all requirements except Requirement (i), the corresponding target ordering is denoted as $\mathbf{L}^{R\backslash i}$.)

In a second step we will look at how the linear ordering can be operationalised, i.e. what function yields that linear order $\mathbf{L}^R$ or an extension thereof. Since the *loset* $\mathbf{L}^R$ is derived from $\preceq$ we will use the symbole $\mathbf{L}^R_{\preceq}$.[*] In general, an ordering derived from $\preceq$ that fulfills a requirement i is denoted $\preceq^i$, if it reflects both i and j (in that order), $\preceq^{i,j}$ etc.

---

[*] This reflects the fact that the underlying ordering relation L^R is derived from $\preceq$: $L^R = L^R_{\preceq}$

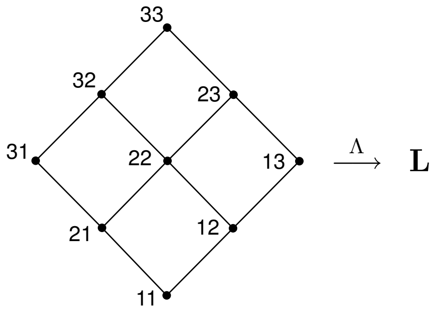

Figure 5.1: Mapping of the event-space to the target ordering $\mathbf{L}^R$

5.2 Weak Order

But before we discuss the requirements and their influence on linear orderings in more detail, *weak orderings* should be introduced [86, 89]. Weak orders have the somewhat special property that, although neither being linear nor complete, they can be linearised to a linear order in one single, unambiguous step. Hence if a procedure yields a weak ordering, then the linearisation does not require any further choices or assumptions with respect to the linearisation function.

Definition and General Application

An order relation that is *anti-symmetric*, *reflexive*, *negatively transitive* is called a *weak order* (see Appendix A). Negative transitivity states that for two comparable elements x, y, *any* third element z is comparable to one or both of them.[†]

$$\text{negatively transitive:} \quad \forall x, y, z \in X,\ xRy \Rightarrow xRz \text{ or } zRy$$

It can be shown that this crucial property (together with the less distinguishing properties of anti-symmetry and reflexivity) leads to the relation $\|\subset\approx$. I.e. in a weakly ordered set all incomparable elements are interchangeable (Theorem 2 in [86], page 5).[‡] In Figure 5.2, the poset depicted on the *right* is not a weak order as it contains two elements b and c for which $b \not\approx c$ although $b \parallel c$. The interchangeability in the weak order on the other hand has a ready interpretation, namely that of a *tie*: for any two elements in (the graph of) the poset, the decision maker must

---

[†] Put differently, if the two comparable elements x, y were taken together and perceived as one, then this combined pair would be comparable to any other element in the weak order. We may already feel that we are close to a linear order.

[‡] In Fishburn book the theorem states $\sim=\approx$ for a weakly ordered set defined using a strict relation. In the present case we use a non-strict relation, and the relation becomes $\approx=\{\| \cup \simeq\}$. In both cases, the theorem includes the statement that all incomparable elements are interchangeable.

either prefer one over the other or be indifferent among them. In the latter case, in that particular characterisation it is not possible to differentiate among the two elements. Hence if the elements are chemicals among which a choice has to be made, the decision maker must express an indifference between two interchangeable elements, since there is no information contained in the poset that would offer ground for differentiation.

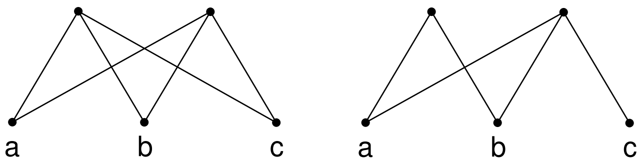

Figure 5.2: *Left*: All elements have same upsets and downsets, hence the poset is a weak order. *Right*: Elements b and c have different upsets, hence poset is not a weak order. Note that when the poset is a weak order, it is also negatively transitive.

Thus, if we want to ensure that linearisation is unambiguous, then all elements that are not comparable in the poset must be related by the relation $\approx$. Since in a weak order all elements which are incomparable are interchangeable, such an order is interesting from a linearisation point of view, as albeit being non-linear, the linearisation is so straightforward that the weak order is just *formally* non-linear: Since all possible events are now either related by $\preceq$ or $\approx$, the weak order can be linearised by a partition of the set of elements based on the equivalence relation $\approx$ (Theorem 2 in [86], page 5):

Theorem 1 (Fishburn) *If $(X, \preceq)$ is a weakly ordered set then for the symmetric complement of $\preceq$ we have $sc(\preceq) = \{\| \cup \simeq\} = \approx$ and the partition $(X/\approx, \preceq)$ is a linearly ordered set.*[§]

The relationship between a weak ordering and the corresponding linearly ordered set as set forth in Theorem 1 will expressed by the symbol $\equiv$: If $(L, \preceq) = (X/\approx, \preceq)$ then $(L, \preceq) \equiv (X, \preceq)$.

Hence a weak order is a suitable penultimate order in a decision procedure, and it is admissible to use a map to a weak order for the purpose of aggregation, just prior to linearisation.

[§] For the purpose of the present work, the theorem, which is written for the strict relation $\prec$ by Fishburn, is adapted to the non-strict relation $\preceq$ by replacing $sc(\prec)$ by $sc(\preceq)$.

5.3 Requirements Check

We can now combine two previous results

1. Instead of finding directly a linear order, we can look for an intermediate weak order which then is readily linearised.
2. Such discussion of the linearisation function can be performed on the event-space.

We start by taking $\mathbf{3} \times \mathbf{3}$ as the event-space and then apply the requirements one after the other, until we have an ordering reflecting all requirements in R (cf. Section 4.4). We start with Requirements 5 (Expandability) and 2 (Symmetry) and later check for the remaining requirements.

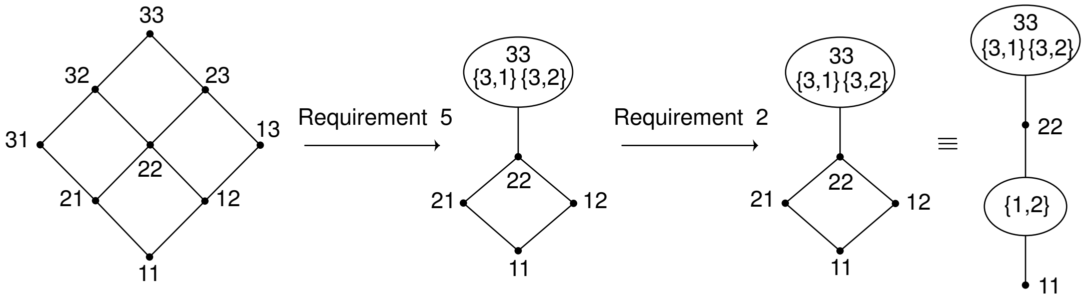

Figure 5.3: The ordered event-space is first extended into an ordering respecting Requirement 5. The ordering is already a weak order. In a second step Requirement 2 is introduced, which does not modify the ordering. Pro memoria: 13 is shorthand for the ordered tuple $(1, 3)$ etc.

Requirement 5 (Expandability) requires that once a chemical has been classified in the highest class, the chemical should remain in that highest class, even if the set of threat scenarios is enlarged. Since classification of a chemical in the highest class is denoted by 3 and since the Requirement must hold for any number of evaluative properties — i.e. also for one evaluative property — it follows that all events that contain the number 3 (once or more) have ultimately to be in the same equivalence class.

Requirement 2 (Symmetry): in the present poset there are three situations to which symmetry applies: for the pair of tuples $(1, 2)/(2, 1)$, $(1, 3)/(3, 1)$ and $(2, 3)/(3, 2)$. In order for this Requirement to be fulfilled, it must be assured that for each pair, the tuples are equivalent in the sense that they are indistinguishable, i.e $(1,2) \approx (2,1)$, $(1,3) \approx (3,1)$ and $(2,3) \approx (3,2)$. This is already the case for all three pairs of tuples. As an immediate result the poset doesn't change if Requirement 2 is applied.

Crucially, the poset that reflects Requirements 5 and 2 is a weak order as can be readily seen from inspection of the corresponding graph in Figure 5.7: each pair of incomparable elements contains only elements that are interchangeable. Hence we can use Theorem 1 and obtain a complete partial order, i.e. a linear order. We continue with the remaining requirements to check for consistency.

Requirement 1 (Order-Preservation) is clearly fulfilled, as we have started with the event-space and have only considered extensions thereof. Hence $\preceq^{5,2,1} = \preceq^{5,2}$ and the requirement is fulfilled.

Requirement 4 (Transparence) is fulfilled since the classification procedure yields individual orderings for the evaluative properties, and if a chemical is eliminated, this is due to the fact that for (at least) one of the evaluative properties, the highest value is attained: $\preceq^{5,2,1,4} = \preceq^{5,2,1} = \preceq^{5,2}$.

Finally we apply *Requirement 3* (three-valuedness). All there is left to do is to regroup the top 3 elements/classes into a single class, as well as the bottom 2 elements, with the help of cut-off values. Three possibilities exist to place the cut-off values. We choose to place one between the class containing $(3,3)$, $\{2,3\}$ and $\{1,3\}$ and the event $(2,2)$, as required by Requirement 5 (Expandability). The other cut-off value is placed somewhat more arbitrarily between 22 and $\{1,2\}$.

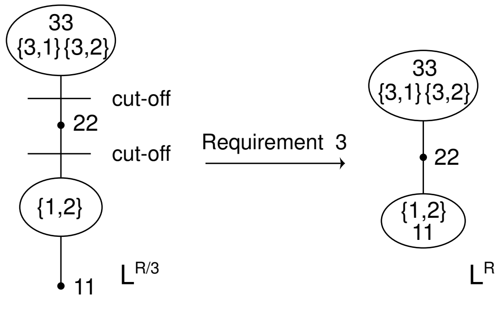

Figure 5.4: The ordering $\mathbf{L}^{R\setminus 3}$ of the event-space $\mathbf{3} \times \mathbf{3}$ after application of Requirements 1, 2 and 5 is mapped to the target ordering $\mathbf{L}^{R}$ reflecting Requirement 3 (three-valuedness).

Further Requirements (cf. Section 4.6)
Non-dictatorship is clearly respected by the ordering obtained, as is Unrestricted Domain. Independence of Irrelevant Alternatives is not applicable, since the ordering executed by the operator is performed on the entire event-space; hence no "irrelevant alternatives" exist. Furthermore, the three requirements do not appear to restrict the ordering any further (i.e. no extension is required for fulfillment). As mentioned before, however, the use of the event-space as paradigmatic for all posets assumes some cardinalisation of the ordinal orderings of the evaluative properties.

The discussion of the Requirements Non-dictatorship, Un-restricted domain and Independence of Irrelevant Alternatives will also be conducted on the operator that yields the target ordering at the end of this chapter.

Relevance of Sequence of Requirements
We have applied the requirements in a particular sequence above, which we may question. The question immediately arises, if such another sequence would have produced another outcome. Somewhat surprisingly, the sequence with which the requirements are imposed on the function appears to be relevant: If Requirement 2 (Symmetry) is applied before Requirement 5 (Expandability) then a different ordering is obtained than if the converse sequence is applied.

Comment 1 *The sequence of imposure of the requirements influences the outcome of the target ordering: If Requirement 2 (Symmetry) is imposed before Requirement 5 (Expandability) then a different target ordering is obtained*

This appears to be a quite important observation, for which, however, we have no ready explanation. We cannot exclude that this fact could turn out as a limitation of the current approach of finding an appropriate aggregation function. The corresponding logical consequences are not clear, however.

5.4 Operationalisation

Although the type of linear ordering which is compatible with the Requirements 1 to 5 for $\mathbf{3} \times \mathbf{3}$ is now known, the actual funtion Λ leading to that linear ordering still has to be found. Of course, the order can be achieved in *that particular case* by giving detailed instruction as to how the events should be ordered, i.e. in essence give a linear order in the way it was just done in Section 5.3. The goal, however, is to have a function that independently of the characterisation (i.e. the number of

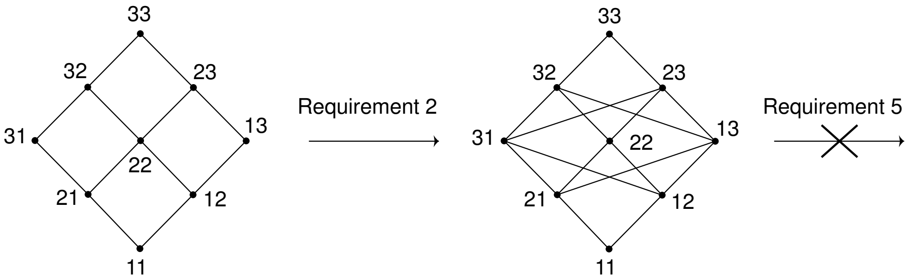

Figure 5.5: If Requirement 2 is applied first, then the partial ordering becomes a weak ordering that allows for no further differentiation among $(2, 2)$ and $(1, 3)$. This precludes the possibility to apply Requirement 5.

evaluative properties and the number of discrete intervals) always yields a classification procedure compatible with the requirements. In short, an *operator* is required that *operationalises* the requirements.

It is important to remember that all the information at hand in the frame of the classification is given by the evaluative properties. There is no information in addition to that given by the evaluative properties (i.e. the threat scenarios), on which the linearisation can be based. Also, all the information given by the evaluative properties is given in the poset, which in turn is embedded in the event-space. Hence any differentiation made by the linearisation function among elements incomparable in the poset has to be based on some properties chemicals exhibit when viewed as elements of the event-space. Such properties are *structural information*, in the sense that they can be defined with respect to the event-space. The goal is hence to find an operator that yields $\mathbf{L}^R$ by using only *structural information* within the poset.

5.4.1 The Sequel Operator

It turns out that a function that operationalises the requirements in the way that at least partly reflects $\mathbf{L}^R$, is the so-called *sequel operator* $S(\prec)$. This operator is derived from the theory of interval order and semiorders [86, 89] and is used in situations where measurement of data gives only a partial picture of an order which is known to be essentially linear. The interesting aspect of the sequel operator is that linearisation is performed in such a way, as to make incomparable chemicals comparable on the basis of structural properties they possess within the partially ordered set.

The definition of the sequel operator is somewhat involved and requires a number of additional ordering relations (see Appendix F). The following definition is a summary of Appendix F, which in turn is taken in large parts from Fishburn's book on interval orders [86]. As noted earlier, note that Fishburn uses the order relation $\sim$ to denote incomparability. In his formalism, however, this also includes identity (since he uses strict order relations). As presently non-strict relations are used, we have $\sim = \|$ (Eqn. 3.6). Hence the relation $\sim$ can be presently replaced with $\|$ for the definition of the sequel operator. We stick, however, to the denotation of $\sim$-relation, in order to facilitate comparison with literature.[¶]

$$\begin{aligned} & sS(\prec)t \text{ if for some } a,\, b \\ \text{either} \quad & s \sim a \prec t \text{ or } s \prec b \sim t \\ \text{but never} \quad & t \sim a \prec s \text{ nor } t \prec b \sim s \end{aligned}$$

The relation holds if s is incomparable to some a which in turn is smaller than t (but not concomitantly the converse) or if s is smaller than some b which is incomparable to t (and not concomitantly the converse). Thus the sequel operator ensures that for two elements x, y the composition of $\sim$ and $\prec$ relation holds only in one "direction" (e.g. $(\sim)(\prec)$) and not concomitantly in the other (i.e. $(\prec)(\sim)$).[‖] It is easily verified that each tuple of events for which $x \prec y$ also satisfies $xS(\prec)y$.

For the two events $(3,1)$ and $(2,2)$, clearly $(2,2) \prec (2,3) \sim (3,1)$. On the other hand there is no solution for $(3,1) \prec x \sim (2,2)$, since this relation holds for $x = (3,2)$ or $x = (3,3)$ (both of which are comparable to $(2,2)$). Analogously $(3,1) \sim x \prec (2,2)$ has no solution. As a consequence $(2,2)S(\prec)(3,1)$ and $(2,2)S(\prec)(1,3)$.

Furthermore $(3,1) \sim (2,2) \prec (2,3)$, while there is no x such that $(2,3) \sim x \prec (3,1)$ or $(2,3) \prec x \sim (3,1)$. Hence $(3,1)S(\prec)(2,3)$ and, by analogy, $(1,3)S(\prec)(3,2)$.

The linear ordering obtained with $S(\prec)$ is shown in Figure 5.6. Note that all we need to do at this point is to introduce cut-off values, in order to obtain L^R. But the same loset as obtained with $S(\prec)$ can be obtained by applying the requirements and replacing Requirement 5 (Expandability) by a weaker consequence of the requirement.

[¶] The relation $\sim = \|$ is obtained for the definition $\sim := sc(\preceq)$. If on the other hand, we regard $\sim$ as the relation containing both incomparable and *identical* elements (as $\sim := sc(\preceq)$ is in the strict case), then $\sim = \{\| \cup \simeq\}$. The relation $\simeq$ adds nothing here, however, since we order events, which can never be identical.

[‖] More plainly, contradiction is avoided.

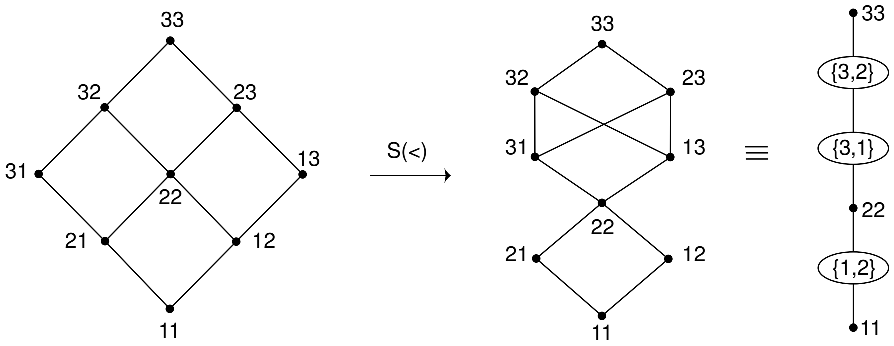

Figure 5.6: Poset 3×3 mapped by sequel order operator $S(\prec)$ to a weak order that is in turn readily linearised

Consequence 3 *The classification must order the tuples* $(2,2)$ *and* $(1,3)$ *in the following manner:* $(2,2) \preceq^R (1,3)$.

If we use Consequence 3 instead of Requirement 5 then we obtain the following sequence of orderings, if the other requirements are applied in the same manner as before.

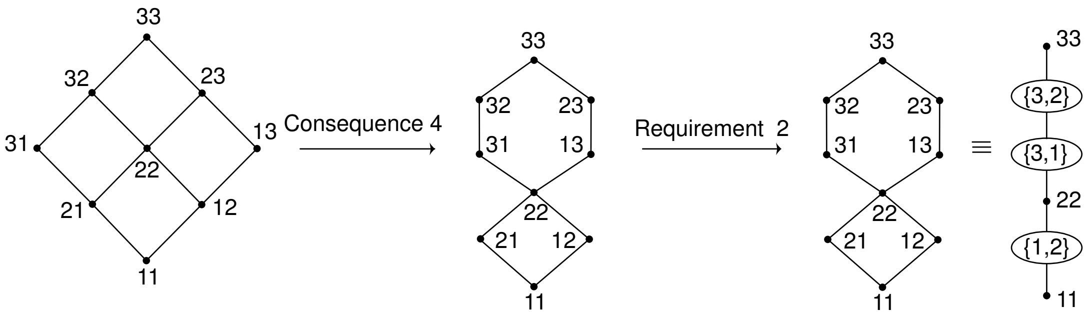

Figure 5.7: The ordered event-space is first extended into an ordering respecting Consequence 3. In a second step Requirement 2 is introduced, which introduces comparabilities between the tuples $(2,3)/(3,2)$ and $(1,3)/(3,1)$.

Hence the target ordering $\mathbf{L}^R$ can be obtained by simply using the sequel operator, and introducing appropriate cut-off values (as demanded by Requirement 3), which are part given by Requirement 5 and partly arbitrary (for the cut-off representing the "green"-"orange" boundary.) In other words, the sequel operator reflects all requirements except Requirement 3 (three-valuedness), if we use a slightly weaker form of Requirement 5. Note that the sequel operator appears to work also on $\mathbf{3} \times \mathbf{3} \times \mathbf{3}$ and beyond for higher dimensions (not shown).

5.4.2 Interpretation of Sequel Operator

The sequel operator implicitly also orders the tuples $(2,2)$ and $(1,3)$: $(1,3)$ is equivalent to $(3,3)$, for which in turn $(2,2)\preceq^5(3,3)$, hence in the linear order $(2,2)\preceq^5(1,3)$.[**]

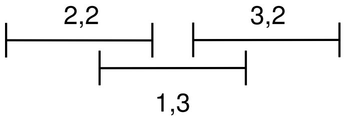

Figure 5.8: Example of operation of sequel operator on tuples. $(2,2) \prec (3,2)$ and $(2,2) \sim (1,3)$, $(1,3) \sim (3,2)$. Hence $(2,2) \prec (3,2) \sim (1,3)$ and not $(1,3) \prec (3,2) \sim (2,2)$ nor $(1,3) \sim (3,2) \prec (2,2)$ and therefor $(2,2)S(\prec)(1,3)$. Similarly $(1,3)S(\prec)(3,2)$.

Note that the only incomparable elements that are made comparable by the sequel operator are $(2,2)$ and $(1,3)$ and $(3,1)$. The illustration of the sequel operator is that they can be viewed as fuzzy measurements on a linear scale. In order to find out if the fuzziness blurs an underlying ordering of i.e. $(2,2)$ and $(1,3)$, we notice that there is another measurement (namely $(3,2)$) of which we know that it is larger than $(2,2)$, but not comparable to $(1,3)$ either. This can be interpreted as $(1,3)$ being the larger measurement, since it must be in a similar range of the linear axis than $(3,2)$, else they would not be incomparable. Graphically, this interpretation can be depicted with the events $(2,2)$, $(1,3)$ and $(3,2)$ depicted as intervals (Figure 5.8).

Another way to look at the sequel operator is to note that the structural feature it uses to differentiate among chemicals is their degree of isolation: the more extremal values a chemical assumes in the evaluative properties, the higher it is ranked by the operator.

5.5 Summary

At the end of Chapter 4, we already noticed that the requirements would turn out to be consistent, as we could find a function in the present chapter that fulfills all of

---

[**] It may be objected at this point that Requirement 5 actually forces all tuples containing an entry 3 to be placed in the same equivalence class. This is, in a weaker form, not necessarily so, since within the class of tuples that are isolated, a separate ranking may still be possible.

them. We have also found that calibration of the assessment procedure is necessary, and that such a process must be conducted separately for each evaluative property considered. We have also found that after application of the requirements, there is, in the present context, very limited choice to be made, in order to select the assessment function.[††] In fact the only such choice consists in placing one cut-off value: The linearisation function Λ is defined by $S(\prec)$ and appropriate cut-off values, of which one is already given by Requirement 5 (Expandability). Hence we have an adequate (mathematical) decision rule for the exposure-based environmental assessment of man-made chemicals (see second part of Problem 1) and other applications in natural sciences.

It is furthermore shown that the sequel operator $S(\prec)$ which is taken from the theory of interval orders, mirrors the requirements made presently, to a large extent. This function operates solely on information contained in the structure of the partially ordered set. This may indicate an interesting territory for finding classification functions that can be used in the context of environmental assessments.

[††] Remember that assessment is synonymous with the mathematical decision rule.

Chapter 6

Combination: Formal Precautionary Procedure

In Chapters 3 and 5 we have extensively discussed the formal structure and properties of the assessment of chemicals. We have thereby used a number of well-known and well-understood requirements from the theory of social choice. In Chapter 2 on the other hand, we have detailed the precautionary approach we have presently chosen, and identified the two threat scenarios *Pandora* and *Bioaccumulation* as the present foundation of an exposure-based assessment of chemicals, which embodies the said approach.

We have thus now discussed the two theoretical blocks announced in the introduction (cf. the scheme depicted in Figure 1.1), which are required for the definition of the *actual precautionary assessment procedure.*

6.1 Combination of the Two Theoretical Blocks

On one hand, the formal analysis has shown that chemical assessment can be formalised within the more general frame of a *classification* of chemicals. This general structure rests on *evaluative properties.* (In accordance with the more restrictive Requirement 3 the evaluative properties must yield linear orderings $\mathbf{L}^{e_1}$ and $\mathbf{L}^{e_2}$ of 3 equivalence classes.)

On the other hand, the *choice of precautionary approach* was based on the consideration of threat scenarios, two of which have been presently chosen. The results from the last and the current paragraph can now be unified, resulting in the embed-

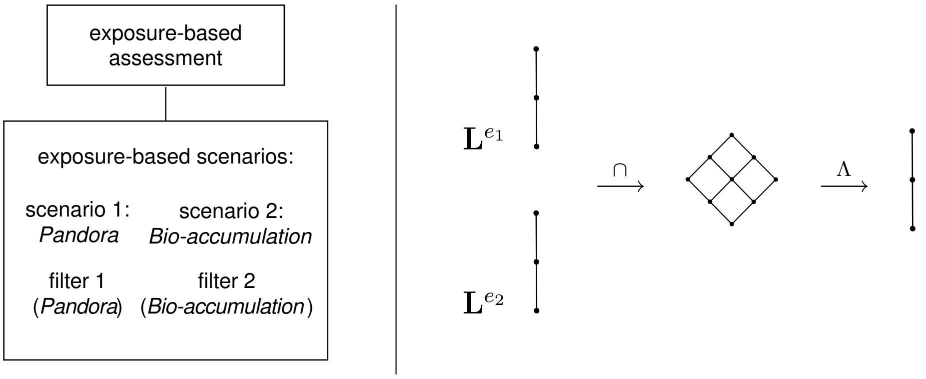

Figure 6.1: On the *left* is shown the result of the choice of precautionary approach, namely the choice of two threat scenarios *Pandora* and *Bioaccumulation*. On the *right* the formal classification is depicted, starting from the two linear orderings $\mathbf{L}^{e_1}$ and $\mathbf{L}^{e_2}$.

ding of the threat scenarios into the classification procedure: leading to the *formal precautionary procedure*.

The main results from these two theoretical blocks are shown Figure 6.1. It is evident from contemplation that somehow the linear orderings $\mathbf{L}^{e_1}$ and $\mathbf{L}^{e_2}$ depicted to the right and the threat scenario detailed on the left must correspond to each other. We cannot, however, directly equate the threat scenarios with the evaluative properties in the classification scheme, for the following reason: the evaluative properties yield linear orderings, whereas the threat scenarios may be higher-dimensional, as is the case for *Pandora* and *Bioaccumulation*.

We have noted this point earlier, at the end of Chapter 2: the *scenarios themselves* are not adequate for the purpose of conducting an assessment: scenarios must first be *quantified* and the quantified scenarios are called *filters* (Definition 1) We are now in a position to address this point in a more formal way. Taking the example of *Pandora*: the threat scenario is defined by two parameters (τ and ρ) and can be represented formally as a poset: $\mathbf{P}^{\{\tau,\rho\}}$. As detailed at the end of Chapter 2, an assessment results in a classification, which in turn yields a linear ordering (see beginning Chapter 3). Hence the quantification of the scenario consists at least in some sort of mapping of its own poset to a linear ordering (the *aggregation* part). Furthermore, this resulting linear ordering may still be modified by introduction of cut-off values, in order to define certain classes required by the assessment (e.g. "requires risk management" or "is completely innocuous"). Once the scenario has been processed to yield linearly ordered classes — among which a chemical can be allocated — the quantification of the scenario is completed. Importantly, no quantification of scenario is possible without linearisation.

We can now give a more formal definition of *filters*: The quantification of the threat scenarios consists in mapping a *poset* (defined by the threat scenario) to a linear ordering. This linear ordering can subsequently be transformed to a discrete linear ordering by introduction of cut-off values. The mapping that performs the linearisation (and possibly a discretisation, too) is called a *filter*. This new formalism permits a more concise consolidation of threat scenarios and evaluative properties.

Consequence 4 *The filters reflecting the threat scenarios in a quantitative manner yield linear orderings and correspond to the evaluative properties in the classification scheme.*

Consequence 5 *Since the evaluative properties are required to yield discrete orderings in the classification scheme, the filters include discretisation, too (beyond linearisation). Also, the calibration of the classification function consists in finding the values defining the discrete intervals for the filters.*

In the parlance of Figure 6.1, the identification of evaluative properties and filters corresponding to scenario can be expressed as

$$\begin{aligned} e_1 &= \text{filter}(Pandora) \\ e_2 &= \text{filter}(Bioaccumulation). \end{aligned} \tag{6.1}$$

It is important to note that a threat scenario can be defined by arbitrarily many parameters; it may also consist of simply one parameter, as we noted when threat scenarios were first introduced at the end of Chapter 2. If this was the case then no function would be required to represent the scenario in a quantitative manner. Hence the quantification of a threat scenario which we perform *via aggregation* of the two parameters for the scenarios *Pandora* and *Bioaccumulation* has no further meaning; beyond the fact that no function is at hand to reflect either of the two scenarios.[*] If such a function were at hand, discretisation would be required nonetheless, in order to satisfy the classification scheme. In this sense, the fact that the discretisation

---

[*] Importantly, although the aggregation of the parameters may resemble the *aggregation of the evaluative properties*, such similarity is accidental.

of the filters representing *Pandora* and *Bioaccumulation* is accompanied by the aggregation of the parameters defining the scenarios, will be abstracted from in the following. The interplay of these two operations will be delved into in the next section, when the construction of filters from indicators is discussed in detail.

Discretisation: Optimal Calibration with Respect to External Reference
As stated at the beginning of Chapter 3, the calibration of an assessment function is necessary for the definition of the assessment. The calibration cannot be conducted only on grounds of the formal structure, and additional information is required. Such is supplied by a *reference*. There are many ways such an operation can be performed, and it should not be ruled out that the discretisation is performed in a manner which is trivial in the present context, i.e. by asking a panel of experts to agree on the discretisation. Also, once such a reference has been established, the process of deriving the discretisation must involve some kind of comparison of the linear ordering resulting from the discretisation, with the linear ordering given by the reference. As also stated at the beginning of Chapter 3, we presently choose to perform the calibration by ensuring optimal correspondence of the linear ordering obtained with the assessment procedure and a reference.

Assumption 2 *The calibration of the procedure is most suitably performed by optimising the assessment function* Π *with respect to a reference.*

In the context of Corrolary 1 optimality must then be given for each of the scenarios separately. I.e. if the function is optimised for the scenario *Pandora*, then the resulting calibration is optimal only with respect to *Pandora*.† Also, optimisation then amounts to finding optimal values for the discrete intervals than define the two dimensions of each threat scenario.

We have at this point not specified how *optimality* is defined, nor what reference we intend to use. We may keep hold of the fact, however, that the present work aims at using as much *empirical* information as possible (see Introduction). In the present case, the relevance of the information with respect to threat potential will be derived from the fact that the information is extracted from existing global regulation, as was announced at the end of the introduction.

---

† It is irrelevant if another calibration for *Pandora*, together with "coordinated" calibrations for the other scenario yield an even better overall match between ordering resulting from classification and the reference.

Problem 2 *Based on the formal precautionary procedure, can a suitable assessment procedure be extracted solely from existing regulation?*

The formal structure of the assessment, together with the precautionary approach chosen in the present work has established that threat scenarios must be modelled as filters yielding linear orderings. We have also seen that the assessment procedure must at some point be calibrated, the orderings must be such that they consist of three equivalence classes, that the calibration must be conducted for each filter separately and that hence the filters *must* include perform some kind of discretisation. The discretisation, is in essence then the quantification of the threat procedure: it consists in finding values that define the discrete intervals for the linear orderings for each filter. So we may restate the problem:

Problem 2 (restated) *To what extent can — within the frame of the formal precautionary procedure — the filters be discretised solely through regard to existing regulations?*

We have not yet established any criteria for how this question could be answered. We just know that the formal precautionary procedure has no answers to that: the answer must be extracted from the result of the discretisation. If existing regulation allows to define the discrete intervals required by the formal precautionary procedure in a unambiguous manner, then the discretisation can be performed without problems. It may be, however, that this is not possible, or only in part. In short, the question is whether the information concerning global threat potential given by existing regulation is sufficient to perform the necessary discretisation.

The calibration of the assessment procedure will answer that question. Before that, however, we need to introduce the joint action of two threat indicators in a filter, i.e. the aggregation of threat indicators in each filter.

6.2 Constructing Evaluative Properties

As stated earlier and shown in Figure 6.1, the evaluative properties are in a one-to-one correspondence with filters (see Eqn. 6.1) which, in turn, correspond to the threat scenarios. At present, for the *Pandora* and *bioaccumulation* scenarios that are discussed no filters are defined in advance: we have no function at hand that maps the posets resulting from characterisation by the two parameters that define the two respective scenarios, into a linear ordering.

There are many ways in which such a mapping can be performed, such as cluster analysis and other methods from the realm of multivariate statistics.

In the present case, we have chosen a straightforward and simple type of aggregation function, which consists in placing a grid over the Cartesian plane spanned by the two parmeters. The grid divides the plane into *nine* areas.

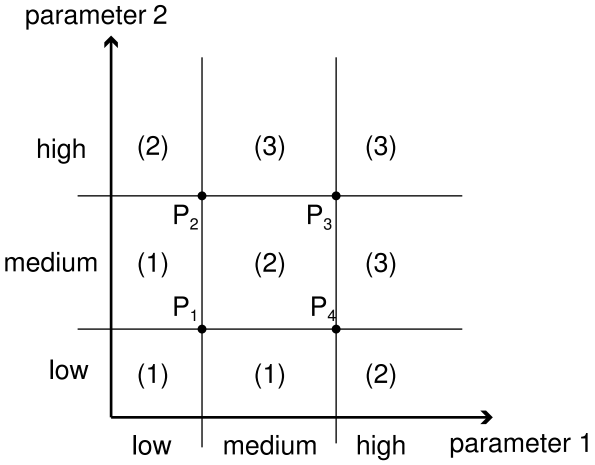

Figure 6.2: The grid that defines the three areas *low*, *medium* and *high* in the Cartesian plane. The pairs of points P_1/P_3 or P_2/P_4 fully define the grid.

The choice of aggregation function for the scenarios amounts to defining the $filters$. Instead of introducing cut-off values in a separate, additional step, the present function has the merit of integrating this step. Note also that the grid is defined by $four$ lines, two of which each define *discrete intervals* for the parameters: the intervals *low*, *medium* and *high*. The choice of three intervals for parameters is somewhat arbitrary, as a linear ordering of *three* could also be obtained by using a 4-by-4 grid. The areas just need to be abstracted accordingly. The argument in favour of 9 areas is that the aggregation abstracts an equal amount of areas into each equivalence class.

The grid divides the plane into nine areas which are subsequently grouped into three classes.[‡] The grid is defined by four lines (as depicted in Figure 6.2), i.e. four coordinates. The pairs of points P_2/P_4 or P_1/P_3 contain all information on the four lines. Hence, in order to describe the positioning of the grid on the Cartesian plane,

---

[‡] Note that the classes 1, 2 and 3 can, but don't necessarily coincide with the classes "green", "orange" and "red". It is only the case if only one threat scenario is studied in the assessment.

it suffices to describe the positioning of either the points P_2/P_4 or of the points P_1/P_3.

We presently choose the latter, as this allows us to define the areas *high* and *low* in a particularly simple manner: The area *low* is given by the union of areas containing points strictly smaller[§] than P_2 and P_4, whereas the area *high* is given by union of points strictly larger than P_2 and P_4. Note that with P_1/P_3 it would be slightly more cumbersome to define these areas.

It is important to note that the choice of this function could well have been another one and hence constitutes an additional assumption, namely that the threat scenarios can be *appropriately quantified* by this function.

Assumption 3 *The aggregation of the parameters defining a threat scenario, to a filter, is suitably performed by placing a* 3×3*-grid over the corresponding Cartesian plane and abstracting the resulting nine areas into three equivalence classes.*

In view of the requirements, there are objections which could be made against this choice. Conversely, the requirements do in no way favour this choice to the detriment of another. It cannot be ruled out either, that the future may yield more appropriate functions, possibly by using more sophisticated criteria to characterise the function (sensitivity, robustness etc.).

Equally, threat scenarios that may be additionally considered in the chemical assessment, may already have a linear representation, i.e. be defined as a filter. As an example, if toxicity was the sole indicator of such an additional scenario (which arguably is not in the spirit of the present assessment procedure), then LD_{50} would be such a quantified measure.

Note that we have hence almost completely defined the discretisation of the evaluative properties. Once the reference and optimality criteria are defined and the resulting cut-off values found, the discretisation is defined. In the formalism introduced in Chapter 4, the coarse-graining function Δ is then defined. And, since the linearisation Λ function has been described in the previous Chapter, the overall assessment function Π then is also defined.

[§] in the sense of the component-wise order, see Eqn. 3.1

Chapter 7

Calibration

The calibration of the classification procedure is performed by comparing a *reference* to the ordering obtained from the classification procedure. The calibration is achieved through selecting the best possible ordering compared to the reference. In order to appreciate the quality of a particular ordering in that regard, it is necessary to have certain *criteria* on which the comparison is based.

7.1 Choice of Reference Chemicals

As the classification procedure produces a ranking of chemicals, and as we choose to compare the result of the procedure to some reference, this reference must itself be some sort of ranking of chemicals. The following choice of reference chemicals is adopted from the selection conducted in the frame-work of the EU-projects which served as a starting-point for the present thesis. The following is meant as a brief overview of this process, and the corresponding references on the provenance of the reference chemicals and the according data is given in [12].*

The classification procedure that is proposed in this work aims at classifying environmental chemicals with respect to threat potential on a global scale. Accordingly the reference should on one hand contain chemicals that can be assumed to have a large threat potential (on a global scale), and on the other hand contain chemicals that can be assumed to have a limited threat potential on a global scale. The reference set should hence consist of two sets: one as paradigmatic for chemicals with

---

*It must also be said in this respect that an earlier calibration method has been proposed in [12], that is, however, less general than the method developed in the next two chapters. In fact the results obtained through that method are included in the results presented in Chapter 9.

a high threat potential, the other as paradigmatic for chemicals with a low threat potential.

We may choose a set of reference chemicals by inquiring e.g. a panel of experts. However, as pointed out in the introduction, the present work proposes to study a precautionary assessment based on as little as possible controversial information. The reference should therefore be as *empirical* as possible, where "empirical" includes scientific, economic and (to a lesser degree) political facts.

Assumption 4 *Existing regulation is with respect to large-scale threat adequately reflected by two particular reference sets: $Hset$ and $Rset$.*

With respect to large-scale environmental threats the universally regulated chemicals — namely those mentioned in the protocols of Montreal, Tokyo, and Stockholm — are taken as paradigmatic for the set of compounds with high necessity for precautionary measures ($Rset$), whereas a set of high-production volume chemicals ($Hset$) is taken as paradigmatic for globally innocuous substances, for which economic benefits together with the apparent absence of global risks lead to broad societal acceptance of global non-regulation, even if the substances are under scrutiny on a local scale, due to effects such as carcinogenicity or other toxic effects. (For all references to substances in the following sections including all sources of data, see [12]).

7.1.1 Reference Chemicals: High Threat Potential

Over the last few decades a number of large-production volume chemicals have been regulated on a global scale in the frame-work of international protocols (Montreal, Kyoto, Stockholm).

Rationale
The fact that a chemical is *globally* regulated is assumed to indicate *unanimity* or *consensus* in the global community's (including governments, NGOs, industrial pressure groups etc.) perception of the need for regulatory action.[†] Regulatory action,

---

[†] In light of the Bush's administration's continuous opposition to the Kyoto treaty, this may appear as a somewhat bold or even erroneous statement. Nevertheless, we assume that a large majority today accepts the significant risk emanating from green-house gases and that opposition to that treaty (and other treaties) is mainly du to procedural considerations or short-term political motivations.

in turn, documents the decision holder's perception that precautionary measures are required. In the frame of the present work, this equates to a large-scale threat potential of the chemicals for which regulatory action has been taken on the level of quasi-universal international protocols.

It may be argued that regulatory action can be taken even in cases where a chemical does not exhibit a large threat potential (as can be the case for a highly toxic chemical). But if this was the case, it is likely that no global regulation would be needed, as these chemicals can adequately be regulated on a local level.

As a result a selected set of globally regulated chemicals are presently taken as the reference for chemicals with a high large-scale threat potential. A total of 43 of such chemicals constitutes the first reference set which is denominated $Rset$.

The 43 substances were selected out of the lists of substances regulated by the three major international agreements on environmental chemicals: the Montreal Protocol (ozone depletion), the Kyoto Protocol (climate warming), and the Stockholm Conference (persistent organic pollutants, POPs). Most of the substances mentioned in these lists were selected according to the following criteria.

- Montreal Protocol: six substances were selected, as the most important compounds of the categories CFCs (CFC-11, -12, -21, -22), HCFCs (F-142*b*), and perchlorinated chemicals (CCl_4).

- Kyoto Protocol: out of the six substances or groups listed in the Kyoto protocol, only xenobiotics were selected, namely five organic chemicals[‡] of the two Kyoto groups hydrofluorocarbons HFC (HFC 134*a*, HFC-142*b*, HFC-152*a*) and perfluorocarbons PFC (CF_4, C_2F_6), as well as sulphur hexafluoride (SF_6).

- Stockholm Conference: all the major POP compounds included in the final document of the Stockholm Conference were selected. The insecticide *heptachlor* was not selected because, due to his very short lifetime, it was very soon substituted by its more stable oxidation product *heptachlor epoxide*. For the groups which include several compounds (eg. DDT group, chlordane group, PCBs, PCDDs, PCDFs) only the congeners that are most frequently referred to in the monitoring literature were chosen.

[‡] The Kyoto gases CO_2, CH_4, and N_2O where not included in the list because they are not xenobiotic and because their major emission sources are not specifically related to chemical production.

For the insecticide toxaphene (a mixture of hundreds of single compounds) the only data available in the used data bases refer to the mixture, and such data was taken.

7.1.2 Reference Chemicals; Low Threat Potential

A second set of chemicals is required to serve as a reference for chemicals that exhibit a low global threat potential. In contrast to the chemicals encountered in the previous section, such chemicals need not be regulated globally. They should also be produced and released in large quantities. The production volume is of primary importance, since only in large quantities, adverse effects make themselves necessarily manifest, whereas in smaller quantities, the threat potential emanating from a chemical may be lower, due to dilution. This implies that non-regulation is only important when seen in the context of production volume: should the volume go up, a high threat potential may result. As a consequence, only non-regulated chemicals that are produced in large quantities can be taken as paradigmatic for chemicals exhibiting a low threat potential.

Rationale
Furthermore, if a chemical is produced in large quantities (and hence is exposed to regulation), but has not been regulated on a global scale (again, the chemical may well be regulated on a regional level[§]), then this indicates that no need for precautionary measure is perceived by the global community. In the frame of the present work, this equates to the chemical not exhibiting a large threat potential.

Hence we assume that chemicals which — though being produced in large quantitites — are not globally regulated, can serve as reference for chemicals that exhibit a low global threat potential and hence are innocuous with respect to large-scale global risk. A set of (35) such chemicals constitutes the reference set for substances which don't exhibit a large threat potential. This set is denominated $Hset$ (see Appendix H). The set was selected by choosing the 35 most produced organic chemicals in the USA, according to their production quantities in the year 1995 [79].

Note also, that non-regulation can also result from the perception of the global community that although a chemical exhibits some signs of threat potential, economic importance and hence considerations overweigh.[¶] This would mean that if economic

---

[§] In the context of this work non-regulated always means non-regulated by international treaties, unless otherwise specified. Of course a chemical can be non-regulated in this sense while being regulated e.g. with regard to safety during production or transportation

[¶] Such was the case for the hydrogenated CFC, which replaced the CFC's. The HCFC were considerations would be absent, then such a chemical would be globally regulated. The problem poses itself, however, that any chemical that is in production is of economic importance (since else it would not be produced), and hence it may be difficult to isolate chemicals that can be said to have a high threat potential, but still are not regulated.

7.2 Criteria for Optimisation

The comparison of a linear ordering with a reference is not given naturally: *criteria* must be established, on grounds of which the comparison is conducted. One way to conduct such a comparison is to measure the coincidence of a linear ordering of chemicals with a reference, by using a simple score function which counts if — in the linear ordering produced by the classification — a chemical is in the *correct* class or not. (The correct class means the class in the linear ordering that corresponds to a particular class in the reference — in the case of the class of regulated chemicals ($Rset$) — the correct class would be the maximal class of the linear ordering, i.e. the class "red" (cf. Figure 7.1).)

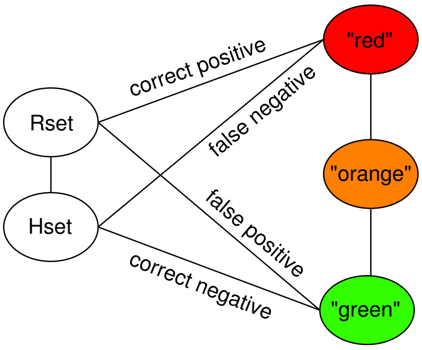

Figure 7.1: The binary ordering of the reference sets on the *left*, juxtaposed with the ternary ordering resulting from classification procedure on the *right*. The connections are labelled with the appropriate denominators and describe the types of relationships between classification inside the binary ordering and the ternary ordering. The relationships between the classes $Rset/Hset$ and the class "orange" is not shown in the graph, as they will not be used during optimisation.

---

believed to be ozone-depleting, too, but less strongly than the CFC's. The decision to use them was largely an economic one, and the HCFC were later banned in the Stockholm convention of 1992.

In turn, correct or false classification can occur both for elements in *Rset* (*positive*) and elements in *Hset* (*negative*).

- correct positive (elements of *Rset* classified in "red")
- correct negative (elements of *Hset* classified in "green")
- false positive (elements of *Hset* classified in "red")
- false negative (elements of *Rset* classified in "green")

In order to find the optimal solution for the calibration, a particular ordering is assessed with respect to the number of occurrences of the four cases listed. Note that with an ordering of chemicals into three classes, it is possible that a chemical is classified neither in the correct nor in the false class: it is then classified in the class "orange".

- Criterion a (maximum number of correct positive): As many elements as possible from *Rset* should be classified as "red"
- Criterion b (maximum number of correct negative): As many elements as possible from *Hset* should be classified as "green"
- Criterion c (minimum number of false negative): As few elements as possible from *Hset* should be classified as "red"
- Criterion d (minimum number of false positive): As few elements as possible from *Rset* should be classified as "green"

The alphabetical numbering of the criteria is accidental and does not express a ranking in the sense of importance of the criteria. Nevertheless, the *sequence* of the criteria in the application during optimisation reflects a priority setting.

7.3 Properties and Sequences of Criteria

There are $4! = 24$ different arrangements in which the Criteria a, b, c, d can be ordered to a sequence. In order to find the most suitable ones of these sequences, we must revert to requirements we have made to the classification procedure, and choose accordingly.

7.3.1 Excluded Sequences

To start, recall that Requirement 3 (three-valuedness) requires the classification procedure to yield three classes. This implies that some elements from the two reference sets $Hset$ and $Rset$ can populate the class "orange" (besides populating the two natural target classes "green" and "red"). Also, any sequence that precludes from the beginning that either elements from $Hset$ or elements from $Rset$ can ever be in the intermediate class, is ruled out.

As a consequence, there are two families of sequences we can immediately rule out, namely all sequences that start with Criterion a and b, respectively.

Proof
We prove the previous statement by assuming the converse: starting e.g. with Criterion a. Then, as consequence of the application of the Criterion, as many elements as possible in *Rset* must be classified as "red". Since no restriction has been made previously to the classification of elements of *Rset*, all elements in the set are immediately classified as "red". This leaves under no circumstances any possibility for an element of *Rset* to be classified as "orange", which indirectly violates Requirement 3 (three-valuedness).

The proof for Criterion b is analogous. Hence, no sequence can start with either Criterion a nor Criterion b and *must* start either with Criterion c or Criterion d.

7.3.2 Suitable Sequences

There are still 12 possible sequences to choose from after the exclusion carried out in the previous section. For reasons detailed in Appendix G, only sequences beginning with *cd* and *dc* are really meaningful in the context of the present work. But such sequences can again be further limited, due to the following observation, which is discussed in more detail in Appendix G:

Observation 3 *Criteria c and d can always be simultaneously and completely fulfilled. Any sequence that starts with cd produces the same result as if c and d were inverted at the beginning of the sequence, to dc.*

Criteria a and b on the other hand cannot be interchanged without affecting the outcome of calibration; and we have two sequences left that are equivalent: *cdab* and *cdba*.

In one case, Criterion b is placed before a and hence more weight is placed on the correct classification of the elements in $Hset$ (to the detriment of the correct classification of *Rset*). As such, the sequence *cdba* is somewhat favouring the *economic benefit* of chemicals in $Hset$ over *precautionary considerations* that are reflected in the correct classification of the chemicals in *Rset*.[‖]

In the other case, Criterion a is placed before Criterion b, and ultimately more weight is placed on *precautionary considerations* rather than on *economic benefit.* As such, the sequence *cdab* classifies the elements slightly more in favour of precaution.

Observation 4 *For the set of Criteria a, b, c, d that are presently used, there are two sequences of criteria that are suitable for calibration (i.e. to define the discrete values required for the definition of discrete intervals): cdba (Sequence* 1*) and cdab (Sequence* 2*).*

7.4 Sequence 1: Slightly more Economic

Sequence 1 can be reflected by applying the following sequence of criteria:

1. Criterion c: No chemical of $Hset$ should be classified as "red".
2. Criterion d: No chemical of *Rset* should be classified as "green".
3. Criterion b: If possible, a chemical from $Hset$ should be classified as "green" (not as "orange").
4. Criterion a: If possible, a chemical from *Rset* should be classified as "red" (not as "orange").

The sequence *cdba* and its constituting criteria shape the equivalence classes which result from the classification. The meaning and interpretation of the equivalence classes as a function of the sequence is studied in the Chapter 8. Remember that the ordering of the criteria c and d is meaningless, since Criteria c and d can be interchanged.

[‖] It doesn't follow, however, that Criterion b can be fulfilled entirely, i.e. that all element in $Hset$ can be classified as "green". On the contrary, in many cases Criterion d forces elements of $Hset$ to be classified as "orange".

7.5 Sequence 2: Slightly more Precautionary

In the case of Sequence 2 the sequence changes, as the focus has shifted to the non-regulated chemicals. As a consequence, the approach can be reflected by applying the following sequence of criteria:

1. Criterion c: No chemical of $Hset$ set should be classified as ”red”.
2. Criterion d: No chemical of $Rset$ set should be classified as ”green”.
3. Criterion a: If possible, a chemical from $Rset$ should be classified as ”red”, not ”orange”.
4. Criterion b: If possible, a chemical from $Hset$ should be classified as ”green”, not ”orange”.

As compared to Sequence 1, the Criteria a and b are inverted. Note that this is the only way to tilt the sequence of Criteria towards the correct classification of the non-regulated chemicals; as seen earlier, the inversion of Criteria c and d has no effect on the outcome of the classification.** Again, the interpretation of the equivalence classes is postponed to Chapter 8.

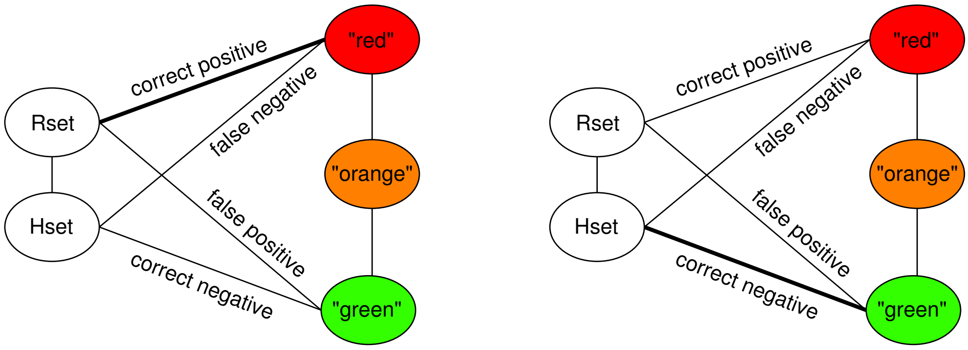

Figure 7.2: On the left hand side the sequence $cdab$ (more precautionary) is illustrated: as many elements from $Rset$ as possible are classified as ”red”. On the right the sequence $cdba$ (more economic) is illustrated: as many elements from $Hset$ as possible are classified as ”green”.

[**] Except for the possibilities to start with c, a and d, b, which we have previously excluded as non-meaningful.

7.6 Limitations of Current Approach

The present work explores the possibility of establishing an assessment in the form of a pre-screening of chemicals, in order to separate chemicals with high threat potential from chemicals with low threat potential. We also make the assumption that in order to define "green", "orange" and "red", we can use two sets of reference chemicals (Assumption 2). This assumption may be wrong for two reasons

1. A chemical which today is not globally regulated and part of $Hset$ can turn out to have adverse effects on a global scale. This may be the case both when a chemical is toxic in some way (and hence regulated locally) and when it is not.

2. There also exists the possibility that a chemical has been globally regulated on political rather than scientific grounds: A chemical can be globally regulated on a scientifically wrong perception of its potential adversity. Had production been permitted large economic gains and no serious global adverse effects would have resulted.

The two possibilities cannot be excluded, and additional criteria may be proposed in the future that address these possibilities. We can, however, study the influence that changes in the reference set would have on the result obtained with an assessment. The corresponding analysis is conducted at the end of Chapter 9, when sensitivity and robustness of the results are discussed.

7.7 Calibration Methods; Discussion Outline

The process that selects the optimal calibration with respect to the reference can be conducted randomly with a *Monte-Carlo simulation*. In this method, linear orderings that result from randomly chosen discretisations for each evaluative property, are compared to the reference. Based on some criteria for the comparison, each such random ordering is attributed a score. The optimal solution is then chosen by picking the ordering (i.e. discretisation) with the highest such score.

Alternatively, a *deterministic method* can be used to find the optimal classification procedure. Such a method aims at finding the overall optimal solution, regardless of how likely it is to be found by random choice.

In the present work, both a Monte-Carlo simulation and a deterministic algorithm were used for calibration. The deterministic algorithm in Appendix J was purposely written for this application and is explained in Chapter 8. The results from the Monte-Carlo simulation are not shown.

We have already been able to concentrate Problem 2 to a more restricted question in Chapter 6: namely if it is possible to perform discretisation of the filters based on existing regulation, if certain technical assumptions are made. In order to answer this question, we can make use of the results obtained from calibration. To see how this can be done, we first recall all assumptions made up to this point.

- The scenarios are modelled with the parameters τ, ρ and τ, K_{ow}, respectively (Assumption 1). The assumption covers the choice as well as the technical definition of the parameters.
- The discretisation is performed by calibrating the filters (evaluative properties) with respect to a reference (Assumption 2). The assumption also covers the criteria of optimality that are used.
- The two parameters that define each scenario are aggregated to a filter by using a simple grid that is placed over the Cartesian plane (Assumption 3).
- Existing regulation can be represented by two reference sets $Hset$ and $Rset$ (Assumption 4). This assumption includes the data used for the reference chemicals, which may be subject to uncertainties of measurement or computation. It also includes some belief that existing regulation represents information beyond purely political motivations, i.e. some realistic measure of the threat potential emanating (or not) from the reference chemicals.

Assumption 4 implies that a solution for discretisation of the evaluative properties — i.e. aggregation and discretisation of the threat scenarios to discrete filters — is optimal with respect to the reference. The solution may not be unique, and it may well be that within the formal classification scheme and with Assumptions 1 to 4, several solutions are possible. Remember that the grid used for definition of the 3 equivalence classes is defined by the two points P_2 and P_4 (Figure 6.2). Hence the solution of calibration concerns these two points. But for such a solution for P_2 and P_4, there are other solutions in the *vicinity* of the points which are equally optimal under the criteria. Hence, using the aggregation function established in Assumption

3, it is not discrete values, but rather *areas* that are obtained as optimal solutions: instead of obtaining exact coordinates for P_2 and P_4 we obtain areas, inside of which we can choose the points. These areas are denoted α_2 and α_4, respectively. Note that the simple definition of the areas *high* and *low* facilitates the calibration under the Criteria a, b, c, d, since the criteria are primarily concerned with ensuring or avoiding the classification of reference chemicals in these two classes.

Problem 2 can hence be restated once again: if the areas are defined by the assumptions in such a way as to yield a reasonable choice of discrete values, then no further assumptions need to be made in order to perform the discretisation, and discrete values can be picked out of the areas in order to define the filters. In other words, the restricting influence of the optimality criteria may be providentially such that the area, i.e. the remaining freedom to choose a definition of the classification procedure (through choice of appropriate values for the boundaries that define the threat scenarios), is practical.

We cannot exclude, however, that the restriction may be quite loose, i.e. the remaining freedom of choice is large. In such a case, since the empirical information has yielded the areas, this choice has to be expected to be more arbitrary with respect to scientific reasoning, i.e. more (*political*). The choice may thus require a significant change in the Assumptions 1 to 4 or even the consideration of additional, new assumptions.

The latter case is one which should be unnecessary, as the assumptions made should suffice for the definition of the filters. Is this not the case, the *set of assumptions should be revisited* as a whole and modifications should be discussed and performed carefully.

Discussion of areas α_2 and α_4
As a result, *Problem 2 can be discussed by inspecting the areas α_2 and α_4*. In case the areas α_2 and α_4 are such that a choice of parameter values can be performed, then the answer to Problem 2 is that indeed such a quantification is possible based on the formal procedure and with the help of an empirical, regulatory reference. The question arises, based on which criteria such a judgment should be made.

To give a summary answer, the discussion of the areas will be conducted empirically: we will avoid establishing criteria of the mathematical type, but will instead just comment on the areas obtained and discuss features of the areas which may not be satisfactory in the current context.

Unsatisfactory Features of Areas

These features relate to the shape of the areas. As an example, if an area is "abnormally" elongated, then we may conclude that the assumptions have not restricted the area sufficiently.

We cannot exclude either that *several solutions for shape and position of the areas are obtained*, which all turn out to be equally optimal. If such a case occurs and the areas corresponding to the different solutions are not overlapping, then we have *a set of (partly) disjoint solutions*, from which we must choose, which again necessitates some additional assumption.*

1. Multiple, disjoint solutions
2. Unsatisfactory shapes of areas

In fact, one reason why multiple solutions are likely to be obtained was elaborated earlier in this chapter: there are two possible sequences of criteria for optimality which cannot be distinguished inside the formal structure. (Pro memoria: the sequence *cdba* favours the correct classification of the reference chemicals in $Hset$, whereas the sequence *cdab* favours the correct classification of the reference chemicals in *Rset* (Chapter 7).) Against the background of Problem 2, we would expect the result from the two optimisation sequences to coincide at least in part. There is no reason why this should be the case, however, from a logical point of view, and it is therefore likely that an additional assumption of the more political type is required, selecting either *cdab* over *cdba* or the converse.

---

*From the analysis of the deterministic algorithm in Chapter 8, we may expect that several optimal solutions can be possible. This is the case when two of the solutions are concomitantly optimal after application of Criterion b, and remain so after application of Criterion a, too. Note that all possible solutions are optimal after Criterion c and Criterion d, since the solutions cannot be compared to each other with respect to correctly placed chemicals: none of the solutions contains chemicals that are incorrectly classified.

Chapter 8

Deterministic Calibration

As shown at the end of the previous chapter, the calibration problem can be restated as finding points P_2/P_4, which yield an *optimal* classification of the reference sets $Hset$ and *Rset*. *Optimality* is defined as maximal overlap of the classification (i.e. linear ordering) of the reference chemicals and the apriori reference ordering of these chemicals. The comparison is based on the two possible and meaningful sequences of the criteria introduced in Chapter 7: *cdab* and *cdba*.

The calibration of the classification procedure is performed with a *deterministic-geometrical* method. The method (i.e. the algorithm) is deterministic as it yields all possible solutions to the problem, and geometrical as the algorithm is geometrical.

For illustration purposes, the working of the algorithm is shown on an example of calibration with very simple reference sets $Hset$ and *Rset* (the reference sets contain only few elements), and the calibration is conducted with respect to the sequence of criteria *cdba*. The working of the algorithm is shown step by step for each criterion. The actual *code* of the algorithm is shown in Appendix J.

8.1 Minimum of False Negatives

Criterion c requires that no element of $Hset$ be classified as *high*. It follows from this requirement that the point P_2 and P_4 cannot lie in the *area below* any elements of $Hset$, since else such an element of $Hset$ would be placed to upper right hand of P_2 or P_4. This, in turn, would yield the classification *high* for that element. (By below we understand the area defined by all points in the Cartesian plane that are smaller than the element.) But for many elements of $Hset$ it is not necessary to observe this requirement: if an element is smaller than another element in $Hset$ (such as is the case for h_1 in Figure 8.1) then the area below this element is included in the area below the larger element (element h_2 in Figure 8.1). As a consequence *only the maximal elements of* $Hset$ need to be considered.

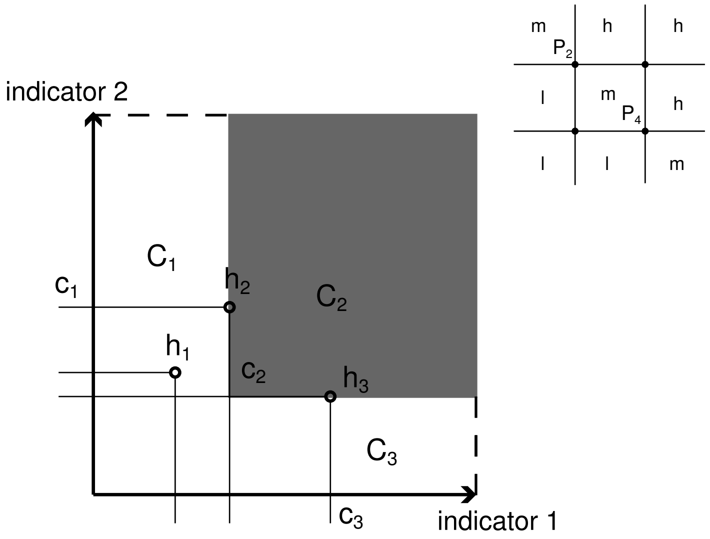

Figure 8.1: Example of three elements of $Hset$ (elements h_1, h_2, h_3). In order to fulfill Criterion c the points P_2 and P_4 may not be smaller than any element of $Hset$. Element h_1 itself is smaller than element h_2, hence it suffices to ensure that P_2 and P_4 don't lie below element h_2. The corner points c_1, c_2 and c_3 define the permissible areas C_1, C_2, C_3 for P_2 and P_4 under Criterion c. The permissible area C_2 is shown in grey for illustration purposes.

As illustrated in Figure 8.1, Criterion c is fulfilled if the placement of either of the points P_2 and P_4 is restricted to the Cartesian plane minus the areas that contain points strictly smaller than the maximal elements of $Hset$. (In the Figure the maximal elements are the elements h_2 and h_3.) The resulting *permissible area under Criterion* c is defined by the union $C_1 \cup C_2 \cup C_3$ where the areas C_i are defined through their bottom left corner points c_i (Figure 8.1). Note that the upper corner of the area is not defined by the data, i.e. the area is unbound above. In the figure, the area is artificially bound for illustration purposes.

Since both points P_2 and P_4 are to be placed in these areas, any *solution* of the calibration under Criterion C (any positioning of the two points in the permissible area) can be described by two areas out of C_1, C_2, C_3. E.g. P_2 in C_1, P_4 in C_3. This engenders redundancy in some cases, as is the case when a point is in C_1 and in

C_2 (when it lies within the intersection of the two areas). In the following an area which is permissible for P_2 is denominated as α_2 and one which is permissible for P_4 as α_4.

area\solution	1	2	3	4	5	6
α_2	C_1	C_1	C_1	C_2	C_2	C_3
α_4	C_1	C_2	C_3	C_2	C_3	C_3

Table 8.1: Six solutions numbered $1-6$ each contain two areas: α_2 (for P_2) and α_4 (for P_4). The solutions are formed by the random combination of the permissible areas C_1, C_2 and C_3.

For Figure 8.1 solution 1 is given by the combination of the permissible area C_1 (for P_2) with itself (for P_4), while solution 2 is given by C_1 (for P_2) and C_2 (for P_4).

8.2 Minimum of False Positives

Criterion d requires that no element of *Rset* be classified as *low*. The points P_2 and P_4 both define the areas *low* as the areas containing all points that are strictly smaller than themselves. As a consequence, in order to fulfill Criterion d it must be avoided that an element of *Rset* lies in these areas. This in turn requires that the points P_2 and P_4 do not lie in the area containing points strictly larger than any element of *Rset*. Again, we need only consider the extremal elements of *Rset* (in this case the set of minimal elements of *Rset*), since if the criterion holds for these elements then it will hold for any element of *Rset*.

In Figure 8.2, the areas D_1, D_2 and D_3 are defined by the points r_1 and r_3, the minimal elements of *Rset*. The element r_2 is in the area than contains points larger than r_1 and hence doesn't need to be taken into account during application of Criterion d.

In analogy to the application of Criterion c, the solutions for the points P_2 and P_4 are given by all possible combinations of the permissible areas D_1, D_2 and D_3. The 6 solutions that were obtained for Criterion c are all adjusted for Criterion d. All solutions are still equivalent (i.e. under Criteria d and c).

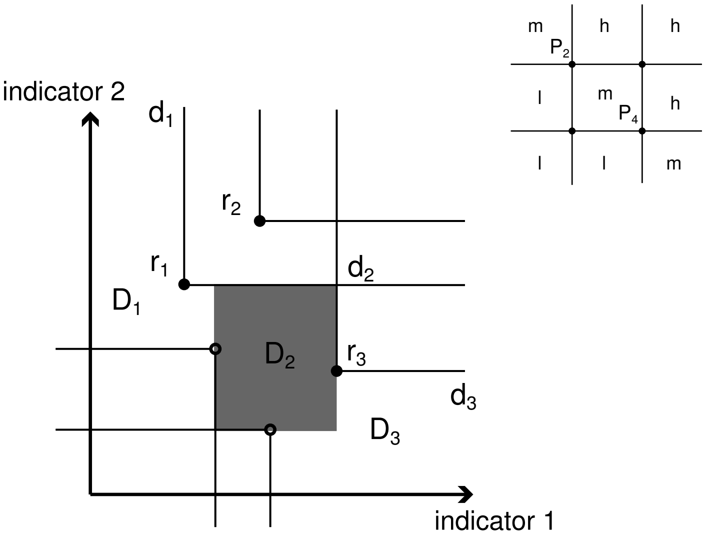

Table 8.2: Example of three elements of *Rset* (r_1, r_2, r_3). In order to fulfill Criterion d, the points may not lie in areas containing elements that are strictly larger than the minimal elements of *Rset* (r_2, r_3). These D_1, D_2, D_3 are defined by the corner points d_1, d_2 and d_3. For illustration purposes, only area D_2 is shown in color.

area\solution	1	2	3	4	5	6
α_2	D_1	D_1	D_1	D_2	D_2	D_3
α_4	D_1	D_2	D_3	D_2	D_3	D_3

Table 8.3: Six solutions numbered $1-6$ each contain two areas: α_2 (for P_2) and α_4 (for P_4). The solutions are formed by the random combination of the permissible, adjusted areas D_1, D_2 and D_3.

Non-rectangular Solutions
Note that the area D_i need not be rectangular. Indeed the areas is not rectangular when the corresponding area C_i contains more than one element of *Rset*.

The examples for illustration of Criterion b and a are of the simpler, rectangular case. The algorithm, however, is capable of calculating the more complex geometries, too.

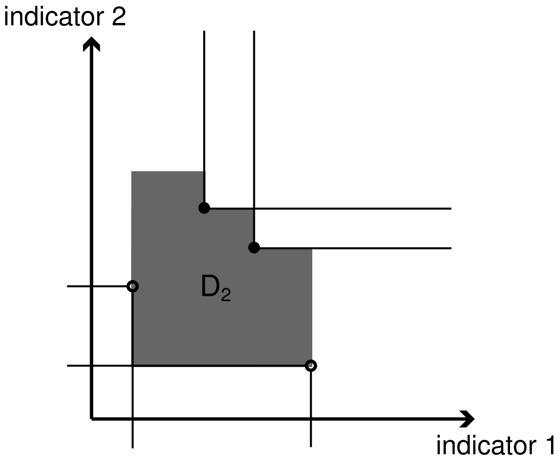

Figure 8.2: Area D_i which is not rectangular: The number of elements of $Rset$ that are in the area larger than the lower left corner, is larger than one. As a result the upper right corner is not given by a single point, but by two points.

8.3 Maximum of Correct Negatives

After having ensured that no element from $Hset$ is classified as *high*, the second step consists in optimising the calibration by classifying as many elements in $Hset$ as possible in the class *low*, with a minimum placed in *medium*. Again, we can use the points P_2 and P_4 as they define the class *low* (independently) and *medium* (together).

In a first step, the algorithm must define the areas which — inside one solution identified during application of Criterion d and c — classify as many elements as possible as *low*. The possibility to adjust for these elements is naturally limited to the elements, the classification of which is influenced by the restriction of one particular area. Hence it can be expected that after each solution is optimised for the Criterion, some solutions have a higher number of elements that are classified as *low* than other solutions; simply because more elements are affected by the restriction of one solution than are affected from another solution.

Then, having adjusted all solutions obtained during application of Criteria c and d, we must choose the solution(s) which *overall*, i.e. in comparison to the other solutions, classify a maximum number of elements from $Hset$ in the class *low*.

8.3.1 First Step: Area Adjustements

Coming back to our example, we look at the optimisation under Criterion b of two out of the six solutions in Table 8.3: the solutions 2 and 1. (Pro memoria: the label α_2 and α_4 indicate the areas that are permissible for the points P_2 and P_4, respectively.) We shall start with solution 2 and adjust the areas $\alpha_2 = D_1$ and $\alpha_4 = D_2$, which we will discuss in detail in the following. (The calibration of solution 1 will be shown at the end of this Section.) For each of these two areas the algorithm must find the restricted area for which a maximum number of elements of $Hset$ are classified as *low*.

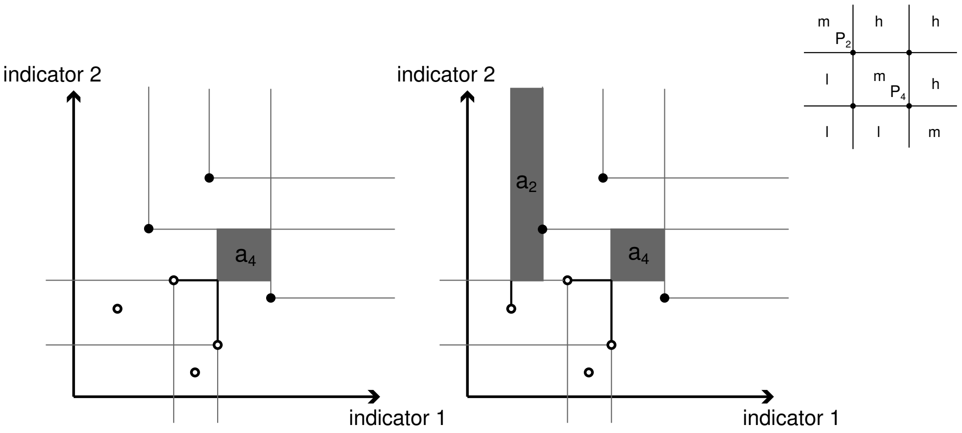

Figure 8.3: Illustration of algorithm on solution 2 (D_1, D_2); adjustments for α_4, then α_2. On the left, the D_2 area is optimised with respect to Criterion b, in order to accomodate P_4 (the α_4-area). The same is subsequently done for the D_1 area, which is designated as α_2 since it hosts point P_2.

In total there are four elements in $Hset$. With the adjustments that have been made to the permissible areas, all four elements are now correctly classified as *low*, and none is classified as *medium*.

There is one observation to make: the adjustment of the α_2 region in the x-direction is not necessary, since the element for which the adjustment is made is already classified correctly. Put differently, the area α_2 is unnecessarily restricted, and no restriction should apply to α_2 in the present example.

Manifestly, and in contrast to Criteria c and d, the sequence of the adjustment plays a role for the operationalisation of Criterion b. We may then ask what occurs if instead of first adjusting α_4 and then α_2, the algorithm does the converse. The

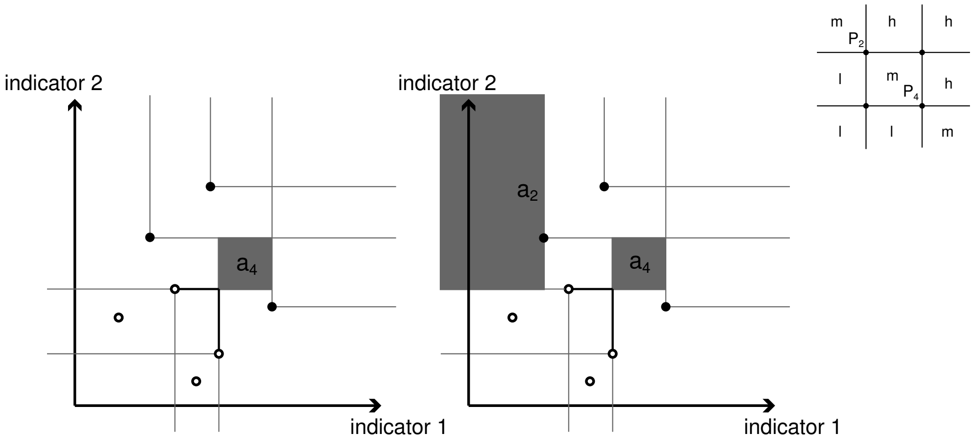

Figure 8.4: Illustration of algorithm on solution 2: corrected for redundancy. α_4 classifies all elements of $Hset$, and hence α_2 need not be restricted.

situation then presents itself differently (Figure 8.5).

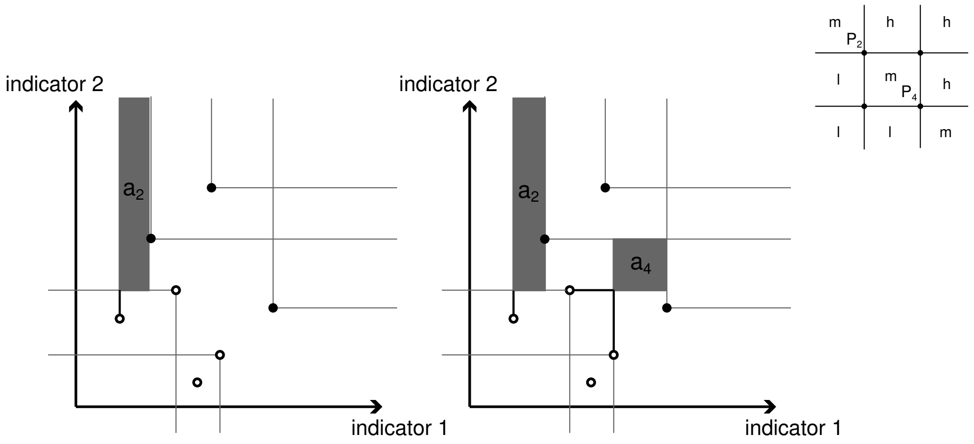

Figure 8.5: Compared to Figures 8.3 and 8.4 the sequence of adjustment is inverted: first adjustment to α_2, then adjustment to α_4. The result is not identical with that obtained for the first sequence (Figure 8.4).

As a consequence, for each of six solutions identified through application of Criteria d and c, the calibration under Criterion b must be conducted twice: once starting with adjustment to α_2 and finishing with adjustment to α_4, the second time opting for the opposing sequence.

In analogy to the adjustment to solution 2, the areas α_2 and α_4 can be adjusted for solution 1 (D_1, D_1). Here, too, the adjustments can be made by first starting with

α_4 and then proceeding with α_2 (Figure 8.6), or the converse (Figure 8.7).

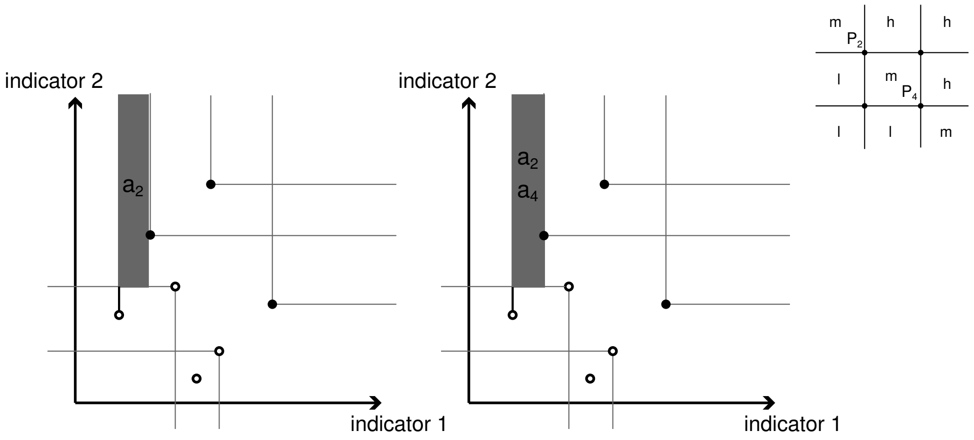

Figure 8.6: Result of optimisation of solution 1 (D_1, D_1) under Criterion d. There are again two possible sequences for the adjustments. The first is depicted in the Figure: adjustements to α_4, then to α_2.

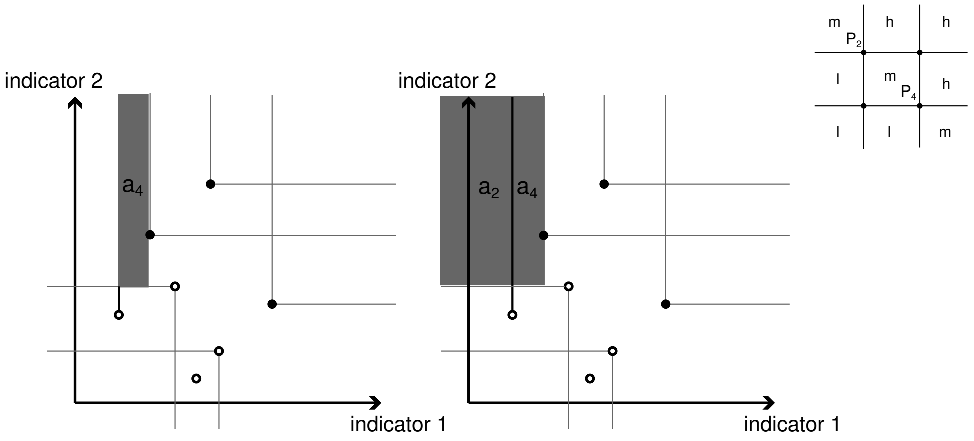

Figure 8.7: The second sequence of adjustments to solution 1 (D_1, D_1), first adjustements to α_2, then to α_4. As in the adjustments to solution 2, the result of the adjustment is not identical to the result obtained in Figure 8.6.

8.3.2 Second Step: Selection of Overall Optimum

We have so far only made the adjustments that were possible for the different solutions, but have not assessed the solutions with respect to how well they perform *overall* under Criterion b. In order to illustrate that second stage of the algorithm for Criterion *b*, we compare the result obtained for the solution 2 (D_1, D_2) with that obtained for solution 1 (D_1, D_1).

Solution 2 classifies all 4 elements in *Hset* as *low*. Solution 1, on the other hand, classifies only one such element as *low*, whereas the remaining three are classified as *medium*. Regardless of the sequence of adjustment, solution 2 (D_1, D_2) has a higher number of elements from *Hset* classified in *low* than solution 1 (D_1, D_1). Accordingly, under Criterion *b* the algorithm chooses solution 2 over solution 1. Nevertheless, under Criterion *b* we now have two sub-solutions of solution 2, which are equivalent with respect to Criterion *b*.

We can now proceed with the result from solution 2. The two sub-results of that solution are equally optimal, so we may continue with both in parallel. Note, however, that the solution resulting from the sequence of adjustement α_4, α_2 contains the solution resulting from the sequence α_2, α_4, and as a consequence we may discard the latter solution.

8.4 Maximum of Correct Positives

Criterion *a* is equivalent to Criterion *b*, only the reference set concerned changes. Here we try to maximise the number of elements from *Rset* that are correctly classified as *high*, rather than *medium*. The algorithm operates analogously to the algorithm used for Criterion *b*: duplicate each solution from Criterion *b* (i.e. actually each sub-solution) to allow for the two sequences in the adjustment, then run the optimisation for each (sub-)solution and pick the solution with the highest number of elements from *Rset* classified as *high*.

In a last step, the adjusted areas for solution 2 can now be compared to all other solutions which have been selected after passing Criterion *b*. The solution which, *overall*, has the highest number of elements in *Rset* classified as *high*, is then selected as the optimum under Criterion *b*. It may not be excluded that more than one solution exhibit that highest value (no example shown).

The solutions that are optimal under Criterion *b* represent the overall optimum under the sequence of Criteria c, d, b, a.

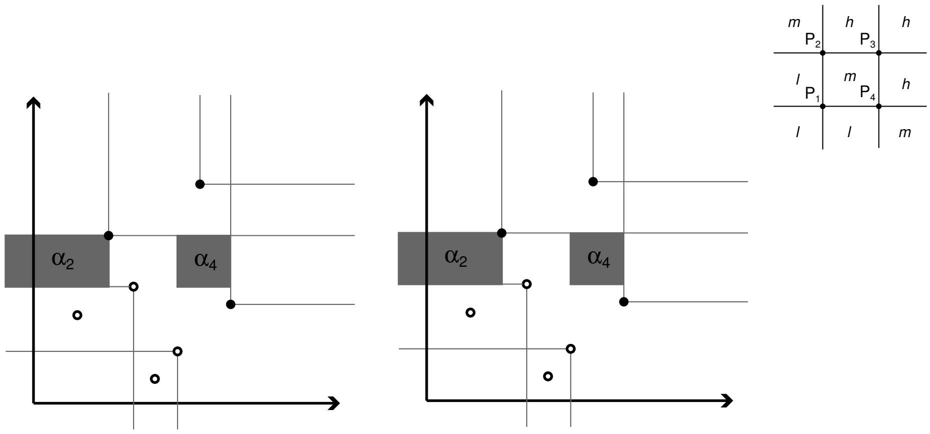

Figure 8.8: Adjustments made to solution 2 (D_1, D_2) under Criterion a. The first sequence of adjustments depicted here is: first adjustements to α_2, then to α_4.

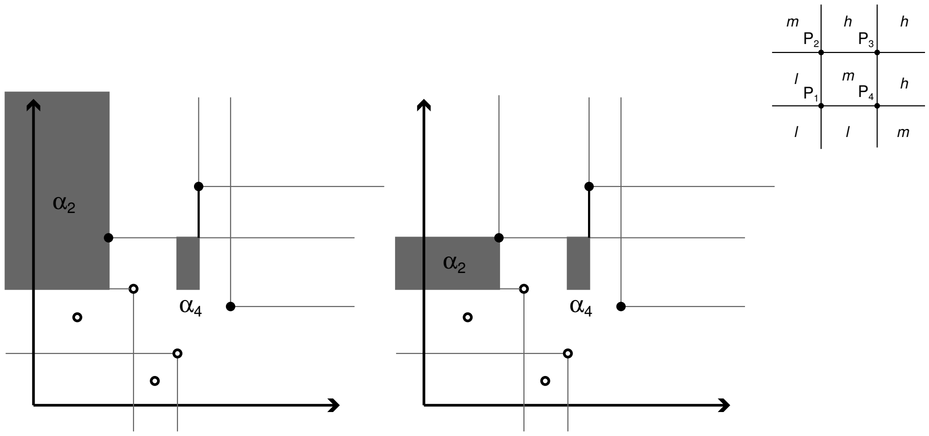

Figure 8.9: The second sequence of adjustments to solution 2 (D_1, D_2) — first adjustements to α_4, then to α_2 — as depicted here, is not identical with the result shown in 8.8.

Chapter 9

Calibration Results

Requirement 4 states that for each of the threat scenarios the classification procedure must yield linear orderings. The requirement is presently fulfilled, and the results obtained for the optimisation (optimal calibration) of the classification procedure are discussed for each scenario separately.

The separate discussion of the results obtained for the two sequences *cdab* and *cdba* is conducted for each threat scenario, i.e. both for *Pandora* and for *Bioaccumulation*. For each scenario the results from the two sequences are then compared. If no solutions from either sequence overlap,[*] then further investigation of the areas corresponding to the two sequences are conducted. Finally the result from calibration (which may have necessitated additional assumptions) is discussed and the discrete values are picked to define the discretisation and hence, ultimately, the assessment procedure. We will conclude with a few general observations.

9.1 *Pandora*

The threat scenario *Pandora* is built from two parameters, the characteristic persistence τ and the characteristic spatial range ρ.[†] The reference chemicals in $Hset$ are marked as black framed grey squares in the following, and the reference chemicals in *Rset* as black squares. The scenario is depicted in the Cartesian plane spanned by the two dimensions that define the senario, in Figure I.1 in Appendix I.

---

[*] i.e. a solution from one sequence overlaps with a solution from the other sequence

[†] See Chapter 2

Results cdba

The areas α_2 and α_4 obtained for the points P_2 and P_4 are shown in grey in the plot in Figure 9.1. In order to characterise the areas α_2 and α_4, it suffices to use the coordinates of two points: namely the lower left corner and the upper right corner of the areas. The corresponding coordinates for the areas α_2 and α_4 obtained for the points P_2 and P_4, are given in Table 9.1. Contrary to what can be expected in the general case, there is a single solution for the areas α_2 and α_4.

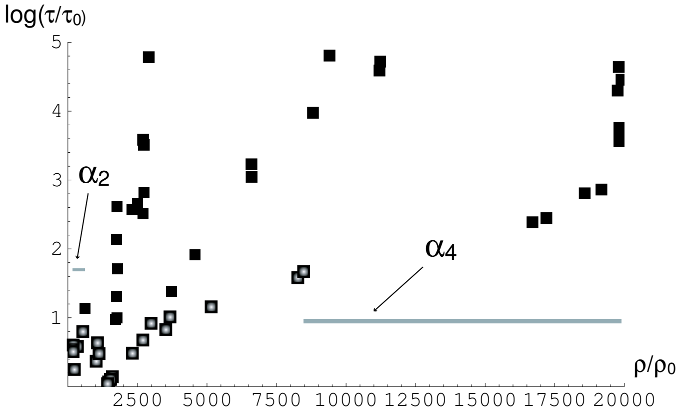

Figure 9.1: Solution from the geometrical algorithm for the areas α_2 and α_4 are shown in grey. The area on the left is α_2, whereas the area on the right is α_4.

area	coordinates
α_2	$(173, 47) - (620, 52)$
α_4	$(8474, 8.3) - (19900, 9.5)$

Table 9.1: Areas α_2 and α_4 inside of which calibration of *Pandora* is optimal

For discretisation of the filters we require values for the boundaries *low/medium* and *medium/high* for each parameter. In order to obtain these, the data in Table 9.1 can be re-organised so as to give the value ranges of the parameters, inside of which lie the boundaries *low/medium* and *medium/high*, for a particular parameter (Table 9.2).

Results cdab

We now turn to the sequence of the criteria that puts more importance on the correct classification of the reference chemicals in *Hset*: the sequence *cdab*. Again, we may expect several solutions from the calibration, as the same argument applies

parameter	*low/medium*	*medium/high*
persistence τ	$8.3 - 9.5d$	$47 - 52d$
spatial range ρ	$173 - 620km$	$8474 - 19900km$

Table 9.2: Ranges for the boundaries *low/medium* and *medium/high* for the two parameters, inside of which calibration is optimal for *Pandora*.

to the algorithm that operationalises the sequence *cdab* then does to the algorithm operationalising the sequence *cdba*.

Surprisingly, the calibration resulting from optimisation based on the sequence *cdab* yields the *exact same areas* as are obtained from the sequence *cdba*. This may be viewed as even more particular than the fact that there is, in this ubiquitous solution,[‡] only a single solution, instead of a set of equivalent solutions. In any case, the first of the problems that were identified as possible anomalies that lead to a revisiting of the assumptions, can be eliminated: there are not several, disjoint solutions to the calibration problem.

The comparison of the calibration with *cdab* and *cdba* would now be visualised by superimposing the areas obtained for *cdab* with the areas obtained for *cdba*. As the areas obtained for the two sequences are identical, the overlap is total (see Figure I.3 in Appendix I). Hence the result of the calibration of the filter representing threat scenario *Pandora* are the areas α_2 and α_4 as depicted in Figure 9.1, and no further restriction is possible at this point.

9.1.1 Discussion

From inspection of the results in Table 9.2, we note that the proxy-measure for the persistence τ has well-defined ranges for the cut-off of the discrete ranges *low*, *medium* and *high*: the ranges span over a period of a few days. Also, we note that the ranges are bounded from below and above.[§]

The proxy measure for the spatial range ρ, on the other hand, yields areas that are less well defined, in the sense that the ranges for the boundaries are much wider.

[‡] in the sense that the solution is identical for both sequences

[§] By bounded from below is meant that the lower bound of the range *low/medium* is larger than the *minimum* of τ in $Hset$ (or Rset). Bounded from above means the equivalent for the higher bound of the range *medium/high*.

Indeed, the range permissible for the boundary *medium/high* stretches over more than half the value-range of the spatial range (i.e. of the circumference of the earth).[¶] It is also noteworthy that the lowest possible value for the *low/medium* boundary for the spatial range is the *overall minimal value* for that parameters in the reference set. (The upper limit for the *medium/high*-boundary for ρ is also the *overall maximal value* for the parameter). In short, neither is the range for the boundary *low/medium* bounded from below, nor is the range for the boundary *medium/high* bounded from above (except by the beginning and the end of entire value, respectively).

As a result, we can single out three items for the discussion of the calibration results:

1. Range for ρ-boundary *low/medium* is not bounded from below
2. Range for ρ-boundary *medium/high* is not bounded from above
3. Range for ρ-boundary *medium/high* is "long"

All points listed relate to Point 1 that is singled out in Chapter 7. Bullet 1 in the list shown above in conjunction with Bullet 2 from the same list could be combined to imply a more dramatic statement: the parameter ρ is irrelevant with respect to calibration.

To see why, assume the grid is chosen precisely in such a way as to place the *low/medium*-boundary at the lowest possible ρ-value and the boundary *medium/high* at the highest possible value. Then we have an optimal solution for which we have *medium* as the value for ρ over all possible chemicals that are in the two reference sets $Hset$ and $Rset$. As a consequence, there is an optimal solution for the calibration that yields a discretisation of the threat indicator ρ *that yields a single interval for* ρ.

It may be tempting to conclude from this observation that the parameter ρ adds nothing to the description of the threat scenario, and the same calibration would have been obtained if only τ had been used in the threat scenario *Pandora*. Indeed, we may come to the conclusion somewhat prematurely that the threat scenario as

---

[¶] This statement is mathematical not really meaningful, since "half the value-range" is not invariant under transformation such as *log* which we use for other parameters, e.g. for τ in the same scenario. In this particular case, it has nonetheless a real-world, namely geographical meaning.

such or at least in its present modelling is not meaningful. This amounts to question Assumption 1. In order to clarify this critical situation, the bullet points 2 and 3 are discussed concomitantly, whereas Bullet 1 is discussed separately.

1. low/medium range for ρ
From theoretical considerations we know that the spatial range ρ is not irrelevant. We simply take into consideration the extremal case of highly-persistent, highly immobile chemicals such as heavy metals or plastics and cement, to convince ourselves of the contrary: Such "chemicals" have very high persistence, but nevertheless do not constitute a high threat, due to their limited range. If such chemicals had been included in $Hset$, then the calibration of $Pandora$ would not be the same if ρ was omitted. Hence, in fact, the parameter ρ is of relevance for the scenario $Pandora$. Interestingly, for the sequence *cdba* inclusion of highly immobile chemicals with long persistence would alter the area α_2 significantly, whereas in the case of sequence *cdab* the area is not altered.

2./3. medium/high range for ρ
Nevertheless, the fact that the range for ρ is not bounded from above (Bullet 2) and spans such a wide range of values for ρ (Bullet 3), is not very satisfactory considering that the calibration is performed in order to find discrete value. These two properties may indicate that indeed Problem 2 can not be answered positively, i.e. the discretisation of the *Pandora* filter *cannot* be performed from existing regulations alone.

The question immediately arises, whether this under-determination stems from the (accidental) selection of reference chemicals, or whether the unsatisfactory shape of α_4 has more fundamental causes. The latter reason would conjure the question whether the threat scenario $Pandora$ is a meaningful one to use, at least in its present definition and with the use of the present calibration function. We hence question Assumption 4 and wonder if a better result could have been obtained with different reference sets.

The shape of the area α_4 arises form the fact that the area to the right of the elements of $Hset$ (see Figure 9.1) is empty, i.e. not populated by elements of $Rset$. Were it populated, the Criterion a would force the area to be bounded from above. The void in the said region may be accidental fact in the sense that it reflects the particular choice of $Rset$.

It may also be, however, that the missing of the upper boundary for the range is due to the scientific impossibility to find a chemical in that region of the scenario. If this region does not necessarily need to be void, then we may revisit the reference set *Rset*, and add a substance that would allow for the range to be bounded from above.

In order to answer this question we can take advantage of a peculiarity of the *Pandora* scenario: the parameter ρ is a function of τ (from [73], see Chapter 7).

$$\rho = e\sqrt{\frac{D}{k}}\tanh\left[\pi r\sqrt{\frac{k}{D}}\right] e^{\frac{\frac{\pi}{2}-\arctan\left[e^{\pi r\sqrt{\frac{k}{D}}}\right]}{\sinh\left[\pi r\sqrt{\frac{k}{D}}\right]}} \quad ; \quad k = \frac{ln2}{\tau} \tag{9.1}$$

Hence we need just look at $\tau(\rho)$. If the curve cuts the horizontal line at the maximal value for τ found for any element of $Hset$ ($Max_\tau(Hset)$), then the area need not be empty of elements of *Rset*. If, on the other hand, the curve $\tau(\rho)$ does not cut the horizontal line, then the area is always empty, and the range is never bounded from above.

The function ρ depends on the *eddy* diffusion constant, which depends on the media of transport. In order to obtain the minimal value of ρ for a given τ, we must use the maximal diffusion constant, which is the constant for diffusion in the air: $D = D_{air}$.

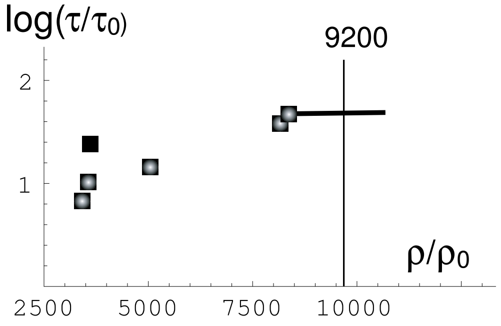

Figure 9.2: Horizontal line at $Max_{tau}(Hset)$ is intersected with the minimal function $\rho(\tau)$, i.e. with $D = D_{air}$. The intersection occurs at $9200km$.

We cannot solve Eqn. 9.1 for τ. But it suffices to insert the value $Max_{tau}(Hset)$. We find that the curve is cut at $9200km$, as shown in Figure 9.2. This implies that there exists an area where elements of *Rset* could be found.

At least this establishes that the anomaly for ρ is not systemic, but could, in theory, be broken if a chemical was in the appropriate area of the plane.

parameter	$low/medium$	$medium/high$
persistence τ	$8.3 - 9.5d$	$47 - 52d$
spatial range ρ	$173 - 620km$	$8474 - 19900km$

Table 9.3: Ranges for the boundaries $low/medium$ and $medium/high$ for the two parameters, inside of which calibration is optimal for $Pandora$.

Further Restrictions
It may be argued, in other words, that the areas α_2 and α_4 we have obtained from both sequences *cdba* and *cdab* indicate that an empirical calibration is not possible and that hence additional, more "political" input may be required in order to obtain a calibration and hence a calibration procedure.

We have noticed that Bullet 1 can be solved by introducing additional chemicals in $Hset$, although this is only possible for the sequence *cdab*. But this is the less pressing of the anomalies: Bullet 2 is more critical.

In order to further restrict, we can, at present only use the previous result: that of the intersection of the curve $\rho(D_{air}, \tau)$ with the horizontal line given by $Max_\tau(Hset)$. This amounts to the statement that we expect the existence of a chemical that has maximal *eddy* diffusion rate and that has a persistence τ of approximately 50 days. This assumption is, from a scientific point of view, not unrealistic. But note that nevertheless, at least one of the assumptions must be questioned, namely Assumption 4.

With this modification, the range for $medium/high$ for ρ is, as consequence, bounded from above.

parameter	$low/medium$	$medium/high$
persistence τ	$8.3 - 9.5d$	$47 - 52d$
spatial range ρ	$173 - 620km$	$8474 - 9200km$

Table 9.4: Ranges for the boundaries $low/medium$ and $medium/high$ for the two parameters, inside of which calibration is optimal for $Pandora$.

9.1.2 Definition of Grid

The areas can only form a *frame* from which the actual value for the boundaries *low/medium* and *medium/high* can be chosen. With the additional assumption made on the hypothetical existence of a reference chemical in $Hset$, we have areas to choose discrete values from, which stand the test of the *unsatisfactory features of areas*, as established at the end of Chapter 7.

The choice is made from within the area and we make the completely arbitrary choice of the median value for each boundary range and each parameter. As a result we obtain the discrete values listed in Table 9.5.

parameter	*low/medium*	*medium/high*
persistence τ	$9d$	$50d$
spatial range ρ	$400km$	$8800km$

Table 9.5: Discrete values for the boundaries *low/medium* and *medium/high* for the two parameters, inside of which calibration is optimal for *Pandora* and which are chosen for the discretisation of the filter .

9.1.3 Summary

The following summary of the discretisation of *Pandora* includes observations as well as additional assumptions that were made, in order to obtain suitable discrete values for the corresponding filter.

- The two sequences *cdba* and *cdab* yield the same results.
- The range *low/medium* for ρ is not bounded from below. This indicates that Assumption 1 to 4 do not suffice to define the filter. The problem can be remedied by considering additional elements for $Hset$.
- The range *medium/high* for ρ is not bounded from above. The problem can be remedied by introducing a *hypothetical chemical* to *Rset*. The corresponding τ- and ρ-values are scientifically reasonable.
- If the horizontal line had not been intersected than it would have been scientifically impossible to find a chemical in the said region, the solution may

have to be dismissed, and the assumptions made earlier revisited in a more involved way.

It should not be missed out that the current modifications to the assumptions are all related to Assumption 4. Better results may also be obtained by modifying some of the other assumptions. Nevertheless, it is shown that the filter *can* be reasonably defined by undertaking limited modifications to *Rset*. It also appears to indicate that the sequence *cdab* is somewhat more suited to the current problem.[‖]

9.2 *Bioaccumulation*

The threat scenario *Bioaccumulation* is defined by two parameters, in this case the persistence τ and the octanol-water distribution coefficient K_{ow}, which is commonly used as a measure for lipophilicity. The reference chemicals in $Hset$ are again shown in grey (with black frame), while the reference chemicals in *Rset* are shown in black.

Results of cdba

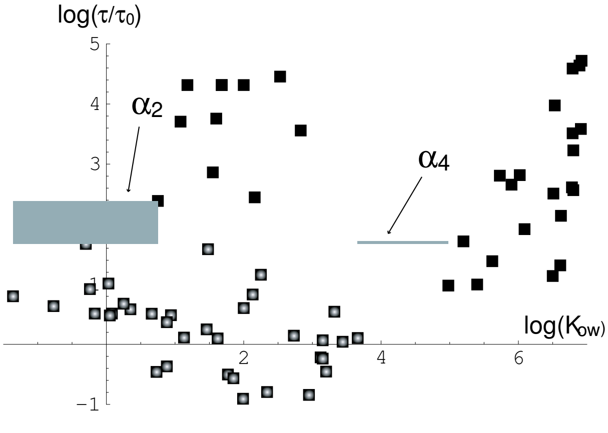

Figure 9.3: Calibration of *Bioaccumulation* with sequence *cdba*. Solution from the geometrical algorithm is shown as the areas α_2 and α_4, the areas permissible for the points P_2 and P_4, respectively.

Here, as in the section on the results for *Pandora*, the two areas permissible for the two points P_2 and P_4 (in order to yield optimal results), are given by the lower left

---

[‖] This statement can be made based on the fact that under sequence *cdab*, the calibration does not change if $Hset$ is enlarged by highly persistent, highly immobile chemicals, whereas this is not the case for the sequence *cdba*.

corners and the higher right corners of the areas. We can again re-organise the data in order to give the boundaries that define *low*, *medium* and *high* values for the two constituting parameters τ and K_{ow} (Table 9.7).

area	coordinates
α_2	$(-1.36, 47) - (0.75, 240)$
α_4	$(3.66, 47) - (4.98, 52)$

parameter	*low/medium*	*medium/high*
persistence τ	$47 - 52d$	$47 - 240d$
$logK_{ow}$	$(-1.36) - 0.75$	$3.66 - 4.98$

Table 9.6: *Above*: The areas α_1 and α_2 which define equivalent optimal solutions with respect to the sequence *cdba*. *Below*: Ranges for boundaries between the discrete intervals *low*, *medium* and *high*, for the two parameters τ and $logK_{ow}$

Results of cdab

As before we next look at the calibration obtained through the sequence *cdab*, i.e. by using a combination of the criteria a, b, c, d that is more in favour of the correct classification of the elements in *Rset*.

As recalled at the beginning of this chapter, the algorithm can yield several, distinct sets of results for optimal calibration. As shown in Figure 9.4, this is the case for the sequence *cdab* applied on the scenario *Bioaccumulation*.

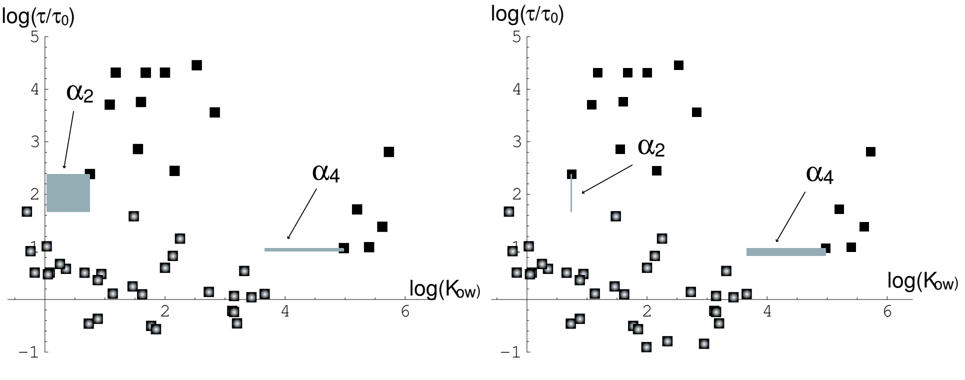

Figure 9.4: Calibration of *Bioaccumulation* with sequence *cdab*. Two distinct solutions from the geometrical algorithm are shown. Note that the areas α_2 and α_4 for the two solutions overlap.

We can again list the boundaries that define *low*, *medium* and *high* values for the two constituting parameters τ and K_{ow}. We need not concern ourselves with the meaning of this two-fold result. For the time being we just re-write the coordinates as the corresponding value ranges and postpone the discussion.

area	coordinates: solution 1	coordinates: solution 2
α_2 for P_2	$(0.03, 47) - (0.75, 243)$	$(0.73, 47) - (0.75, 243)$
α_4 for P_4	$(3.66, 8.3) - (4.98, 9.0)$	$(3.66, 6.8) - (4.98, 9.5)$

	solution1		solution2	
parameter	*low/med*	*med/high*	*low/med*	*med/high*
pers. τ	$8.3 - 9.0d$	$47 - 243d$	$6.8 - 9.5d$	$47 - 243d$
$logK_{ow}$	$0.03 - 0.75$	$3.66 - 4.98$	$0.73 - 0.75$	$3.66 - 4.98$

Table 9.7: *Above:* The areas α_1 and α_2 which define equivalent optimal solutions with respect to the sequence *cdba*. *Below:* Ranges for boundaries between the discrete intervals *low*, *medium* and *high*, for the two parameters τ and $logK_{ow}$.

cdab vs. cdba

As discussed in the introductory part of this chapter, the formal structure of the assessment does not favour the sequence *cdab* over *cdba* or vice-versa. In the case of *Pandora* the results from the two sequences (luckily) coincided. In the case of *Bioaccumulation* this is not the case. At least the areas representing the two solutions for *cdab* coincide since else we would have to compare (and make a choice among) *three* possible solutions. But neither of the two solutions for *cdab* coincides with the results from sequence *cdba*. So in any case a choice between *cdba* and *cdab* *must* be made.

Before making this choice, we may first look more closely at the results obtained for the two sequences of criteria. This may provide indications on and ultimately a basis for an adequate choice of sequence.

Discussion cdba

A number of observations can again be made upon inspection of Figure 9.3. First, all ranges for boundaries are bounded from above. In the case of the parameter $logK_{ow}$, the boundary for *low/medium* is not bounded from below: the boundary can be exactly the minimal value for $logK_{ow}$ over all chemicals, in which case the

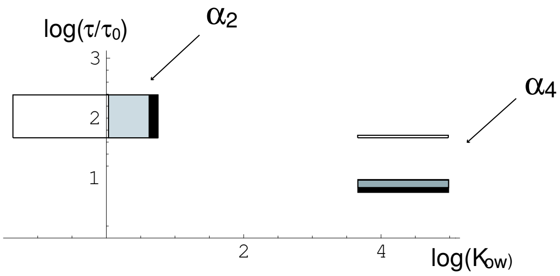

Figure 9.5: Solution from the geometrical algorithm for the areas α_2 and α_4 combined for the two sequences *cdba* and *cdab*. The solution from *cdba* is shown in white, the two solutions from *cdab* in grey and black. The reference points have been omitted in view of simplification.

interval *low* is not populated by any chemical of the two reference sets. As a second observation, we note that the areas α_1 and α_2 overlap in the y-direction. This means that for the parameter τ the boundary *low/medium* and *medium/high* can coincide, i.e. of the exact same value or *degenerate*. This is not without problems, since when such an coincidence occurs, the interval *medium* vanishes for τ and the classification becomes significantly more sensitive.

- Range for K_{ow}-boundary *low/medium* is not bounded from below
- The areas α_2 and α_4 overlap: boundaries for τ can be degenerate

With respect to the boundary *low/medium* for $logK_{ow}$, it can be shown that the calibration algorithm does not exclude the possibility of the range to be bounded from below. The problem of degeneracy, however, remains, is more involved and cannot be readily solved.

Discussion cdab
Inspection of the result and corresponding tables and figures, indicate three important features of the results

- There are two sets of results for the areas α_2 and α_4.
- In both sets, all ranges are bounded from above, and with the exception of the lower range for the paramter $log(K_{ow})$, all ranges are bounded from below.
- The range for the upper boundary of *medium/high* for τ appears large $(47 - 243d)$.

Point 2 does not appear to pose a problem. Also, the fact that two solutions are obtained, can be solved (Point 1). First we note that neither of the areas of Solution 2 are subsets of the areas of Solution 1, nor the converse. But the areas α_4 are very similar for the two solutions. For the area α_2, the solutions overlap (which must not necessarily be the case). Hence, we have the possibility of using a single solution to represent the result from *cdab*, by intersecting Solution 1 and 2.

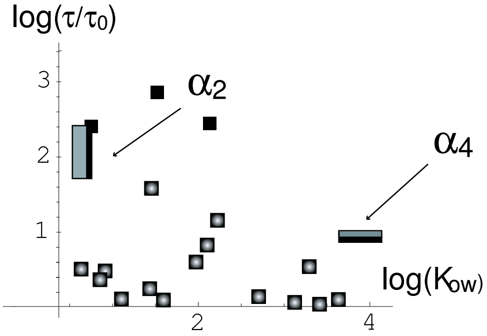

Figure 9.6: Combination of the two sets of results for the scenario *Bioaccumulation* and the sequence *cdab*. Solution 1 is shown in grey, and Solution 2 in black. The area α_4 for solution 1 is enlarged in the for illustration purposes.

Hence, we can use the coordinates of the intersection of the two solutions for further discussion. Remains Point 3, namely the fact the upper boundary of τ has a large interval.

parameter	*low/medium*	*medium/high*
persistence τ	$8.3 - 9.0d$	$47 - 243d$
$logK_{ow}$	$0.73 - 0.75$	$3.66 - 4.98$

Table 9.8: Intersection of the two sets of results from *cdab*: the value ranges thereby obtained are subsets of all results for *Bioaccumulation*

9.2.1 A Choice to be Made

Since the two solutions for *cdab* intersect, the results from the two sequences *cdba* and *cdab* can be directly compared, and chosen among. This, as we have seen, requires additional assumptions. In the spirit of the formal structure of the classification problem we wish, however, to make a minimum of political assumptions (which could be done in simply stating that a correct classification of the elements $Hset$ is preferable).

In order to attain a choice, structural features can be used, based on which the two solutions can be differentiated. One such feature is that (as noted in the discussion of *cdba*) in the case of *cdba* the boundaries for *low/medium* and *medium/high* can be degenerate, i.e. may coincide. This can lead to unsatisfactory situations where the classification of a chemical may be either *high* or *low*, just depending on a very small change in the value for τ. Although no sensitivity analysis has been performed at this point (it will follow in the next section), degeneracy can still be ruled objectionable. This can be taken as a first indication that the solution from *cdab* is better suited than the solution from *cdba*.

Independently from this consideration, the solution from *cdab* is characterised by an interesting finding: the boundaries of the ranges for τ are exactly the same than the ones found in *Pandora* (for both *cdba* and *cdab* since the solutions coincide there). It may appear difficult to attach a special meaning to the coincidence of the τ-boundaries. Nevertheless, we may recall Corrolary 3, which appears to be somewhat ackward now, and welcome the fact that with the solution from *cdab* the discretisation of the threat indicator τ is now identical for both threat scenarios *Pandora* and *Bioaccumulation*.

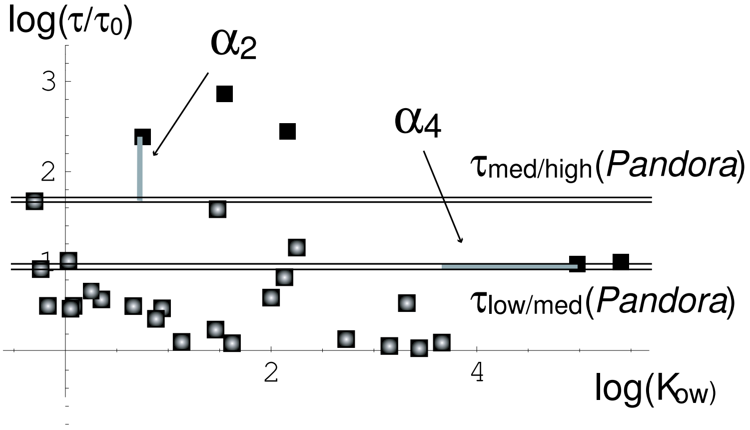

Figure 9.7: The boundaries for τ obtained for *Pandora* are shown in the plane for *Bioaccumulation*. The boundaries obtained for τ in *Pandora* cut the areas found for α_2 and α_4 in *Bioaccumulation*.

With both criteria (the other being the wish to avoid degeneracy) the favoured sequence is *cdab*.

If we only consider the argument against degeneracy, however, then Point 9.2 as mentioned above in the discussion of *cdab* remains unanswered: the range for the boundary of τ remains large. Hence we choose (or may choose) the boundaries from *Pandora* in order to narrow down the area.

Note that this does not infringe on the Requirement 2, as the requirement applies to the *classification of chemicals*, i.e. how the result of the classification is influenced by the sequence of application of the threat scenarios. If we use information on the discretisation of τ from another scenario, then this implies that we have performed the *calibration* of that other scenario earlier. As much as this can influence the classification of a chemical, it does not so on grounds of the sequence of application of (calibrated) filters, i.e. the asymmetric application of the filters occurs exactly once, when the filters are defined, and hence prior to any classification of chemicals.

9.2.2 Definition of Grid

Accordingly, we can use the data from Table 9.8 as the basis for the choice of the discrete values. Furthermore, the upper-limit of the τ-range from *Pandora* is used to restrict the area α_2.

No further restrictions of the areas is possible at this point, unless further assumptions are introduced, and hence it is appropriate at this point to choose the discrete values required for the calibration of the filter corresponding to the scenario *Bioaccumulation*.

The choice is made from within the area and we again make the completely arbitrary choice of the median value for each boundary range and each parameter. As a result we obtain the discrete values shown in Table 9.9.

parameter	*low/medium*	*medium/high*
persistence τ	$9d$	$50d$
$logK_{ow}$	0.74	4.32

Table 9.9: Discrete values for the boundaries *low/medium* and *medium/high* for the two parameters, inside of which calibration is optimal for *Bioaccumulation* and which are chosen for the discretisation of the filter .

The definition of the discretisation values for *Bioaccumulation* can be summarised in the following way:

- Two solutions are obtained from *cdab*, one from *cdba*
- The two solutions from *cdab* overlap but none of these overlap with the solution from *cdba*.
- The overlap of the solutions from *cdab* is chosen on grounds of avoiding degeneracy.
- The overlap of solutions from *cdab* is further narrowed down by using the boundaries for τ obtained in *Pandora*.

9.3 Sensitivity and Robustness

Before proceeding to a compendium of the results obtained for calibration, we must first assure ourselves that the results are *stable*. The results are derived from a reference, and as such the choice of reference influences the outcome.We, *must*, however, choose a reference in the present setting (or in any calibration for that matter) and it is somewhat absurd to require stability in case of major changes to the reference (i.e. by choosing completely different substances).

Beyond this fundamental influence, we may also question the stability of the results, inside of a given set of reference chemicals: the numerical data that is used may be questionable and if changed for some chemicals may lead to results different to the ones presently obtained. In this sense it is important to study the results under considerations of sensitivity and robustness. The former relates to changes in the *number of chemicals in reference sets*, whereas the latter concerns changes in the *values used for characterisation* of a given number of chemicals.

Data Sensitivity
The results for calibration are expressed as *two areas*. Each of the areas is reflected by two points, which gives a total of 4 coordinates per area and 8 coordinates per solution. Each of these coordinates is reflecting the value of some chemical (from either $Hset$ or *Rset*). Hence there are 8 chemicals to consider and discuss, and we can leave out the others for the time being. In order to discuss the 8 coordinates, the following nomenclature is adopted: $lb_{low/med}(\tau)$ for the $\underline{l}$ower $\underline{b}$oundary of the range *low/medium*, $ub_{low/med}(\tau)$ for the $\underline{u}$pper $\underline{b}$oundary of the range *low/medium*. Analogously $lb_{med/high}(\tau)$ for the $\underline{l}$ower $\underline{b}$oundary of the range *medium/high* and

$ub_{med/high}(\tau)$ for the upper boundary of the range *medium/high*. The corresponding coordinates for the parameter ρ are defined analogously, as is the case for the parameter $logK_{ow}$ in the case of the scenario *Bioaccumulation*.

9.3.1 *Pandora*

$lb_{low/med}(\tau)$: The lower boundary of the range *low/medium* for τ is defined by the element *acetone* ($\tau = 8d$). If the element has a smaller τ value, the boundary can drop down to $\tau = 7d$ (*benzene*), if the value is larger, it can increase up to $\tau = 10d$ (*propyleneoxide*). ⇒ sensitivity low

$ub_{low/med}(\tau)$: The upper boundary for the same range is give by the τ value of *dieldrin* ($\tau = 10d$). If the chemical takes a lower value, there is no bound, except $\tau = 0d$. Moving up, the value can go as high as $\tau = 10d$, i.e. the same value (up to rounding) for *heptachlor epoxide*. ⇒ sensitivity medium

$lb_{med/high}(\tau)$: The lower boundary for the range *low/medium* is bounded by the substance *ethylene oxide*. The value can move up, but not down, since it is close to *ethylene dichloride* with $\tau = 38d$. ⇒ sensitivity medium

$ub_{med/high}(\tau)$ The upper boundary is defined by *endrin* ($\tau = 52d$), and may drop no further than the τ-value for *PCB* 28 ($\tau = 24d$) and up to $\tau = 83d$ (*PCB* 52). ⇒ sensitivity low

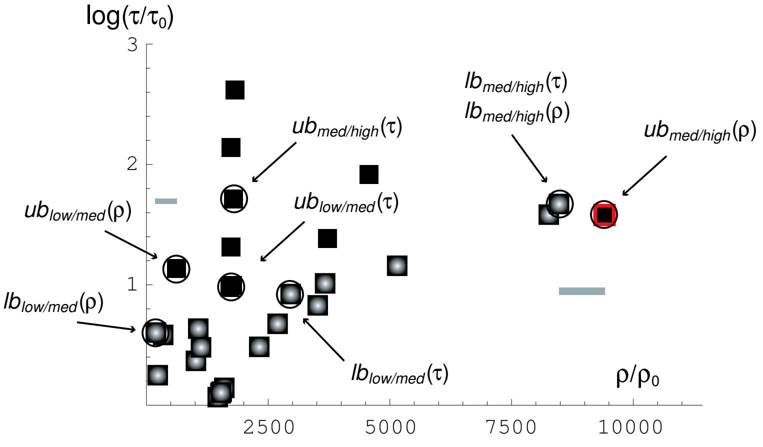

Figure 9.8: Relevant chemicals for sensitivity of τ and ρ values in scenario *Pandora*. Although we would expect eight such chemicals, there are only seven, as *Hset* has a *maximum* (which defines two boundaries). The hypothetical substance added at ($\rho = 9200km, \tau = 50d$) is marked in red.

$lb_{low/med}(\rho)$: The lower boundary of the range *low/medium* is given by the minimum elements of $Hset$ with respect to ρ; *acetic acid*, *adipic acid* or *bisphenol A*, which all take the same value. Movement below is bounded by $\rho = 0$, and above at $\rho = 173km$, since regardless of which of the four chemicals moves, the remaining three stay at that minimum. ⇒ sensitivity medium

$ub_{low/med}(\rho)$: The upper boundary is given by the substance *aldrin* ($\rho = 620km$), may move down to $\rho = 0km^{**}$, and up to $\rho = 1770km$, the value for *dieldrin*. ⇒ senstivity medium

$lb_{med/high}(\rho)$: The lower boundary of the range is defined by the *maximal element* of $Hset$ with respect to ρ. If the element (*ethylene oxide*) has a higher value for ρ, $lb_{low/med}(\rho)$ would simply move up. If the element has a lower value, then $lb_{low/med}(\rho)$ would only move down to $8300km$, the value taken by *ethylene dicloride*, the next largest elements of $Hset$ with respect to ρ. ⇒ sensitivity medium

$ub_{med/high}(\rho)$: The upper boundary is defined by the hypothetical chemical included earlier. No sensitivity, except if ρ is calculated differently from τ.

All sensitivities are either low or medium in all cases below on order of magnitude.

9.3.2 *Bioaccumulation*

There are multiple solutions for the calibration of the scenario *Bioaccumulation*, and hence the sensitivity could be conducted for each of these solutions separately. We will presently abstain from this, and only conduct the analysis for the solution which was chosen.

$lb_{low/med}(\tau)$: The lower boundary for the range *low/medium* of τ is given by the substance *propylene oxide*. This is the same as for *Pandora*, and since the discussion of the boundaries of τ is independent of the boundaries the second parameter. Hence the discussion can been adopted from that section. ⇒ sensitivity low

$ub_{low/med}(\tau)$: The upper boundary for the *low/medium*-range of τ is given by the *minimal* τ-value within the reference set *Rset*, as in *Pandora*. As before, the discussion is identical to the discussion for *Pandora*. ⇒ sensitivity medium

$lb_{med/high}(\tau)$: The lower boundary for the range *medium/high* of τ is the maximal τ-value within the reference set $Hset$. The discussion, again, is identical to the discussion for *Pandora*. ⇒ sensitivity medium

---

[**] more accurately: $173km$

$ub_{med/high}(\tau)$: The only deviation from discussion of $Pandora$ occurs for the upper bound of the $medium/high$-range of τ. In case the original value at $\tau = 243 days$ is taken (HFC $152a$), the value can move down as far as $47 days$, the value of the $maximal$ element of $Hset$ with respect to τ. On the other hand, it can move up, but no further than $\tau = 280d$, as this is the value for the substance $CFC - 12$. We have presently chosen the same value as in $Pandora$, however. ⇒ sensitivity low.

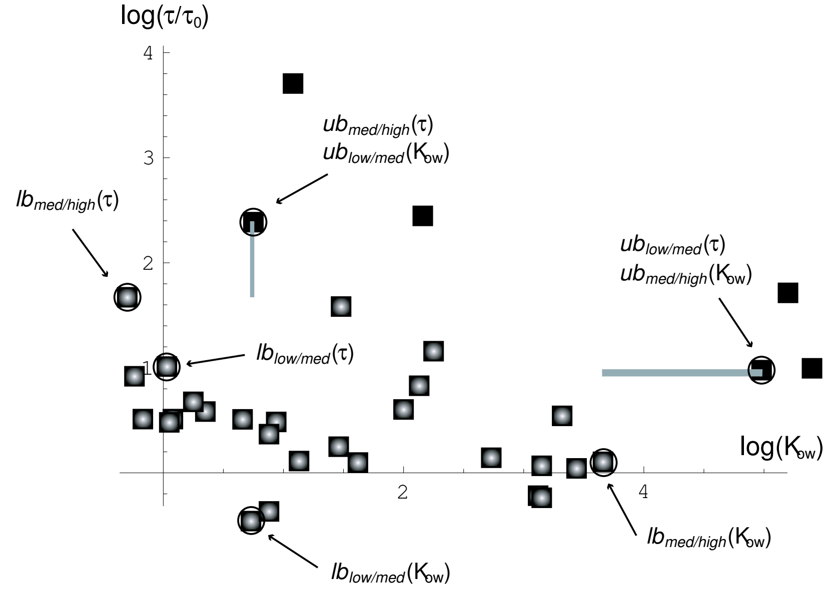

Figure 9.9: Relevant chemicals for sensitivity of τ and K_{ow} values in scenario $Bioaccumulation$. There are only six such chemicals, as the two substances $dieldrin$ and HFC $152a$ are both $minima$ within subsets.

$lb_{low/med}(logK_{ow})$: The lower boundary for the range $low/medium$ for K_{ow} is given by the substance $vinyl$ $acetate$ at $logK_{ow} = 0.73$. It can move to the left until $logK_{ow} = 0.66$ ($caprolactam$) and to the right until $logK_{ow} = 0.88$ ($butyraldehyde$). ⇒ sensitivity low

$ub_{low/med}(logK_{ow})$: The upper boundary is given by the substance HFC $152a$, as for $Pandora$. Since the solution chosen for $Bioaccumulation$ results from the sequence $cdab$, the value can move as low as conceivalbe, i.e. $K_{ow} = 0$. If the value drops below $logK_{ow} = 0.73$, then the value $lb_{low/med}(logK_{ow})$ must change, too. Upwards, the boundary can move up to $logK_{ow} = 1.06$ for the substance $CFC - 22$. ⇒ sensitivity medium

$lb_{med/high}(logK_{ow})$: This lower boundary (at $logK_{ow} = 3.66$) can move down to $logK_{ow} = 3.44$ and up to infinity (or at least the maximal value for $logK_{ow}$ obtained). ⇒ sensitivity medium

$ub_{med/high}(logK_{ow})$: The upper boundary ($logK_{ow} = 3.66$) can drop as low as conceivable, but cannot move further up than $logK_{ow} = 5.20$. ⇒ sensitivity medium

It appears that the calibration for *Bioaccumulation* is more sensitive than the calibration for *Pandora*, due to the sensitivity to changes in K_{ow}. It must be pointed out, that although the values for e.g. $ub_{med/high}(logK_{ow})$ may be lower than expected, it is unlikely to be smaller by more than one order of magnitude. Similarly, $lb_{med/high}(logK_{ow})$ can be larger, but is unlikely to be so by more than one order of magnitude, in which case the boundary obtained for the discretisation is still fairly good.

All in all, it appears that the current results are not very sensitive to changes in values for the reference chemicals. Nevertheless even better results could possibly be obtained by using a different construction of the evaluative properties (change in Assumption 3). Note however that a more elaborate method for the sensitivity analysis may have to be contemplated, in order to confirm the low sensitivity for the two scenarios.

9.3.3 Changes in Reference Sets

Pandora

By inspection of the areas α_2 and α_4 we may ask if the areas could move significantly by deletion of elements from the reference sets. In the case of *Pandora*, the effects of deletion are the same as the effects of moving values, if only in one direction. It appears from this point of view that the calibration is robust (see sensitivity).

Bioaccumulation

The same argument holds as for changes in the reference set in the case of *Pandora*. There is an additional difficulty, however, which arises from the fact that there are more possibilities to place the areas α_2 and α_4. The general criterion of selection for the Criteria *a* and *b* is the number of correctly classified chemicals. The choice among the possibilities for α_2 and α_2 is made precisely using the number of correctly classified chemicals. As a result, changing the number of chemicals in a particular area may change the shape, and even the positioning of the areas. Such is indeed the case, as is shown in the following example with the application of sequence *cdba*.

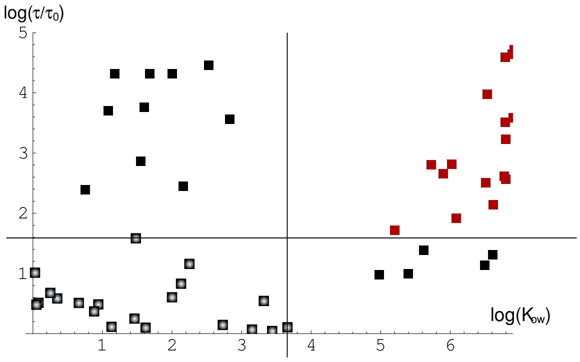

Figure 9.10: Application of Criterion *a* during calibration of *Bioaccumulation*. From the two blocks of elements of *Rset* depicted in black (upper left, lower right), only one can be correctly classified. The other is classified as "orange". The elements depicted in dark red are classified as "red" anyhow. Presently the set on the upper left contains more elements than the set on the lower right and hence application of Criterion *a* positions the grid so as to classify the block on the upper left as "red" and the block to the lower right as "orange".

When applying Condition *a* during the calibration of *Bioaccumulation*, the algorithm chooses to correctly classify the subset on the upper left in Figure 9.10 as it contains more elements than the subset on the lower right, which consequently is classified as "orange". Had the subset on the right contained more elements, the subset on the upper left would have been classified as "orange". Hence the classification is accidental.

Luckily, the sequence presently chosen is *cdab*, for which the above-mentioned, accidental classification does not occur. Nevertheless, by using an algorithm that *counts* the number of chemicals in a particular section of the Cartesian plane, the algorithm and hence the calibration can turn out not to be very robust. In any case, the analysis shows that particular care should be placed on the choice of chemicals and the avoidance of *redundancies* in the reference sets (as is the case when choosing two very closely related derivatives).

9.4 Concluding Remarks

The present chapter sets out to find limiting values that define discretisation. Such values have been identified and hence the discrete classes are now defined for each of the evaluative properties (i.e. for each of the filters corresponding to threat scenarios). Since a linearisation function has been found in Chapter 5, the assessment function Π is now completely defined.

We are now in a position to answer Problem 2 (restated). In the case of both *Pandora* and *Bioaccumulation*, the formal structure plus the assumptions made (which together define the formal precautionary procedure) *do not allow* for a discretisation of the filters, if calibration is based solely on existing regulation. This implies that a suitable assessment procedure does not result from existing regulation, within the formal precautionary procedure adopted.

Problem 2 can be answered affirmatively, however, if we allow for the following three additional assumptions

- A hypothetical reference substance can be added to *Rset*, in an area of the threat scenario *Pandora* that is scientifically possible.
- A solution can be dismissed on grounds that we obtain degenerate boundaries for one of the threat indicators inside a scenario.
- In case boundaries for a parameter are known from a scenario, than these boundaries can be used for another scenario.

Furthermore, both results from calibration seem to indicate that the results obtained from the sequence *cdab* are more suitable and hence require less additional assumptions or modifications on existing assumptions then the results obtained with *cdba*. This would mean that within the formal precautionary procedure and based on existing regulation, a prioritised correct classification of the elements that are globally regulated is preferable to the correct classification of chemicals that are not globally regulated. In a limited sense this puts more weight on "precautionary" considerations.

It must be pointed out, however, that a number of assumptions are made herein, that could (and should) be further challenged in the face of the results obtained. The fact that the Problem could be answered with the help of a few modification does not indicate that the problem has been solved optimally and therefore conclusively.

It should be noted that only Assumption 4 has been questioned presently. The other assumptions may also be responsible for the failure to sufficiently restrict the calibration in order to get suitable discretisation intervals. This may be the case for Assumption 3, which is shown to be subject to issues of robustness. The remaining two assumptions, namely Assumption 2 and Assumption 1 appear to be less questionable at present.

Chapter 10

Classification chemicals

10.1 Classification of Reference Chemicals

Before a test of the assessment procedure is performed on a set of test chemicals, the classification of the reference chemicals is first discussed. The classification procedure is given by the sequel operator introduced in Chapter 5. In the present case the number of evaluative properties is 2; for the two threat scenarios *Pandora* and *Bioaccumulation*, we have the following definition of classes (see Figure 5.4):

"green": for the tuples $(1,1)$, $(1,2)$, $(2,1)$

"orange": for the tuple $(2,2)$

"red": for the tuples $(1,3)$, $(3,1)$, $(2,3)$, $(3,2)$ and $(3,3)$

The reference chemicals in $Hset$ are excluded from placement in the class "red" through the optimisation algorithm, as the elements in *Rset* are excluded from the class "green". In case a reference chemical is not optimally classified, it is in "orange".

The classification of all 78 chemicals is given in Appendix I. The following is a short synopsis.

The filter corresponding to the scenario *Pandora* classifies 9 substances as "orange" (both from the reference sets $Hset$ and *Rset*): *propyleneoxide*, *terephtalic acid* DME, *ethylene dichloride*, *ethylene oxide*, *aldrin*, $cis - chlordane$, *dieldrin*, *heptachlor epoxide*, PCB 28. The filter corresponding to the scneario $Bioaccumu-$ *lation* classifies two substances as "orange": *ethylene dichloride* and *terephtalic acid* DME. Accidentally then, the substances in "orange" in the overall procedure are equal to the substances in "orange" for *Bioaccumulation*.

10.2 Classification of Test Chemicals

A set of eleven chemicals is chosen as a proper test for the procedure (Appendix H). As a test, the application to this set should illustrate the usefulness of the present approach as an assessment tool. In this sense, the test set contains chemicals that are presently discussed for (global) regulation. The application procedure then aims at showing that such chemicals as are contemporarily discussed can be distinguished on grounds of their environmental threat potential.

The chemicals show some a priori evidence of persistence or bioaccumulation or long-range transport or are produced or emitted on a large scale. Some of these special chemicals might be regulated on a national level. As an application of the filter series technique, they are submitted to a screening by the assessment procedure developed earlier. The test chemicals were sourced by Morosini, and the following description of the substances is taken from [12].

Hexachlorocyclohexanes: alpha-HCH, beta-HCH, gamma-HCH

substance	*Pandora*	*Bioacc.*	*overall*
alpha-HCH	3	3	"red"
beta-HCH	3	3	"red"
gamma-HCH	3	3	"red"
endosulfan	1	1	"green"
carbaryl	1	1	"green"
carbofuran	1	1	"green"
HMDS	1	1	"green"
OMCTS (D4)	1	2	"green"
DMCPS (D5)	1	2	"green"
HBB	3	3	"red"
DBDE	3	3	"red"

Figure 10.1: Classification of eleven test chemicals. Classification is given for *Pandora*, *Bioaccumulation*, as well as for the overall classification procedure.

As a first test the two-filter procedure was applied to three stereoisomers of the insecticide *hexachlorocyclohexane*: *alpha*-HCH, *beta*-HCH and *gamma*-HCH (*lindane*). These are the major components of the once widely used so-called technical HCH (*benzene hydrochloride*, BHC). They are also the most frequently detected HCH isomers in environmental samples and in human fat and milk. Technical HCH is now banned in most industrialised countries, where in contrast *gamma*-HCH (the only insecticidal isomer) is used as an almost pure substance (*lindane*). In the US the production of *lindane* ceased in 1976. *alpha*-HCH and *gamma*-HCH are almost ubiquitous in environmental samples from every continent, including polar and pristine regions (from [12]).

The three chemicals received a 3 score in the *Pandora* as they did for the *Bioaccumulation* filter and hence a "red" score from the assessment. Note that although they are widely considered as POPs in scientific literature, the HCHs are not included in the Stockholm Convention.

Endosulfan
The substance is a polychlorinated cyclodiene insecticide, the use of which is permitted in most countries because of its relatively rapid degradation in air and water and because of its lower tendency to bioaccumulate if compared to DDT or HCHs (from [12]).

It passes both filters, getting a 2 both in the *Pandora* and in the *Bioaccumu−lation* filters, and hence a "green" from the assessment

Carbamates: carbaryl, carbofuran
Carbaryl and *carbofuran* are the most widely used *carbamate* insecticides. Because of their rapid degradation in air and water (due to photooxidation, photolysis, hydrolysis, and biodegradation) and their low tendency to bioaccumulate, their potential for persistence and long-range transport is supposed to be low (from [12]).

Both chemicals pass both filters, getting a 2 both in the *Pandora* and in the *Bioaccumulation* filters, and hence the overall score "green".

Silicones: Hexamethyldisiloxane (HMDS), octamethylcyclotetrasiloxane (OMCTS or D4), decamethylcyclopentasiloxane (DMCPS or D5)
On the basis of environmental monitoring and general ecotoxicological considerations, a possible role of *silicon* compounds as a general new class of environmental chemicals has been postulated. It is then interesting to test the precautionary filter procedure on some of these compounds, such as *hexamethyldisiloxane* (HMDS),

octamethylcyclo-tetrasiloxane (OMCTS or D4), and *decamethylcyclopentasiloxane* (DMCPS or D5), which are man-made special representatives of the *silicones*, commonly referred to as *polymethylsiloxanes* (from [12]).
They all pass the *Pandora* filter (1 score), and get one 1 and 2 scores (for OMCTS or D4 and DMCPS or D5) in the *Bioaccumulation* filter.

Brominated chemicals: hexabromobenzene (HBB), decabromodiphenyl ether (DBDE)
Brominated compounds such as *polybrominated diphenyl ethers* (PBDE) are widely used as flame retardants in consumer products. They have been detected in environmental and human milk samples in industrialized countries, with increasing concentrations in the last decades. Two of them are submitted to the two-filter procedure. *Hexabromobenzene* (HBB) is used as flame retardant in polymers. It is not expected to be degraded by direct photolysis, hydrolysis, chemical oxidation or biological activity (from [12]).

Hexabromobenzene is retained by both the *Pandora* and the *Bioaccumulation* filter (two 3 scores), resulting in the overall classification "red".

Decabromodiphenyl ether (DBDE) is used as flame retardant in textiles, rubbers and virtually every class of polymers (ABS, PVC, polyamides, polyesters, polyolefins, etc.) It degrades in air, water, and soil only in the presence of sunlight. Hydrolysis and biodegradation have not been reported. The potential for bioaccumulation is reported to be low or moderate (from [12]).

It is retained in both the *Pandora* filter (3 score) and the *Bioaccumulation* filter (3 score) and hence is classified as "red".

10.3 Summary

We have tested eleven chemicals from five classes of chemicals, and have obtained classifications of chemicals in "red" as well as in "green". Interestingly, all chemicals that were classified as "red" in the overall assessment, are also classified as "red" (i.e. (3)) in the individual scenarios. We note on the other hand that two chemicals, namely *OMCTS* and *DMCPS*, are both classified as "green" (i.e. innocuous), although the chemicals have been classified as "orange" for the scenario *Bioaccumulation*.

The validity of the present assessment is not easy to justify upon inspection of the result obtained by the test chemicals. On the other hand, the result obtained does not appear to invalidate the procedure either, i.e. by being overly restrictive.

Chapter 11

Concluding Remarks

At the outset of this thesis we proposed to define and discuss an exposure-based approach to the assessment of environmental chemicals. The discussion should be conducted in a way as to answer the key question of suitability of the procedure. The definition of suitability we gave was that the procedure should be consistent, but also that the assessment should be applicable in the sense that it got by with as little as possible assumptions of the more political type. We also announced that, in case suitability was not given right away, the use of a formal approach would isolate problematic areas and would help narrow down the degrees of freedom, so to speak, that need to be addressed in order to arrive at a suitable procedure. Finally, we noted that even though the thesis does not claim to comprehensively discuss the precautionary principle in the environmental assessment of chemicals or even to find the most appropriate definition of the principle for the current task, if a suitable exposure-based assessment resulted form the present work, this would give an indication on the legitimacy of the principle in the context of the assessment of environmental chemicals. Or rather: the key problematic areas found in the analysis of the assessment procedure could indicate targets for future work, in order to operationalise a precautionary approach in the environmental assessment of chemicals or beyond.

We chose the strategy of addressing the main objectives of the work by first conducting two separate theoretical investigations: The first (Block 1) concerning the derivation of exposure-based indicators starting from an analysis of current chemical assessment and the second (Block 2) concerning the formal aspects of the assessment. The first part was mostly taken from work previously done by present and past members of the supervisor's research group, whereas the second part rests predominantly on contributions by the author. Based on these findings we proposed to discuss three main problems (Problems 1, 2 and 3).

Block 1*: Chemical Assessment and Precautionary Approaches*

Chapter 2 started with an overview of the current practice in the assessment of environmental chemicals. In view of the major challenges that have emerged over recent decades with such assessments, we moved to the concept of exposure-based threat indicators as it has been developed previously in the group.* Such indicators were further investigated and bundled into threat scenarios, a method which has also been previously described, although this step is further formalised in the present work. The chapter then established two scenarios which relate to exposure: the scenario *Pandora* and the scenario *Bioaccumulation* (Assumption 1). We noted that whereas such threat scenarios are multi-dimensional, they actually needed to be linearised at some point, if they were to form the basis for an assessment.

We also gave a brief overview on the various definitions of the precautionary principle and showed how the concept of exposure satisfies all properties or dimensions that are associated with the principle, at least as given in a recent definition by Sandin [56]. We then discussed formal problems that have been encountered in the past with the precautionary principle, and showed that the strongest of these results relies on the occurrence of probabilities in the definition of the principle. The current exposure-based assessment, although it can be viewed as fulfilling a particular definition of the precautionary principle,† does not contain any sort of probabilities and it appears therefore that such limitations of the principle as reported recently in literature do not concern the present approach.

Block 2*: Formal Aspects of the Assessment Problem*

Chapters 3, 4 and 5 focused on formal aspects of the assessment problem. Chapter 3 introduced the formal (and fundamental) concept of partial orderings and partially ordered sets, as well as a number of special orderings and operators. Next, Chapter 4 started with a brief theory of the simplification of partially ordered sets. As a method of simplification we introduced *coarse-graining*, i.e. the replacement of continuous values with discrete classes. This amounts to breaking down the original decision rule Π into two subrules: the coarse-graining function Δ and the actual linearisation function Λ that takes the coarse-grained poset as its domain. These considerations

---

*The concept was developed mainly by Scheringer, see [26].

†namely that of the principle as the maximum probable loss rule

led to the concept of event-space, i.e. the visualisation of the set of all events that can occur. Such a graphical representation was only possible for a limited number of events, and such was the case when each dimension was limited to a restricted number of classes: We showed how the event-space can be graphically depicted in the present case, for up to three dimensions, if only three equivalence classes per dimension are permitted. This concept being established, we moved on to show that the event-space could be used as a tool to demonstrate the consistency of a set of requirements (Consequence 1). In order to arrive at this results we established Conjecture 1.

The second part of Chapter 4 identified five requirements deemed relevant for the present classification problem. The requirements were of two kinds: general rationality requirements (Requirement 1 (Order-preservation) and Requirement 2 (Symmetry)) as well as three requirements that are more specific to the present problem (Requirement 3 (three-valuedness), Requirement 4 (Transparence) and Requirement 5 (Expandability)). We also mentioned Arrow's requirements of *Non-dictatorship*, *Independence of Irrelevant Alternatives* and *Unrestricted Domain*. Starting from these requirements we derived Consequence 2, namely that if calibration was performed with the help of a reference and concomitantly for all evaluative properties, then Requirements 2 (Symmetry) and 5 (Expandability) are no longer compatible. This, we observed, stipulates that the calibration of the assessment procedure has to be conducted independently for each evaluative properties, i.e. independently for each of the filters (Corrolary 1). It also stipulates that each evaluative property (threat scenario) must yield three equivalence classes (Corrolary 2). The chapter concluded with a classification scheme for the assessment problem, which resumed the formal results developed in the chapter in a graphical manner.

The last part of the second theoretical block, which is discussed in Chapter 5, concerned the linearisation function Λ, by making use of the Consequence 1: instead of separately deriving a linearisation function, which then is tested for consistency with respect to the requirements and consequences/corrolaries thereof, these formal statements were used to map successive extensions of the event-space. We thereby obtained an ordering of the event-space that fulfilled all formal statements (requirements, consequences, corrolaries) and subsequently introduced a linearisation function Λ that leads to this ordering. This function was identified as the sequel operator $S(\prec)$, combined with cut-off values.

Combination of the Two Blocks

Chapter 6 consolidated the two theoretical blocks. The main result consisted in the identification of the filters (i.e. the quantitative, linear representations of threat scenarios) with the evaluative properties in the classification scheme. The domain of the linearisation function was thus defined: the discrete, partially ordered set $\mathbf{P}^{\{\tau,\rho\}}$. It was also shown that the equation of filters and evaluative properties directly entails the discretisation of filters, as discretisation is required for the evaluative properties. We then picked up an observation made at the outset of Chapter **??**, namely that the calibration of the assessment procedure is generally required and introduced a formal assumption with regard to the best way such a discretisation could be conducted (namely by performing a calibration with respect to a reference, Assumption 2).

Such discretisation is typically performed on continuous values. Here, we encounter the special situation, however, where no such domain of continuous values exists for the filters: the threat scenarios first have to be linearised by a function to filters. The last part of Chapter 6 introduced a simple function, which in fact already performs the discretisation (Assumption 3).

Hence the coarse-graining function Δ was found to be part of the linearisation function which defined the filters. The function chosen to this end places a 3×3-grid on the Cartesian plane defined by the two threat indicators that in each case define the threat scenario. Together with the classification scheme (which includes the aggregation of threat indicators to filters and subsequent discretisation) this allowed to define the assessment function Π that reflects the assessment procedure, as at this point formal aspects as well as the particular choice of approach are reflected.

The calibration method was introduced in Chapter 7 based on the four criteria (a, b, c and d) and was designed in way as to yield areas, inside of which solution were equally optimal. As the reference for the calibration, we used two reference sets, which have been described in previous work and are paradigmatic for low and high threat potential, respectively (Assumption 4). Chapter 7 also showed that two sequences (*cdba* and *cdab*) of the four criteria exist, which are equivalent with respect to the formal structure.

Problem 1: What is an adequate formal structure and what is an adequate (mathematical) decision rule for an exposure-based environmental assessment of man-made (xenobiotic) chemicals?

An adequate formal structure for the description of chemicals turns out to be given by the partially ordered set. Also (as has been shown previously by other authors) the ultimate result of the assessment is a linear ordering and the decision rule can be viewed as a mapping Π from a partially ordered to a linearly ordered set. Furthermore, this mapping can be subject to requirements: Decision Theory offers a number of requirements that can be used in the sense of restrictions in order to define a classification scheme as a general frame for precautionary chemical assessments. In order to arrive at such a scheme, we established a total of five requirements. These requirements are on one hand more general rationality criteria (such as Requirement 1 (Order-preservation) and Requirement 2 (Symmetry)) and on the other hand more specific to the problem of environmental assessment of chemicals, such as for the Requirements 3 (three-valuedness), 4 (Transparence) and 5 (Expandability). It appears that the formal structure of the procedure is consistent, if certain technical assumptions are made (concerning the use of cardinal values for the evaluative properties). Hence, the adequate formal structure is found by describing the basis of the assessment problem as that of a set of chemicals characterised with evaluative properties and giving rise to a partially ordered set. Such a set is then mapped by decision rule Π, which must respect a number of requirements.

But Problem 1 also concerns finding an appropriate linearisation function (i.e. the decision rule Π). Typically, such a function is first established on grounds of some scientific understanding on how such a function should be, to be only later scrutinised with respect to formal statements. Somewhat surprisingly it is found that deriving such a function independently is not necessary, as we can define a linear ordering simply by applying the requirements on the event-space. (In order to be able to perform this task, the concept of event-space is necessary). Based on the linear ordering thereby obtained, a function can easily be found which reproduces that ordering. It is noteworthy that instead of choosing a function and checking for requirements, we are hence able to discuss the requirements and assumptions first, in order to attain the linear ordering that is compatible with the requirements. The remaining task then simply consists in finding a function that maps the domain to that particular ordering. In the present work, the requirements are sufficient for the definition of the linearisation function, except for the cut-off values, which have to be set in order to arrive at three equivalence classes, and hence the search for a suitable linearisation function has been transferred to a good part to that of defining the appropriate requirements. Another somewhat novel feature of this work is that this strategy relies on the use of special ordering relations such as weak orders and interval orders.

Problem 2: Based on the formal precautionary procedure, can a suitable assessment procedure be extracted solely from existing regulations?

The first answer to the problem concerns consistency. It is shown that there exists a function that fulfills all requirements made in the work, if we admit cardinal values to describe the discrete intervals. The present work also establishes the link to levels of indifference, as introduced in the rather different context of election problems, in response to Arrow's Impossibility Theorem. An important observation with respect to consistency was also made in Chapter 4, where we proved that the fulfillment of requirements restricts the type of optimisation that can be used in calibration: it is not permitted to conduct the optimisation for all evaluative properties (i.e. for all filters) concomitantly. Rather, each filter must be optimised (or calibrated) separately in order for all requirements to be fulfilled. This result is significant as if it had not been obtained, we may well have chosen another form of optimisation, which in turn would have led to inconsistency.

The second part of the answer concerns the applicability of the procedure. This question relates to the demand made in the introduction concerning the type of information which should be used: The classification scheme demands that the assessment function be calibrated, with the help of a reference (Assumption 2). As indicated earlier, for such reference, information on existing regulation is to be used. The question then arises immediately, if existing regulation is sufficient on its own to permit calibration. The main problem Problem 2 can hence be restated:

Problem 2 (restated): To what extent can — within the frame of the formal precautionary procedure — the filters be discretised solely through regard to existing regulations?

This central question can be answered by using the results obtained from a specially developed algorithm. The algorithm is designed in a way as to yield all possible solutions which are optimal with respect to the reference, leading to areas of solutions. This now offers a possibility to judge on the suitability of a procedure: If the optimal areas are such that discrete values can be chosen with limited arbitrariness, then the procedure can be said to be suitable, as no more assumptions need to made to define the assessment procedure. We found two rough criteria which allow, based on the shape and relative positioning of the two relevant areas α_2 and α_4, to judge upon a given procedure: 1. Occurrence of multiple, disjoint solutions and 2. Unsatisfactory shapes of areas.

Based on these crtieria it is found that the discretisation of the filters (and hence the quantification of the threat scenarios) is under-determined, if the calibration is based on existing regulations. As such, the assessment procedure is not suitable.

In order to make it suitable we must hence revisit the formal statements, i.e. the requirements and assumptions made. It thereby turns out that the procedure can be made significantly more suitable through altering some of the assumptions and adding two additional ones: One assumption is made with respect to the choice of reference (and can be viewed as a supplement to Assumption 4), whereas a second and a third assumption concern additional criteria for the selection of a suitable procedure. The three additional assumptions made in order for the assessment procedure to be suitably defined are

- Hypothetical reference chemicals can be added to the sets of reference chemicals to reflect the maximal spatial range for a persistence of around $50d$.
- A solution for calibration can be ruled out if it leads to degenerate discretisation of an indicator.
- In case an indicator occurs in several threat scenarios, discretisation of the indicator in one threat scenario can be used to define discretisation in another threat scenario.

It is beyond the scope of the present work to analyse these assumptions in detail, except to state that they don't appear to be entirely unreasonable. But note that the first bullet point relates to Assumption 4, whereas the second and the third point do not relate to any of the technical assumptions in a direct manner. The choice of these modifications is somewhat accidental. Hence, in a future analysis, modifications on the other assumptions would have to be discussed as well. The most appropriate candidate for such a discussion of the assumptions indeed appears to the author to be Assumption 3. Remember that this assumption deals with the choice of the aggregation function for the indicators which define the threat scenarios.

An analysis of the criteria yielded that there are two sequences which are equivalent within the formal frame, as seen above. Strangely, the results of calibration indicate that *cdab* yields better results than *cdba*: Suitability of the procedure is much better if the sequence *cdab* is used. In the sense of the interpretation of the two sequences, this implies that the formal structure and the technical assumptions define an assessment procedure more suitably if the assessment gives more weight to precautionary considerations rather than to economic considerations. Remember that an indication of the same preference of *cdab* over *cdba* is found during calibration of *Pandora* as the inclusion of highly immobile chemicals does not change the calibration for *cdab*, but does so for *cdba*. The observed superiority of *cdab* over *cdba*, however, is not further discussed in the work and should not be overrated.

The results found as answers to Problem 2 are only meaningful if they are stable. We hence conducted a sensitivity as well as a robustness analysis at the end of Chapter 9. These analyses yielded a low sensitivity for the scenario *Pandora*, and slightly higher sensitivity for *Bioaccumulation*. With regard to robustness, the result was satisfactory, too. The results of calibration hence appear to be rather stable.

It should not be overlooked that there is an entirely different type of stability, that has only been superficially touched upon in the present work, namely that of the requirements made in Chapter 4. The results presented herein may vary if requirements are changed or indeed if requirements are added or withdrawn. We gave a brief discussion of this problem at the end of Chapter 4.

Problem 3: What does the exposure-based assessment procedure yield when applied to a set of appropriate test chemicals?

The assessment of eleven test chemicals, which are presently discussed for regulation, identifies a set of chemicals that should be regulated based on the exposure-based assessment procedure. These are the *hexachlorocyclohexanes* and two brominated chemicals (*HBB* and *DBDE*). Endosulfan, two carbamates and three silicones are found to be innocuous with respect to global environmental threat. The significance of this result is presently difficult to estimate, as some points relating to the definition of the procedure may still have to be more comprehensively discussed. It shows, however, that chemicals can be classified into different classes with respect to their global threat potential. And as much as some chemicals appear to pose such a threat, others don't appear to do so.

Concluding Remarks and Outlook

The present thesis establishes a formal setting for the assessment of environmental chemicals. It also proposes a decision rule Π that satisfactorily reflects the formal setting and that appears to be free of inconsistencies. Serious limitations are uncovered, however, when the crucial step of calibration is performed. This necessitated a number of adjustments to the setting and to the technical assumptions made herein. Nevertheless, it appears that a consistent and suitable decision rule that is based on the concept of exposure can be established.

To the extent this is the case, the present work indicates how the precautionary principle, if only in a narrow definition, can be applied to the assessment of environmental chemicals. The thesis demonstrates that within the formal setting that is mathematically founded, an assessment procedure can be defined with the help of formal requirements and a number of technical assumptions, and hence relies on more political input in only a very limited way. Furthermore, the formal setting allows to define and detail each of the assumptions that are necessary for the definition of the procedure, independently from each other. This allows to discuss the problematic aspects of the definition of such a procedure in a systematic way, even if some key questions cannot be answered conclusively.

One key problem concerning the definition of the procedure has been localised in the aggregation of threat indicators to threat scenarios (Assumption 3). Another important problem may be seen in Requirement 3, which may well be weakened, so that the final ordering can accomodate more than three classes. The other requirements appear to be less subject to debate. Assumption 4 may of course be subject to an overhaul, although the sensitivity analysis appears to indicate that small changes in either the data or the set of substances itself, may not have a large impact. The same applies to Assumption 1; it may turn out that alternative technical definitions of the threat indicators are more appropriate from a scientific point of view. But this would just change the data for the reference sets and hence this alteration can be discussed with similar arguments than Assumption 4 (i.e. sensitivity, robustness). Finally, it cannot be excluded that an alternative to Assumption 2 can be used, thereby using a different method to perform discretisation.

We must also point to the curious finding in Chapter 5 namely that the sequence of application of requirements on the event-space appears to be of significance. We have at present no orientation on how to view this finding.

On a more methodological level, we also have developed a method of depicting the

Hasse diagrams of all events that can possibly occur in our setting. This method may be used in other topics related to chemical assessment, or more generally, in classification problems. Similarly, the introduction of special ordering relations such as weak orders and interval orders that are shown to be powerful in the present context may serve to establish such structures in assessment problems in environmental sciences and elsewhere.

Appendix A

Basic Definitions

reflexive:	$\forall x \in X,\ xRx$
irreflexive:	$\forall x, y \in X,\ not(xRy)$
symmetric:	$\forall x, y \in X,\ xRy \Rightarrow yRx$
anti-symmetric:	$\forall x, y \in X,\ xRy$ and $yRx \Rightarrow x = y$
asymmetric:	$\forall x, y \in X,\ xRy \Rightarrow not(yRx)$
complete:	$\forall x, y \in X,\ x \neq y,\ xRy$ or yRx
transitive:	$\forall x, y, z \in X,\ xRy$ and $yRz \Rightarrow xRz$
negatively transitive:	$\forall x, y, z \in X,\ xRy \Rightarrow xRz$ or zRy
semi-transitive:	$\forall x, y, z, w \in X,\ xRy$ and $yRz \Rightarrow xRw$ or xRz

partial ordering	*transitive*, *reflexive* and *anti-symmetric*
strict partial ordering	*transitive*, *irreflexive*
weak ordering	*reflexive*, *anti-symmetric* and *neg. transitive*
strict weak ordering	*asymmetric* and *neg. transitive*
linear ordering	*reflexive*, *transitive*, *anti-symmetric* and *complete*
strict linear ordering	*transitive*, *asymmetric*, and *complete*

Appendix B

Symbols

$A, B, C \dots$	Capital Roman letters for sets
$\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$	Bold capital Roman letters for ordered sets
$a, b, c \dots$	Lower case Roman letters for elements of sets
$\Pi, \Delta, \Theta, \dots$	Capital Greek letters for functions
$\alpha, \beta, \gamma, \dots$	Lower Greek case letters for enumeration

$\{m, n\}$	An unordered tuple of values m and n
(m, n)	An ordered tuple of values m and n
mn	Shorthand for tuple of values m and n
$E = \{e_1, e_2, \dots\}$	Set of evaluative properties (parameters)
R	A set of requirements
X	Set of chemicals
X_{ref}	Set of reference chemicals
$X \times X$	Cartesian product of X with itself
$P, Q, P', L \subseteq X \times X$	Binary relations on X (subsets of $X \times X$)
$\preceq$	Binary relation; the component-wise order relation
$\preceq^{i,j,\dots}$	Component-wise order that reflects requir. i, j etc.
$\parallel$	Binary relation; the incomparability relation
$\approx$	Binary relation; the interchangeability relation
$\simeq$	Binary relation; the identity relation
$\mathbf{P} := (X, P)$	Poset using continuous values for $\{e_1, e_2, \dots\}$
$\mathbf{Q} := (X, Q)$	Extension of a poset
$\mathbf{P}' := (X, P')$	Extension of a poset as a result of coarse-graining
$\mathbf{L} = (X, L)$	Linear extension of a poset
$\Delta : \mathbf{P} \rightarrow \mathbf{P}'$	Coarse-graining
$\Delta_i : \mathbf{P} \rightarrow \mathbf{P}'$	Coarse-graining of a particular eval. property e_1
$\Lambda : \mathbf{P}' \rightarrow \mathbf{L}$	Linearisation function
Π	The overall decision rule or assessment function
$\Pi := \Lambda \circ \Delta$	Composite map of coarse-graining and linearisation: assessment function
$S = \{s_1, s_2, \dots\}$	Set of threat scenarios
$\mathbf{P}^{e_j}, \mathbf{P}'^{e_j}, \mathbf{Q}^{e_j}, \mathbf{L}^{e_j}$	Posets pertaining to a specific ev. property e_j
$\mathbf{P}^{s_j}, \mathbf{P}'^{s_j}, \mathbf{Q}^{s_j}, \mathbf{L}^{s_j}$	Posets pertaining to a specific threat scenario s_j
$\mathbf{P}^{\{s_j, s_k\}} := (X, L^{s_j} \cap L^{s_k})$	Poset defined by the two scenarios s_i, s_j
$\Lambda : \mathbf{P}^{\{s_j, s_k\}} \rightarrow \mathbf{L}$	The linearisation function from a poset defined by two scenarios s_i and s_j to a linear ordering
C^R	Class of all functions that respect the requir. in R
C^{O_i}	Class of functions from i^{th} optimisation method

Appendix C

Glossary

aggregation function	Synonym for linearisation functions
assessment procedure	A rule that attributes chemicals to different categories
assessment function	Formal (mathematical) representation of an assessment procedure
assumptions	Restrictions that are imposed on the classification function, beyond requirements
boundary range	The range inside of which a boundary between two discrete intervals for a particular parameter can lie.
calibration	Setting of boundaries/constants/cut-off values etc. to define a function with respect to reference (presently setting for boundaries for discretisation)
characterisation	The description of the objects (i.e. chemials) to be described at the outset of the classification. The *poset* defined by the evaluative properties reflects the characterisation.
choice of approach	Choice that is made with respect to the definition of the precautionary definition.
classification function	The function that reflects the overall decision rule (Π, see symbols list)
classification procedure	General frame for assessement, i.e. any rule that classifies a set of objects (presently: chemicals) into different categories
consistency	logical compatibility of requirements
decision rule	A rule that attributes objects to different categories
discrete filter	The function that aggregates the parameters to single discrete values.
discretisation	The process by which a continuous value range is partioned into intervals.
evaluative properties	The quantities used to evaluate a set of objects
filter	The function that aggregates parameters defining the threat scenario to single values.
formal assessment procedure	A assessment procedure that is formally derived
formal precautionary procedure	Choice of approach embedded in formal structure
formal structure	The logical structure for the assessment

hazard	Synonym for threat
interval	A range within which a parameter adopts identical values
linearisation function	A function that aggregates value tuples (vector entries) to a single value
lower boundary	The lower end of a boundary range
optimisation	Synonym for calibration in case the latter is performed based on optimality with respect to a reference
parameter	Any sort of measure that yields single values.
quantification	Aggregation of the tuples of the values that are given by the threat scenario, to single values.
requirements	Conditions that the classification function is required to fulfilled
threat	An unwanted state of the world (i.e. the environment): *global* environmental damage
threat (potential)	A measure for a chemical's potential to exhibit *global* environmental damage
threat indicator	An indicator, i.e. parameter for a threat
threat scenario	A more general indication for a threat, i.e. one which can be two and higher dimensional.
upper boundary	The higher end of boundary range

Appendix D

Building Blocks

Assumption 1 *The two threat scenarios Pandora and Bioaccumulation can be used for a precautionary assessment of chemicals. The two scenarios are each (best) defined by two indicator: for Pandora the indicators* τ *and* ρ *and for Bioaccumulation the indicators* τ *and* K_{ow}*.*

Assumption 2 *The calibration of the procedure is most suitably performed by optimising the assessment function* Π *with respect to a reference.*

Assumption 3 *The aggregation of the parameters defining a threat scenario, to a filter, is suitably performed by placing a* 3×3*-grid over the corresponding Cartesian plane and abstracting the resulting* 9 *areas into* 3 *equivalence classes.*

Assumption 4 *Existing regulation is completely reflected by two particular reference sets:* $Hset$ *and* $Rset$*.*

Consequence 1 *In order to prove that a set* R *of requirements is consistent it suffices to show that a single linear ordering of the corresponding event-space respects all requirements in* R*.*

Consequence 2 *Assume the classification function is optimised concomitantly for all evaluative properties with the help of a reference then if follows that the Requirements 2 (Symmetry) and 5 (Expandability) are not compatible.*

Consequence 3 *The classification must order the tuples* $(2,2)$ *and* $(1,3)$ *in the following manner:* $(2,2) \preceq^{R} (1,3)$*.*

Consequence 4 *The filters reflecting the threat scenarios in a quantitative manner yield linear orderings and are equal to the evaluative properties in the classification scheme.*

Consequence 5 *Since the evaluative properties are required to yield discrete orderings in the classification scheme, the filters include discretisation, too (beyond linearisation). Also, the calibration of the classification function consists in finding the values defining the discrete intervals for the filters.*

Conjecture 1 *If the linear ordering of a set* U *respects a set* R *of requirements, then any linear ordering of subset* S *of* U *also respects the set* R*.*

Corrolary 1 *If the Requirements 2 und 5 hold, then it follows that the optimisation of the classification procedure with respect to a reference has to be conducted for each evaluative property* e_i *(i.e. for each threat scenario) separately.*

Corrolary 2 *It follows from Corrolary 1 and Requirement 3 that each evaluative property must yield a linear ordering of three equivalence classes.*

Corrolary 3 *In the special case that a parameter occurs in the context of two or more threat scenarios, it follows from Corrolary 1 that a unique discretisation of that parameter is generally not possible.*

Definition 1 *Filters are the (linear) quantified representations of threat scenarios.*

Definition 2 *The order relation* P *which reflects the initial characterisation of the chemicals is a partial ordering. I.e. it is a transitive, reflexive and anti-symmetric order relation.*

Definition 3 *The order relation* $P = \preceq$ *used in the characterisation is: "has — for all properties* $e_1, e_2, \ldots$ *taken into consideration — smaller values than".*

Definition 4 *The order relation* $(L_{\preceq})$ *reflecting the final classification of the chemicals should be a linear ordering, i.e. a complete partial ordering (transitive, reflexive, and complete).*

Observation 1 *At present only two threat scenarios have been chosen, and such choice does not constitute a willful limitation. Rather, additional scenario could be considered, if deemed relevant.*

Observation 2 *In order for allow for chemicals from the reference sets to be misplaced, one class is introduced in the ordering of the assessment procedure, in addition to the two classes corresponding to the classes in the reference set.*

Observation 3 *Criteria c and d can both be fulfilled completely. Any sequence that starts with cd produces the same result as if c and d were inverted at the beginning of the sequence, to dc.*

Observation 4 *For the set of Criteria* a, b, c, d *that are presently used, there are two sequences of criteria that are suitable for calibration (i.e. to define the discrete values required for the definition of discrete intervals): cdba and cdab.*

Problem 1 *What is an adequate formal structure for the environmental assessment of man-made (xenobiotic) chemicals?*

Problem 2 *Based on the formal precautionary procedure, can a suitable assessment procedure be extracted solely from existing regulation?*

Problem 3 *What does the exposure-based asessment procedure yields when applied to a set of appropriate test chemicals?*

Requirement 1 *The function* Π *representing the classification procedure is an order homomorphism*

$$\forall a, b \in F, a \preceq b \Rightarrow \Pi(a) \preceq \Pi(b)$$

This implies that any function that is a component of Π *is also order-preserving.*

Requirement 2 *The classification procedure* Π *is symmetric if*

$$\forall x, y \in X, \Pi(e_1(x), \ldots; e_1(y), \ldots) = \Pi(e_{\Psi(1)}(x), \ldots; e_{\Psi(1)}(y), \ldots)$$

where $\Psi(i)$ *is a permutation function on the index set* $\{1, 2, \ldots, n\}$.

Requirement 3 *The final order should yield three distinct classes.*

Requirement 4 *If a chemical is classified in the maximal class, the classification should be traceable to the threat scenarios that are the source of this classification.*

Requirement 5 *If a set of evaluative properties* $E = \{\ldots e_k\}$ *has classified a chemical in the maximum class, then this should remain so if the set of evaluative properties is expanded* $E = \{\ldots e_k, e_{k+1}, \ldots\}$.

Appendix E

Proof of Consequence 1

As detailed in the Introduction, the classification procedure has to be calibrated with the help of an outside reference. We will now study the case where this optimisation is performed by comparing the *final linear ordering* of the classification procedure with the outside reference. We will show that such an optimisation is not possible without violating either Requirement 5 or Requirement 2.

For notational reasons and compared to the main text, the descriptors for the evaluative properties are changed from $e_1, e_2, \ldots$ to $s_1, s_2, \ldots$.

The general scheme of the classification it the following: based on a poset ($P^{S_K} = P^{\{s_1,\ldots,s_k\}}$) that reflects the characterisation of a set of chemicals with the help of k threat scenarios, first a coarse-graining is performed on the poset to yield the coarse-grained poset $P'^{S_K} = P'^{\{s_1,\ldots,s_k\}}$. This poset, in turn, is mapped to a linear ordering itself $L^{S_K} = L^{\{s_1,\ldots,s_k\}}$ consisting of 3 equivalence classes: $E_3^{S_K} \geq E_2^{S_K} \geq E_1^{S_K}$.

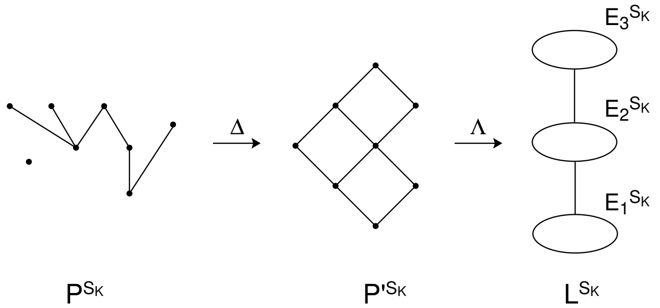

Figure E.1: The classification scheme in its most general form: The poset P^{S_K} is mapped to P'^{S_K} (by the coarse-graining function Δ), which in turn is mapped to the linear ordering L^{S_K} by Λ.

As was mentioned explicitly in the introduction, the classification is designed in a way as to make the use of additional evaluative properties (i.e. threat scenarios) possible at any moment. Furthermore, we can make no *apriori* statement on to how such an additional threat scenario could look like, or particular characteristics such a scenario could have.

If such an additional scenario is indeed accounted for then the set of evaluative properties becomes $S_{K+1} = \{s_1, \ldots, s_{k+1}\}$, with the according final linear ordering $\mathbf{L}^{S_{K+1}}$ consisting of the equivalence classes $E_3^{S_{K+1}} \geq E_2^{S_{K+1}} \geq E_1^{S_{K+1}}$.

Requirement 5 (Expandability) can be rewritten in this notation as

$$E_3^{S_K} \subseteq E_3^{S_{K+1}} \tag{E.1}$$

i.e. all chemicals which were ordered in the highest category by the classification procedure Π^{S_K} must — under any circumstances and independently of the type or nature of the scenario $k+1$ — also be ordered in the highest equivalence class of $\mathbf{L}^{S_{K+1}}$.

The relation expressed in Eqn. E.1 is independent of k, i.e. holds equally for $k-1$, $k-2$ etc. As a consequence Requirement 5 (Expandability) also requires

$$E_3^{S_1} \subseteq \ldots \subseteq E_3^{S_{K-1}} \subseteq E_3^{S_K} \tag{E.2}$$

as we could have equally well have started with $k-1$ and added the k^{th} scenario (rather than start with k and go to $k+1$ threat scenarios). Analogously we could have started with $k-2$ scenarios and then added first the $k-1^{th}$ and then the $k-2^{th}$ scenario.

The number k describes the sequence of the scenarios. Now assume that some scenario, say *Pandora*, is the starting point of a characterisation, i.e. it has $k=1$. But equally well, *Pandora* could have been the second scenario to be considered, i.e. have $k=2$. Indeed, Requirement 2 (Commutativity) states that the sequence of scenario should be of no importance.

In order to clarify whether this is the case with the present optimisation we look at the example of a classification procedure which uses two evaluative properties. In the example the evaluative properties are given by two parameters p_1 and p_2:

Now let us assume that the calibration consists of the following operation: for each of two parameters p_1, p_2 the optimisation yields 3 distinct classes. Hence in the

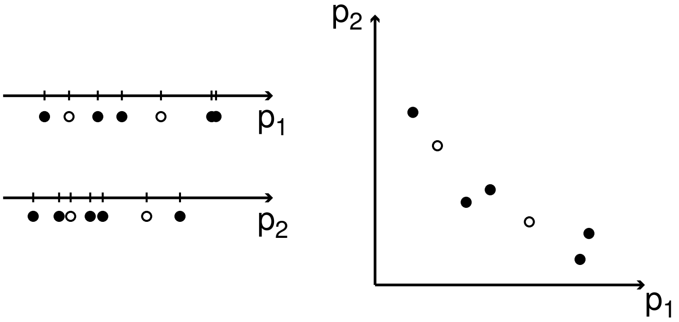

Figure E.2: Characterisation of a set of objects depicted in the Cartesian plane. On the *left* the characterisation is given for each of the properties separately.

Cartesian plane spanned by p_1, p_2 the calibration amounts to optimally placing a 3-by-3 grid over the plane.

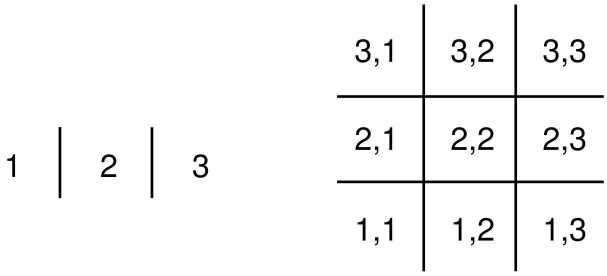

Figure E.3: On the *left* the "grid" is depicted for a the one-dimensional case, on the *right* the grid for the more familiar two-dimensional case.

Furthermore, we define the 3 equivalence classes $E_3^{S_K} \geq E_2^{S_K} \geq E_1^{S_K}$ in the following way:

$$\begin{array}{ll}
\text{in } L^{S_1} & E_3^{S_1} := \{x \in X \mid \Delta(x) = 3\} \\
\text{in } L^{S_1} & E_2^{S_1} := \{x \in X \mid \Delta(x) = 2\} \\
\text{in } L^{S_1} & E_1^{S_1} := \{x \in X \mid \Delta(x) = 1\} \\
\text{in } L^{S_2} & E_3^{S_2} := \{x \in X \mid \Delta(x) = (2,3), (3,2), (3,3)\} \\
\text{in } L^{S_2} & E_2^{S_2} := \{x \in X \mid \Delta(x) = (1,3), (3,1), (2,2)\} \\
\text{in } L^{S_2} & E_1^{S_2} := \{x \in X \mid \Delta(x) = (1,2), (2,1), (1,1)\}
\end{array}$$

In order to evaluate a particular calibration, we establish the two following criteria:

1. No white elements in $E_3^{S_K}$
2. A maximum of the black elements in $E_3^{S_K}$

<u>Case *a*</u>: We begin by calibrating p_1. With respect to Eqn. E.1 only the border between the intervals 2 and 3 is of importance. As shown in Figure E.4 two (black) elements are in the equivalence class $E_{a,3}^{S_1}$.

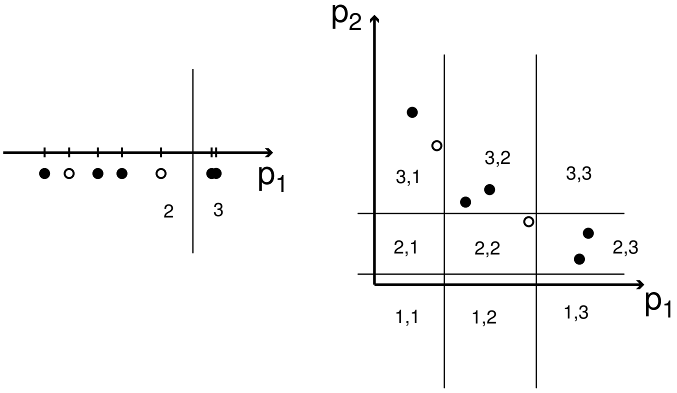

Figure E.4: *Left*: Calibration of parameter p_1. *Right*: Total calibration p_1 and p_2

If now we expand the set of evaluative properties by the property p_2 and in turn optimise then we obtain the calibration depicted on the right-hand side of Figure E.4. As the two black elements which are in $E_{a,3}^{S_1}$, the highest category after calibration of p_1, the calibration is order-preserving.

<u>Case *b*</u>: We now start by calibrating the classification procedure with p_2 as the only evaluative property. Again, with respect to Eqn. E.1 only the border between 2 and 3 is importance. After optimisation we find one black element in the equivalence class $E_{b,3}^{S_K}$.

Expanding the set of evaluative properties by the property p_1, we obtain the same optimisation as in case *a* since nothing distinguishes the situation on the right-hand side of Figure E.5 from that on the right-hand side in Figure E.4. But in that calibration, the black element which is the single element of $E_{b,1}^{S_K}$ is not in the highest equivalence class ($E_{b,3}^{S_2}$), hence Eqn. E.1 is violated. In order to remedy, we must ensure that $E_{b,3}^{S_1} \subseteq E_{b,3}^{S_2}$ which is only possible by conducting the optimisation while keeping an eye on the ordering we obtain when only considering p_2. This, in

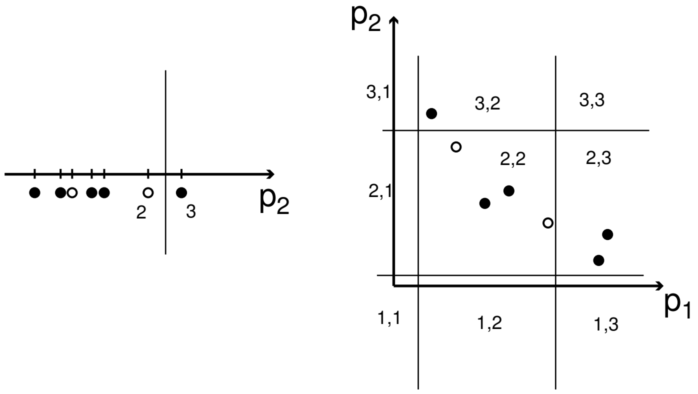

Figure E.5: *Links*: Kalibrierung des Parameters p_2. *Rechts*: Kalibrierung gesamthaft mit p_1 und p_2

turn, excludes the possibility that the two elements that were elements of $E_{a,3}^{S_2}$ in case a are also elements in $E_{b,3}^{S_2}$. Hence

$$E_{a,3}^{S_2} \neq E_{b,3}^{S_2} \tag{E.3}$$

The calibration for the cases a and b are not the same. (As a matter of fact the calibration in case a is better than in case b as more element are in the highest class $E_3^{S_2}$). This is a violation of Requirement 2 that requires precisely the complement, namely

$$E_{a,3}^{S_2} = E_{b,3}^{S_2} \tag{E.4}$$

Thus we come to the conclusion that the hypothesis expressed at the start must be wrong: Requirements 5 and 2 are inconsistent if the calibration is performed with respect to the final linear ordering $\mathbf{L}^{\{s_K\}}$.

Summary

1. The optimal calibration for a set of evaluative properties S_k is generally not respected in the optimal calibration for a set S_{k+1}.

2. Requirement 5 has to be enforced, i.e. the calibration has to be changed in such a way as to make the Requirement fulfilled.

3. Generally speaking, the calibration of a classification procedure depends on the sequence of evaluative properties.

4. In order to fulfill Requirement 2 calibration on final linear ordering must be forsaken.

The consequence is that the calibration has to be performed for each scenario separately (Corrolary 1).

Furthermore, assume a parameter p_i appears in different threat scenario and the definitions (function) of the threat scenarios are such that the parameter p_i is coarse-grained, too. Then, as the calibration is performed for each scenario separately, the parameter will be calibrated for each threat scenario it appears in separately. Hence the parameter will have a separate coarse-graining for each threat scenario (Corrolary 3).

This is somewhat surprising, as one could require the coarse-graining of a parameter to be uniform, i.e. independent of the threat scenario it appears in.

Appendix F

Definition of the Sequel Operator $S(\prec)$

In order to introduce the sequel operator, we must first define two additional order relations; the two *compositions* of $\prec$ and $\sim$. The relation $\sim$ was briefly introduced in Section 3.3.4, whereas $\prec$ is the strict form of the component-wise order $\preceq$. The composition can take two forms, the left ($\prec^-$) and right ($\prec^+$) compositions

$$\prec^- = (\sim)(\prec) : x \prec^- y \text{ if } x \sim a \prec y \text{ for some } a \tag{F.1}$$

$$\prec^+ = (\prec)(\sim) : x \prec^+ y \text{ if } x \prec b \sim y \text{ for some } b \tag{F.2}$$

In order to better visualise the compositions (F.1) and (F.2), it is useful to view x, y, a, b as *intervals*. An interval is the outcome of a measurement which has a beginning and an end. Two intervals are comparable if an interval x ends before the another interval y starts: $x \prec y$. For any other situation the two are incomparable: e.g. if x begins after y begins but ends before y ends, then the intervals x and y are incomparable.

Now consider the situation where x is incomparable to y, but there is an interval a that it is also incomparable to x but is smaller then y. Then it is possible to differentiate between x and y since $a \prec y$ but $a \sim x$. The composition of the relation $\prec$ and $\sim$ can now be used to order x and y

In Figure F.1 both compositions are depicted. With the goal of using the composition as a tool to order x and y, it is necessary that the composition relations holding between the two elements are not contradicting. I.e. for the purpose of ordering x

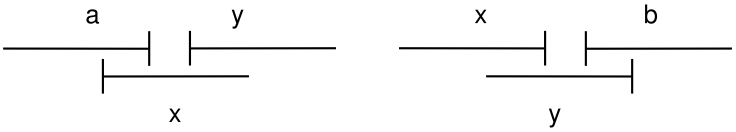

Figure F.1: In both situations x and y are incomparable since in both cases x has not ended when y starts. For the same reasons $a \sim x$ and $b \sim y$. But $a \prec y$ and $x \prec b$

and y with the help of the compositions $\prec^-$ and $\prec^+$ it is of use if we have $x \prec^- y$ and/or $x \prec^+ y$, but never $y \prec^- x$ nor $x \prec^+ y$. The sequel operator reflects precisely this idea.

$$S_0(\prec) := [\prec^- \cup \prec^+] \cap \mathrm{not}[\succ^- \cup \succ^+] \tag{F.3}$$

In words: for x, y $xS_0(\prec)y$ if either $x \prec^- y$ or $x \prec^+ y$ but *in no case* $x \succ^- y$ or $x \succ^+ y$. Thus the sequel operator ensures that the composition is used exclusively as a tool to compare two incomparable elements if among the possible compositions the resulting order is not contracting.

The sequel operator $S_0(\prec)$ maps the event-space to an extension of the linear ordering $\mathbf{L}^R$ and thus fulfills the requirements made in the formal outline of the classification procedure (see Section 4.4), with the exception of Requirement 3 (three-valuedness), which is readily fulfilled by introducing the appropriate cut-off values. Note that the sequel operator also works on $\mathbf{3} \times \mathbf{3} \times \mathbf{3}$ and beyond (without proof).

Appendix G

Preclusion of Sequences

After the exclusions stated in Section 7.3.1, the following sequences are left

$$\begin{gathered} cbad, cbda, dacb, dabc \\ cadb, cabd, dbac, dbca \\ cdab, cdba, dcab, dcba \end{gathered}$$

Elimination of sequences starting with cb and da
Now suppose we start with Criterion c and then continue with Criterion b. In this case, we first ensure that no elements of $Hset$ are erroneously classified in "red". In Criterion b we then ensure that of the elements in $Hset$ (which are now all classified in "green" and "orange"), as many as possible are classified in "green". So all elements of $Hset$ are classified in "green", as the previous criteria contain nothing to prevent that from happening. Hence any sequence that starts with the sequence cb could equally well start simply with b. Hence we can rule such sequences out, as — by analogy — we can rule out any sequence starting with Criterion d and a.

Elimination of sequences starting with ca and db
The sequence c, a translates into: first we ensure that no element in $Hset$ is classified as "red", then next we ensure that as many elements from $Rset$ as possible are classified as "red". The sequence cd on the other hand ensures first that no element in $Hset$ is classified as "orange" and then that no element in $Rset$ is classified as $Rset$. The Requirements made in Chapter 4.4 do preclude neither sequence ca nor the sequence cd. We may feel, however, that the sequence c, d is more appropriate to our problem, since we first consider the two criteria which deal with preventing false classification, and then proceed in a second step to ensure optimality. If we considered a sequence that starts with ca, then the hierarchy between prevention

of falses and reaching of positives would be broken. The same argument holds for excluding any sequence starting with db.

Interchangeability of Criterion c and Criterion d
Criterion c and d ensure that no elements of $Hset$ get classified as "red" and no elements of *Rset* get classified as "green", respectively. But in order to ensure that e.g. an element of $Hset$ is not classified as "red" it can be either classified as "green" or as "orange". If this is ensured for every element in $Hset$ (Criterion c) it is never the case that causally an element of *Rset* will be classified as "green". Conversely, ensuring that no element in *Rset* is classified as "green" (Criterion d) never forces an element of $Hset$ to be falsely classified. As a consequence

- Criteria c and d can both be fulfilled totally
- Criteria c and d are interchangeable without affecting the calibration

Non-interchangeability of Criterion a and Criterion b
Criteria a and b on the other hand ensure that a maximal number of elements are classified *correctly.* Now to ensure that an element in $Hset$ is classified correctly, i.e. as "green" (Criterion b) it may well be that an element from *Rset* is consequently classified as "orange", e.g. the element is smaller then the correctly classified element from $Hset$. This precludes the possibility for that specified element in *Rset* to be classifed as "red". As a consequence the sequence of Criteria b and *a cannot be* interchanged without affecting the calibration.

Appendix H

List of Chemicals

H.1 High Production Volume Chemicals

substance	$\tau/days$	$log(\tau/[\tau])$	rho/km	$log(K_{ow})$
Ethylene	1	0.11	1545	1.13
Propylene	0	-0.5	761	1.77
Methyl t-butyl ether	3	0.48	2317	0.94
Ethylene dichloride	38	1.58	8264	1.48
Benzene	7	0.83	3518	2.13
Vinyl chloride	1	0.1	1520	1.62
Ethylbenzene	1	0.07	1464	3.15
Styrene	0	-0.84	512	2.95
Methanol	4	0.64	1064	-0.77
o-Xylene	1	-0.22	1054	3.12
m-Xylene	0	-0.45	803	3.2
p-Xylene	1	-0.24	1031	3.15
Formaldehyde	4	0.58	343	0.35
Terephthalic acid DME	14	1.16	5150	2.25
Terephthalic acid	4	0.6	192	-2.49
Ethylene oxide	47	1.67	8474	-0.3
Toluene	1	0.14	1599	2.73
Cumene	1	0.11	1530	3.66

Ethylene glycol	6	0.8	536	−1.36
Acetic Acid	3	0.51	173	−3.42
Phenol	2	0.25	232	1.45
Propylene oxide	10	1.01	3670	0.03
1,3-Butadiene	0	−0.91	478	1.99
Isobutylene	0	−0.79	544	2.34
Acrylonitrile	5	0.68	2694	0.25
Vinyl acetate	0	−0.46	779	0.73
Acetone	8	0.92	2990	−0.24
Butyraldehyde	0	−0.37	797	0.88
Cyclohexane	1	0.04	1426	3.44
Adipic acid	3	0.51	173	−3.48
Nitrobenzene	0	−0.56	480	1.85
Bisphenol A	3	0.54	173	3.32
Caprolactam	3	0.51	187	0.66
n-Butyl alcohol	2	0.37	1016	0.88
Isopropyl alcohol	3	0.48	1121	0.05

The values for *Henry* constants of four relatively polar compounds were corrected taking into account their pKa values. Also, only their corresponding salts (i.e. non-dissociated) forms at pH 8 were considered (The pH of marine water). These chemicals are *acetic acid* (pKa 4.76); *adipic acid* (pKa 4.44); *terephtalic acid* (pKa 3.51) and *phenol* (pKa 9.9). The *Henry* constant was changed by using the formula given in Schwarzenbach's textbook [91]. The adjustments are significant for *acetic acid*, *adipic acid* and *terephtalic acid* (around 3 orders of magnitude). For *phenol*, the adjustment was negligeable.

For the *octanol/water partition coefficient* K_{ow}, the values had not been corrected for $pH = 8$ initially and hence the calibration has been conducted on the original, uncorrected values. The adjustments are identical to the ones for the *Henry* constant, and hence can have a significant impact on calibration. The calibration has been tested for these adjustments, and seen not to be influenced by the changes in K_{ow}-values for the three substances *acetic acid* , *adipic acid* and *terephtalic acid* (not shown). The K_{ow} values listed are the new, corrected values. The values used in calibration are -3.42 for *acetic acid*, -3.48 for *adipic acid*, -2.49 for *terephtalic acid* and 1.46 for phenol.

H.2 Globally Regulated Chemicals

substance	$\tau/days$	$log(\tau/[\tau])$	rho/km	$log(K_{ow})$
Aldrin	14	1.14	620	6.5
Dieldrin	10	1	1770	5.4
Endrin	52	1.71	1788	5.2
cis-Chlordane	21	1.31	1739	6.61
trans-Chlordane	138	2.14	1740	6.62
trans-Nonachlor	410	2.61	1758	6.78
Heptachlor Epoxide	10	0.98	1720	4.98
$4,4'$-DDT	3843	3.58	2702	6.91
$2,4'$-DDT	3254	3.51	2731	6.79
$4,4'$-DDD	655	2.82	2740	6.02
$4,4'$-DDE	323	2.51	2700	6.51
D48 ($2,3,7,8$-TCDD)	368	2.57	2305	6.8
F83 ($2,3,7,8$-TCDF)	9466	3.98	8810	6.53
F94 ($1,2,3,7,8$-PeCDF)	38939	4.59	11200	6.79
D54 ($1,2,3,7,8$-PeCDD)	60837	4.78	2907	7.56
F114 ($2,3,4,7,8$-PeCDF)	52440	4.72	11226	6.92
D66 ($1,2,3,4,7,8$-HxCDD)	184814	5.27	3350	7.8
D67 ($1,2,3,6,7,8$-HxCDD)	416018	5.62	3135	8.2
D70 ($1,2,3,7,8,9$-HxCDD)	416019	5.62	3136	8.22
F131 ($1,2,3,4,6,7,8$-HpCDF)	1230477	6.09	12521	7.92
D85 (OCDD)	116959	5.07	3061	8.21
F 135 (OCDF)	6612156	6.82	11670	8.6
Hexachlorobenzene	641	2.81	18578	5.73
Mirex	43751	4.64	19801	6.89
PCB 28	24	1.38	3717	5.62
PCB 52	83	1.92	4568	6.09
PCB 101	1693	3.23	6599	6.8
PCB 118	1113	3.05	6604	7.12
PCB 138	64115	4.81	9404	7.44
PCB 153	119917	5.08	9411	7.75

PCB 180	1330855	6.12	11201	8.27
Toxaphene	455	2.66	2500	5.9
CFC-11	28599	4.46	19900	2.53
CFC-12	280	2.45	17201	2.16
CFC-21	728	2.86	19179	1.55
CFC-22	5088	3.71	19811	1.08
F-142*b*	5717	3.76	19816	1.6
Carb. Tetrachloride	3621	3.56	19817	2.83
Carb. Tetrafluoride	20676	4.32	19851	1.18
Hexafluoroethane	20677	4.32	19852	2
HFC 134*a*	20688	4.32	19853	1.68
HFC 152*a*	244	2.39	16700	0.75
Sulphur hexafluoride	20679	4.32	19880	1.68

H.3 Test Chemicals

substance	$\tau/days$	$log(\tau/[\tau])$	rho/km	$log(K_{ow})$
alpha-HCH	79.6	1.90	6209	3.80
beta-HCH	101.7	2.01	2169	3.78
gamma-HCH	113.6	2.06	5332	3.72
endosulfan	0.2	−0.70	428	3.83
carbaryl	4.3	0.63	200	2.36
carbofuran	3.8	0.58	188	2.32
HMDS	6.0	0.78	3321	4.20
OMCTS (D4)	8.2	0.91	3883	5.10
DMCPS (D5)	5.4	0.73	3139	5.20
HBB	636.3	2.80	4577	6.07
DBDE	1937.2	3.29	1648	5.24

Appendix I

All Results for the Calibration

I.1 Threat Scenario *Pandora*

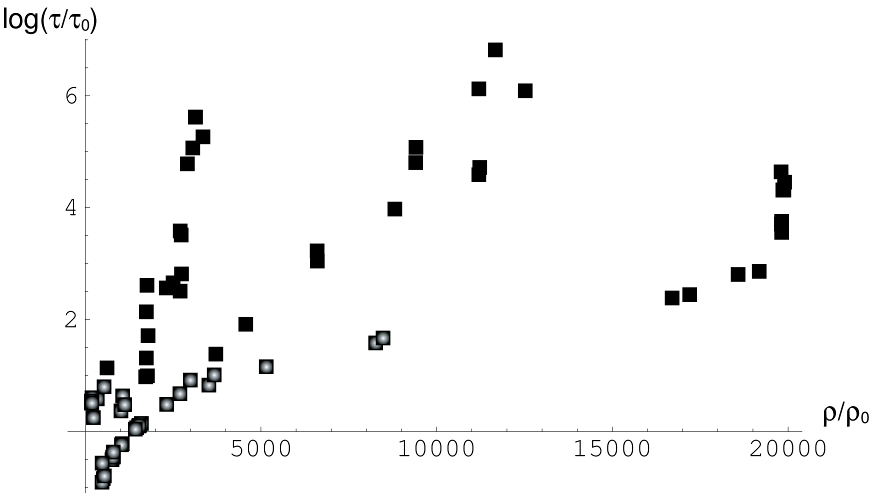

Figure I.1: The reference chemicals in *Hset* are depicted in black-framed grey, the reference chemicals in *Rset* in black. The axis $log(\tau/\tau_0)$ and ρ/ρ_0 define the plane reflecting the threat scenario *Pandora*.

I.2 Results *Pandora*, *cdab*

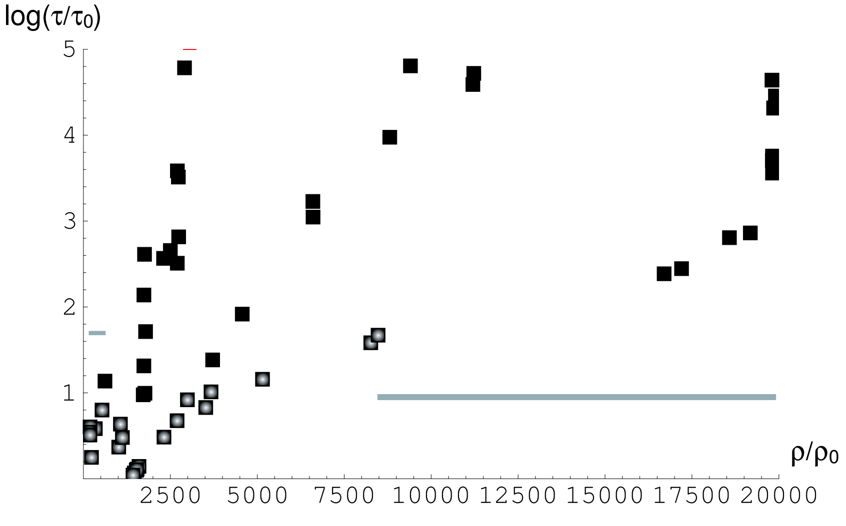

Figure I.2: Solution for α_2, α_4 for the sequence *cdab*

I.3 Results *Pandora*, *cdab* and *cdba* Combined

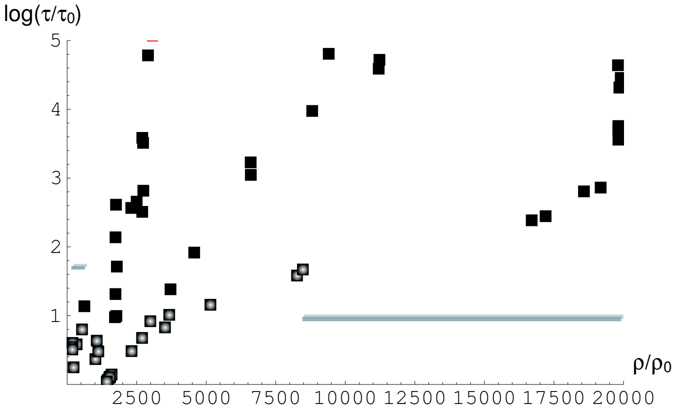

Figure I.3: Solution from the geometrical algorithm for the areas α_2 and α_4 combined for the two sequences of conditions *cdba* and *cdab*. The results coincide.

I.4 Threat Scenario *Bioaccumulation*

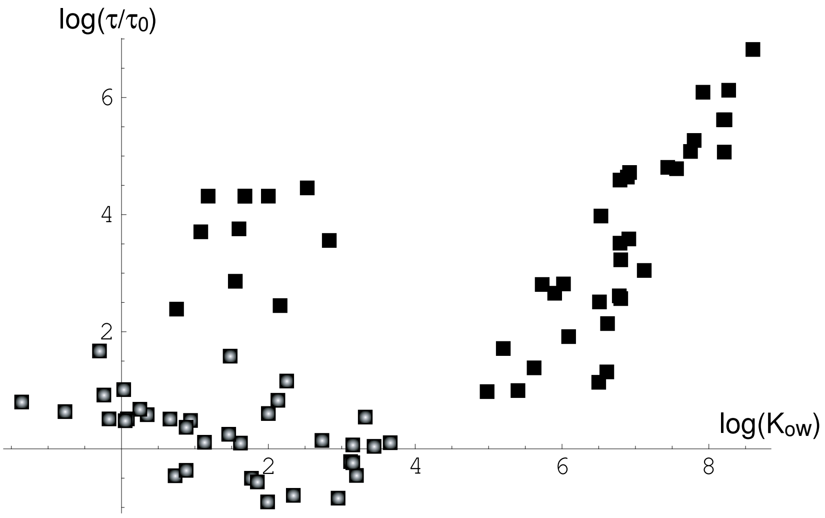

Figure I.4: The reference chemicals in $Hset$ are depicted in black-framed grey, the reference chemicals in $Rset$ in black. The axis τ and K_{ow} define the threat scenario *Bioaccumulation*.

class	Pandora	Bioaccumulation	overall
"red"	$Rset\setminus\{\text{"orange"}\cap Rset\}$	$Rset\setminus\{\text{"orange"}\cap Rset\}$	$Rset\setminus\{\text{"orange"}\cap Rset\}$
"orange"	propylene oxide, terephtalic acid DME, ethylene dichloride, ethylene oxide, aldrin, cis-chlordane, dieldrin, heptachlor epoxide, PCB 28	ethylene dichloride, terephtalic acid DME	ethylene dichloride, terephtalic acid DME
"green"	$Hset\setminus\{\text{"orange"}\cap Hset\}$	$Hset\setminus\{\text{"orange"}\cap Hset\}$	$Hset\setminus\{\text{"orange"}\cap Hset\}$

Figure I.5: Classification of reference chemicals in $Hset$ and $Rset$

Appendix J

Code Algorithm

J.1 Functions

```
(*ELEMENTS e_1,e_2,...,e_n OF elementset ARE INSERTED IN combset, WHICH
HAS THE FORM {{{e_1},{}},{{e_2},{}},....,{{e_n},{}}}*)

comb1:=For[combset:={};i=1,i<(Length[elementset]+1),i++,
For[k=i,k<(Length[elementset]+1),k++,combset=
Append[combset,{{elementset[[i]],{}},{elementset[[k]],{}},{}}]]]

(*FOR A SET arbset, SUPREMUM; (MIN(X),MIN(Y))*)

setmeet:={arbset[[i,1,1]],arbset[[i,2,2]]}

(*SET maximal OF MAXIMAL ELEMENTS A SET arbset*)

maximal:=For[maximalset:={};k=1,k<(Length[arbset]+1),k++,
For[n:={};j=1,j<(Length[arbset]+1),j++,
If[arbset[[j,1]]>=arbset[[k,1]],
If[arbset[[j,2]]>=arbset[[k,2]],l[j]={arbset[[j,1]],arbset[[j,2]]},],];

If[Length[l[j]] >1,n=Append[n,l[j]];l[j]=.,]];
a[j]:=Length[n];If[a[j]<2,maximalset=Append[maximalset,n[[1]]],]]

(*MINIMAL ELEMENTS OF AN ARBITRARY SET*)
```

```
minimal:=For[minimalset:={};k=1,k<Length[arbset]+1,k++,
For[n:={};j=1,j<Length[arbset]+1,j++,
If[arbset[[j,1]]<=arbset[[k,1]],
If[arbset[[j,2]]<=arbset[[k, 2]],
lj={arbset[[j,1]],arbset[[j,2]]},Null],Null];

If[Length[lj]>1,n=Append[n,lj];(lj)=.,Null]];
a[j]:=Length[n];If[a[j]<2,minimalset=Append[minimalset,n[[1]]],Null]]

(*SUPREMUM OF DOWNSET *)

supdownsetalgo02:=For[m:={};k=1,k<(Length[downset]+1),k++,
For[n:={};j=1,j<(Length[downset]+1),j++,
If[downset[[j,1]]>=downset[[k,1]],
If[downset[[j,2]]>=downset[[k,2]],l[j]={downset[[j,1]],downset[[j,2]]},],];

If[Length[l[j]]>1,n=Append[n,l[j]];l[j]=.,]];
a[j]:=Length[n];If[a[j]<2,m=Append[m,n],]];

e:=Sort[m];supremum:={e[[(Length[m]),1,1]],e[[1,1,2]]}

(*SUPREMUM (supremum) OF A SET upset*)

infupsetalgo:=For[m:={};k=1,k<(Length[upset]+1),k++,
For[n:={};j=1,j<(Length[upset]+1),j++,
If[upset[[j,1]]<=upset[[k,1]],
If[upset[[j,2]]<=upset[[k,2]],l[j]={upset[[j,1]],upset[[j,2]]},],];

If[Length[l[j]]>1,n=Append[n,l[j]];l[j]=.,]];
a[j]:=Length[n];If[a[j]<2,m=Append[m,n],]];

e:=Sort[m];infimum:={e[[1,1,1]],e[[(Length[m]),1,2]]};


(*DOWNSET (downset) OF A SET numset, WITH RESPECT TO A GENERAL POINT (point);
```

```
FOR NON-STRICT RELATION*)

downsetalgoGeneral := For[downs := {}; r=1, r<Length[numset]+1, r++,
If[numset[[r, 1]]<= point[[1]],
If[numset[[r, 2]] <= point[[2]], downs=Append[downs, numset[[r]]];
downset := Sort[downs], Null], Null]]

(*DOWNSET (downset) OF A SET numset, WITH RESPECT TO THE POINT
zielset[[i,f,2]]; FOR NON-STRICT RELATION*)

downsetalgo:=For[downs:={};r=1,r<(Length[numset]+1),r++,
If[numset[[r,1]]<=zielset[[i,f,2,1]],
If[numset[[r,2]]<=zielset[[i,f,2,2]],downs=Append[downs,numset[[r]]];
downset:=Sort[downs],],]]

(*DOWNSET (downset) OF A SET numset, WITH RESPECT TO THE POINT
zielset[[i,f,2]]; FOR STRICT RELATION*)

strictdownsetalgo:=For[downs:={};r=1,r<(Length[numset]+1),r++,
If[numset[[r,1]]<zielset[[i,f,2,1]],
If[numset[[r,2]]<zielset[[i,f,2,2]],downs=Append[downs,numset[[r]]];
downset:=Sort[downs],],]]

(*DOWNSET (downset) OF A SET numset, WITH RESPECT TO THE POINT
zielset[[i,f,2]]; FOR NON-STRICT RELATION. FOR A VARIABLE i2*)

downsetalgo2:=For[downs:={};r=1,r<(Length[numset]+1),r++,
If[numset[[r,1]]<=zielset[[i2,f,2,1]],
If[numset[[r,2]]<=zielset[[i2,f,2,2]],downs=Append[downs,numset[[r]]];
downset:=Sort[downs],],]]

(*UPSET (upset) OF A SET numset, WITH RESPECT TO A GENERAL POINT (point);
FOR NON-STRICT RELATION*)

upsetalgoGeneral := For[ups:={}; r=1,r<Length[numset]+1, r++,
If[numset[[r,1]]>=point[[1]],
```

```
If[numset[[r, 2]]>=point[[2]], ups=Append[ups, numset[[r]]];
upset:=Sort[ups],Null],Null]]

(*UPSET (upset) OF A SET numset, WITH RESPECT TO THE POINT zielset[[i,f,2]];
FOR NON-STRICT RELATION*)

upsetalgo:=For[ups:={};r=1,r<(Length[numset]+1),r++,
If[numset[[r,1]]>=arbset[[i,f,1,1]],
If[numset[[r,2]]>=arbset[[i,f,1,2]],ups=Append[ups,numset[[r]]];
upset:=Sort[ups],],]]

(*UPSET (upset) OF A SET numset, WITH RESPECT TO THE POINT zielset[[i,f,2]];
FOR STRICT RELATION*)

strictupsetalgo:=For[ups:={};r=1,r<(Length[numset]+1),r++,
If[numset[[r,1]]>arbset[[i,f,1,1]],
If[numset[[r,2]]>arbset[[i,f,1,2]],ups=Append[ups,numset[[r]]];
upset:=Sort[ups],],]]

(*UPSET (upset) OF A SET numset, WITH RESPECT TO THE POINT zielset[[i,f,2]];
FOR NON-STRICT RELATION. FOR A VARIABLE i2*)

upsetalgo2:=For[ups:={};r=1,r<(Length[numset]+1),r++,
If[numset[[r,1]]>=zielset[[i2,f,1,1]],
If[numset[[r,2]]>=zielset[[i2,f,1,2]],ups=Append[ups,numset[[r]]];
upset:=Sort[ups],],]]

(*EXTRACTION OF CHARACTERISTIC NUMBER FROM A SOLUTION-SET arbset;
AT POSITION position*)

extract:=For[n:={};i=1,i<(Length[arbset]+1),i++,
n=Append[n,arbset[[i,3,position]]]];

(*RETAINS THE SOLUTIONS INSIDE arbset WITH HIGHEST CHARACTERISTIC NUMBER,
DELETES THE OTHER SOLUTIONS. OUTPUT: zielset*)
```

```
solutionextract:=For[zielset:={};i=1,i<(Length[arbset]+1),i++,
If[arbset[[i,3,position]]==Max[n],zielset=Append[zielset,arbset[[i]]],]]

(*FOR NUMBER OF AREAS lcorners AND TOTAL OF i SOLUTIONS, SOLUTIONS FOR
P_2 CAN BE LISTED AS SOLUTION 1 COMBINED WITH SOLUTION 1 to SOLUTION
lcorners FOR P_4, THEN SOLUTION 2 TO SOLUTION lcorners, ETC. UNTIL SOLUTION
lcorners IS COMBINED WITH ITSELF. relevantnumber YIELDS THE i WHERE - IN THIS
LIST - A SET OF SOLUTIONS STARTS WITH A NEW SOLUTION FOR P_2 STARTS*)

rightneighbour:=
  If[i>(lcorners),
    If[i>(2*lcorners-1),
      If[i>(3*lcorners-3),
        If[i>(4*lcorners-6),
          If[i>(5*lcorners-10),
            If[i>(6*lcorners-15),
              ,
              relevantnumber=6*lcorners-14],
            relevantnumber=5*lcorners-9],
          relevantnumber=4*lcorners-5],
        relevantnumber=3*lcorners-2],
      relevantnumber=2*lcorners],
    relevantnumber=lcorners+1]

(*AS FOR rightneighbour, EXCEPT relevantnumber INDICATES THE i WHERE THE
PREVIOUS, DIFFERENT SOLUTION FOR P_2 OCCURED.*)

leftneighbour:=If[i<(lcorners),
    If[i>(2*lcorners),
      If[i>(3*lcorners-2),
        If[i>(4*lcorners-5),
          If[i>(5*lcorners-9),
            If[i>(6*lcorners-14),
              ,
              relevantnumber=6*lcorners-13],
            relevantnumber=5*lcorners-8],
```

```
     relevantnumber=4*lcorners-4],
    relevantnumber=3*lcorners-3],
   relevantnumber=1],
  relevantnumber= 1]
```

J.2 Algorithm Criterion c

```
(*DATA-UPLOAD. HERE EXAMPLE pandora: CONTAINS DATA FOR PARAMETERS tau,rho: FIRST
THE SET filtertype IS OF THE FORM {{Rset},{Hset},{Allset}}, WHERE Rset, Hset AND
Allset ALL CONTAIN TUPLE OF logtau,rho-VALUES FOR EACH ELEMENTS THEY CONTAIN.
Allset  IS THE UNION OF Hset AND Rset.*)

filtertype:=pandora;Rset:=filtertype[[1]];Hset:=filtertype[[2]];
Allset:=filtertype[[3]]

(*MINIMAL AND MAXIMAL VALUES FOR BOTH x AND y*)

minAllX:=Sort[Allset][[1,1]];minAllY:=Sort[Allset,#1[[2]]>#2[[2]]&][[
Length[Allset],2]];maxAllX:=Sort[Allset][[Length[Allset],1]];maxAllY:=
Sort[Allset,#1[[2]]>#2[[2]]&][[1,2]]

(*SET maximal OF MAXIMAL ELEMENTS IN Hset; THEN ADDS EACH ONE PERIPHERAL
ELEMENT IN x- AND y-DIRECTION. THE PERIPHERAL ELEMENTS ARE THE FIRST AND THE
LAST CORNERS GOING FROM LEFT TO RIGTH IN THE PLANE, SEE Figure 2,
THE CORNERS c_1 AND c_3.*)

arbset:=Hset;maximal;extmaximalH:=Sort[maximalset];
;extmaximalH=Append[extmaximalH,{extmaximalH[[Length[extmaximalH],1]],minAllY}
];extmaximalH=Prepend[extmaximalH,{minAllX,extmaximalH[[1,2]]}]

(*SET corners CONTAINING ALL CORNERS FROM WHICH SOLUTION-SETS ARE BUILT*)

For[pairsetc:={};i=1,i<Length[extmaximalH],i++,pairsetc=Append[pairsetc,
```

```
{extmaximalH[[i]],extmaximalH[[i+1]]}]];arbset:=pairsetc;corners:={};
For[i=1,i<(Length[pairsetc]+1),i++,corners=Append[corners,setmeet]];arbset=.

(*SET zielset CONTAINING ALL POSSIBLE SOLUTIONS FOR CRITERION C. THE SET
CONTAINS SOLUTIONS OF THE FORM {{area alpha2},{area alpha4}}. THE
SOLUTIONS EXPRESSED WITH THE AREAS C_1,C_2,... (EACH CORRESPONDING TO
CORNERS c_1,c_2,... SEE Figure 2) ARE:  {{area alpha2=C_1},{area alpha4=C_1}},
{{area alpha2=C_1},{area alpha4=C_2}},...,{{area alpha2=C_2},{area alpha4=C_1}},
...,{{area alpha2=C_n},{area alpha4=C_n}}*)

elementset:=corners;comb1;zielset=combset;numset:=Hset;arbset=zielset;

(*CALCULATES THE CHARACTERISTIC NUMBER OF CRITERION C, I.E.
THE CARDINALITY  OF THE SUBSET OF Hset FOR WHICH CRITERION IS
FULFILLED. (JUST A CHECK AS ALL ELEMENTS OF Hset FULFILL THE
CONDITION.) THE CHARACTERISTIC NUMBER IS INSERTED AS A THIRD
SUBSET OF A SOLUTION SET: {{area alpha2},{area alpha4},{charact. number}}*)

For[i=1,i<(Length[zielset]+1),i++,For[laenge:=0;f=1,f<3,f++,strictupsetalgo;
laenge=(Length[upset]+laenge);zielset=ReplacePart[zielset,{maxAllX,maxAllY},
{i,f,2}]];zielset=ReplacePart[zielset,Append[zielset[[i,3]],(Length[Hset]=
laenge)],{i,3}]];areasetc=zielset

(*CALCULATES areasetc BY EXTRACTING THE SOLUTIONS WITH MAXIMAL CHARACTERISTIC
NUMBER: ALL SOLUTIONS*)

arbset=areasetc;position=1;extract;solutionextract;areasetc=zielset
```

J.3 Algorithm Criterion d

```
(*sumofi THE LENGTH OF THE SOLUTION SET areasetc, COMBINATORICALLY
FROM NUMBER OF SOLUTION AREAS lcorners*)
```

```
lcorners=Length[corners];For[sumofi=0;l=1,l<(lcorners+1),l++,sumofi=sumofi+l]

(*THE ALGORITHM FOR CRITERIA b IS SPLIT INTO TWO PARTS*)

(*1ST ADJUSTMENT: ADJUSTMENT FOR ELEMENTS OF Rset THAT ARE BELOW THE AREA
areasetc[[i,f]], AS WELL AS FOR ELEMENTS OF Rset THAT ARE TO THE LEFT OF
THE AREA areasetc[[i,f]]*)

zielset=areasetc;

(*ADJUSTMENTS FOR ELEMENTS BELOW areasetc[[i,f]]*)

f=1;
For[i=1,i<(Length[areasetc]+1),i++,
rightneighbour;If[i==sumofi,rightneighbourpoint={maxAllX,areasetc[[i,f,1,2]]},
rightneighbourpoint={areasetc[[relevantnumber,f,1,1]],areasetc[[i,f,1,2]]}];

numset=Rset;point={maxAllX,rightneighbourpoint[[2]]};downsetalgoGeneral2;
downset=Sort[downset];

If[downset\not={},If[downset[[1,1]]<zielset[[i,f,2,1]],
zielset=ReplacePart[zielset,downset[[1,1]],{i,f,2,1}],],]]

f=2;
For[i=1,i<(Length[areasetc]+1),i++,
rightneighbour;If[i==sumofi,rightneighbourpoint={maxAllX,areasetc[[i,f,1,2]]},
rightneighbourpoint={areasetc[[relevantnumber,f,1,1]],areasetc[[i,f,1,2]]}];

numset=Rset;point={maxAllX,rightneighbourpoint[[2]]};downsetalgoGeneral2;
downset=Sort[downset];

If[downset\not={},If[downset[[1,1]]<zielset[[i,f,2,1]],
zielset=ReplacePart[zielset,downset[[1,1]],{i,f,2,1}],],]]

(*ADJUSTMENTS FOR ELEMENTS TO THE LEFT OF areasetc[[i,f]]*)
```

```
f=1;

For[i=1,i<(Length[areasetc]+1),i++,
leftneighbour;If[i==1,leftneighbourpoint=areasetc[[i,f,1]],leftneighbour-
point={areasetc[[i,f,1,1]],areasetc[[relevantnumber,f,1,2]]}];

numset=Rset;point={leftneighbourpoint[[1]],maxAllY};downsetalgoGeneral2;
downset=Sort[downset,#1[[2]]<#2[[2]]&];

If[downset\not={},If[downset[[1,2]]<zielset[[i,f,2,2]],zielset=
ReplacePart[zielset,downset[[1,2]],{i,f,2,2}],],]];

f=2;
For[i=1,i<(Length[areasetc]+1),i++,
leftneighbour;If[i==1,leftneighbourpoint=areasetc[[i,f,1]],leftneighbour-
point={areasetc[[i,f,1,1]],areasetc[[relevantnumber,f,1,2]]}];

numset=Rset;point={leftneighbourpoint[[1]],maxAllY};downsetalgoGeneral2;
downset=Sort[downset,#1[[2]]<#2[[2]]&];

If[downset\not={},If[downset[[1,2]]<zielset[[i,f,2,2]],
zielset=ReplacePart[zielset,downset[[1,2]],{i,f,2,2}],],]]

(*DEPTH OF areaset[[i,f]] IS INCREASED BY ONE, IN ORDER TO ALLOW FOR
SEVERAL SUB-AREAS TO BUILD-UP THE areaset[[i,f]]. SEE FIGURE NON-
RECTANGULAR SOLUTION IN FIGURE 6.*)

For[i=1,i<(Length[zielset]+1),i++,For[f=1,f<3,f++,
zielset=ReplacePart[zielset,{zielset[[i,f]]},{i,f}]]]

(*2ND ADJUSTMENT: ADJUSTMENT FOR ELEMENTS OF $Rset$ THAT ARE LARGER THAN
areaset[[i,f,1,1]], THE LOWER-LEFT CORNER OF AN AREA AND SMALLER THAN
areaset[[i,f,1,2]], THE UPPER-RIGHT CORNER OF THE AREA. AREA THEREBY DEFINED
IS areaRset. ONLY THE MINIMAL ELEMENTS OF Rset INSIDE OF areaRset ARE
CONSIDERED.*)
```

```
numset=Rset;
For[i=1,i<(Length[zielset]+1),i++,
  For[f=1,f<3,f++,
    point=zielset[[i,f,1,1]];upsetalgoGeneral2;upsetprov=upset;
    point=zielset[[i,f,1,2]];downsetalgoGeneral2;
    areaRset=Sort[Intersection[upsetprov,downset]];
arbset=areaRset;minimal;minimalset;minimalset=Sort[minimalset,#1[[1]]<#2[[1]]&];
Length[minimalset];pzwei=zielset[[i,f,1,2]];

(*MULTIPLICATION OF THE AREA IN ORDER TO HAVE ENOUGH SUB-AREAS TO ACCOMMODATE
 ALL SOLUTIONS CORRESPONDING TO THE NUMBER OF ELEMENTS IN minimalset.
 THEN ADJUSTS THE INDIVIDUAL AREAS WITH RESPECT TO THE ELEMENTS IN RSET
 THAT ARE INSIDE THE AREA.*)

For[c=1,c<(Length[minimalset]+1),c++,gam=Insert[zielset[[i,f]],
zielset[[i,f,1]],1];zielset=ReplacePart[zielset,gam,{i,f}]];

(*ADJUSTMENT OF SUB-AREAS. FIRST FOR FIRST SUB-AREA (CORRESPONDING TO FIRST
ELEMENT IN minimalset), SECOND FOR LAST SUB-AREA (CORRESPONDING TO
(Length[minimalset]+1) OF THE ELEMENTS OF minimalset) AND FINALLY ADJUSTMENT
OF ALL INTERMEDIATE SUB-AREAS. IN TOTAL (Length[minimalset]-1+2) SUB-AREAS,
I.E. ADJUSTMENTS*)

If[minimalset=={},,zielset=ReplacePart[zielset,{minimalset[[1,1]],
pzwei[[2]]},{i,f,1,2}]];

If[minimalset=={},,zielset=ReplacePart[zielset,{pzwei[[1]],
minimalset[[Length[minimalset],2]]},{i,f,(Length[minimalset]+1),2}]];

For[h=1,h<(Length[minimalset]),h++,zielset=ReplacePart[zielset,
{minimalset[[(h+1),1]],minimalset[[h,2]]},{i,f,h+1,2}]],]];

(*FINAL ELIMINATION OF SOLUTIONS FOR WHICH THERE ARE ELEMENTS OF Rset
IN THE DOWNSET OF areaset[[i,f,1]], I.E. OF THE LOWER-LEFT CORNER OF ANY OF
THE TWO SOLUTIONS f=1 OR f=2.*)
```

```
For[i=1,i<(Length[areasetd]+1),i++,
(zahli=Length[areasetd]+1- t);
point=areasetd[[i,2,1]]; downsetalgoGeneral2;
If[downset\not={},areasetd=Delete[areasetd,zahli],point=areasetd[[i,1,1]];
downsetalgoGeneral2;If[downset\not={},areasetd=Delete[areasetd,zahli],]]]

(*0 IS INSERTED AS VALUE FOR NUMBER OF ELEMENTS OF Hset CLASSIFIED AS "RED"*)

zielset=ReplacePart[zielset,Append[zielset[[i,3]],0],{i,3}]]
```

J.4 Algorithm Criterion b

```
zielset:=areasetd;arbset:=areasetd;solset:=areasetd;numset:=Hset;
zahlI=Length[areasetd];

For[i=1,i<(zahlI+1),i++,

(*ADDITION OF ONE SOLUTION PER EACH i SOLUTION, TO ALLOW FOR
SEQUENCE OF ADJUSTMENT (f=1,f=2), AND AS WELL AS FOR SEQUENCE (f=2,f=1)*)

zielset=Append[zielset,zielset[[i]]];

(*ADJUSTMENTS FOR AREAS WITH RESPECT TO condition B.
SEQUENCE OF ADJUSTMENT (f=1,f=2)*)

f=1;

(*ADJUSTMENT TO THE LEFT OF THE AREA areasetd[[i,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA h SEPARATELY. AREAS TO THE LEFT
IS OBTAINED AS THE COMPLEMENT OF TWO DOWNSETS. ALSO, LENGTH OF
DOWNSET THAT IS CORRECTLY CLASSIFIED IS CALCULATED laengeleft*);

For[h=1,h<(Length[zielset[[i,f]]]+1),h++,
```

```
point=zielset[[i,f,h,2]];downsetalgoGeneral2;downsetprov=
Sort[downset,#1[[2]]>#2[[2]]&];
point={zielset[[i,f,h,2,1]],zielset[[i,f,h,1,2]]};downsetalgoGeneral2;
downsetleft=Sort[Complement[downsetprov,downset],#1[[2]]>#2[[2]]&];

If[downsetleft\not={},zielset=ReplacePart[zielset,
downsetleft[[1,2]],{i,f,h,1,2}],];

laengeleft=Length[downsetleft];

(*ADJUSTMENT FOR THE AREA BELOW AREA areasetd[[i,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA SEPARATELY. AREAS BELOW IS OBTAINED
AS THE COMPLEMENT OF TWO DOWNSETS. ALSO, LENGTH OF
DOWNSET THAT IS CORRECTLY CLASSIFIED IS CALCULATED; laengedown.
ALSO, LENGTH OF OVERALL DOWNSET; laengedownset[h]*)

point=zielset[[i,f,h,2]];downsetalgoGeneral2;downsetprov=
Sort[downset,#1[[2]]>#2[[2]]&];
point={zielset[[i,f,h,1,1]],zielset[[i,f,h,2,2]]};downsetalgoGeneral2;
downsetdown=Sort[Complement[downsetprov,downset],#1[[1]]>#2[[1]]&];

If[downsetdown\not={},zielset=ReplacePart[zielset,
downsetdown[[1,1]],{i,f,h,1,1}],];

laengedown=Length[downsetdown];

laengedownset[h]=laengeleft+laengedown;

(*COMPUTE MAX-VALUE FOR THE LENGTH OF ANY DOWNSET FOR THE h-SOLUTIONS:
YIELDS laengemax*)

If[h==1,laengemax=laengedownset[h],];
If[laengedownset[h]>laengedownset[h-1],laengemax=laengedownset[h],];
point=zielset[[i,f,h,2]];downsetalgoGeneral;

];
```

```
(*SO FAR CORRECT BOUNDS HAVE BEEN IDENTIFIED FOR EACH OF THE h-AREAS.
THE LAST SECTION ALSO COMPUTED THE MAXIMAL LENGTH OF ANY h-AREA, I.E.
THE BEST PERFORMANCE ATTAINED BY ANY OF THE AREAS. THE AREA WHICH
PERFORMS THIS TASK OPTIMALLY, HOWEVER, IS NOT IDENTIFIED YET. IN THE NEXT
SECTION, THE LENGTH OF THE CORRESPONDING ELEMENT OF Hset IS COMPUTED.
THE h-AREAS FOR WHICH THIS NUMBER IS NOT MAXIMAL, ARE ELIMINATED. ALSO,
THE REFERENCE SET downsetprov7 IS COMPUTED, WHICH WILL BE USED IN f=2 TO
COMPUTE THE SUBSET OF Hset WHICH IS NON-REDUNDANT.*)

(*ELIMINATION OF SUB-OPTIMAL AREAS IN i,f*)

For[t=1,t<(Length[zielset[[i,f]]]+1),t++,zahlt=(Length[zielset[[i,f]]+1-t]);
point=zielset[[i,f,zahlt,2]];downsetalgoGeneral2;If[Length[downset]<
laengemax,zielset=Delete[zielset,{i,f,zahlt}],];];

(*COMPUTE SET OF RELEVANCE WITH RESPECT TO REDUNDANCY FOR f=2*)

point=zielset[[i,f,(Length[zielset[[i,f]]]),2]];downsetalgoGeneral2;
downsetprov7=downset;

(*NUMBER OF ELEMENTS IN Hset WHICH ARE CORRECTLY CLASSIFIED AFTER
ADJUSTMENT OF f=1 IS RECORDED AS laengeoverall*)

laengeoverall=laengemax;

f=2;

(*ADJUSTMENT TO THE LEFT OF THE AREA areasetd[[i,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA h SEPARATELY. AREAS TO THE LEFT
IS OBTAINED AS THE COMPLEMENT OF TWO DOWNSETS. ALSO, LENGTH OF
DOWNSET THAT IS CORRECTLY CLASSIFIED IS CALCULATED laengeleft*);

For[h=1,h<(Length[zielset[[i,f]]]+1),h++,
point=zielset[[i,f,h,2]];downsetalgoGeneral2;downsetprov=
Sort[downset,#1[[2]]>#2[[2]]&];
```

```
point={zielset[[i,f,h,2,1]],zielset[[i,f,h,1,2]]};downsetalgoGeneral2;
downsetleft=Sort[Complement[downsetprov,downset],#1[[2]]>#2[[2]]&];

(*downsetleft IS CORRECTED FOR ELEMENTS WHICH ARE ALREADY IN downsetdown OF f=1*)

downsetleft=Sort[Complement[downsetleft,downsetprov7],#1[[2]]>#2[[2]]&];

If[downsetleft\not={},zielset=ReplacePart[zielset,
downsetleft[[1,2]],{i,f,h,1,2}],];

laengeleft=Length[downsetleft];

(*ADJUSTMENT FOR THE AREA BELOW AREA areasetd[[i,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA SEPARATELY. AREAS BELOW IS OBTAINED
AS THE COMPLEMENT OF TWO DOWNSETS. ALSO, LENGTH OF
DOWNSET THAT IS CORRECTLY CLASSIFIED IS CALCULATED; laengedown.
ALSO; LENGTH OF OVERALL DOWNSET*)

point=zielset[[i,f,h,2]];downsetalgoGeneral2;downsetprov=
Sort[downset,#1[[2]]>#2[[2]]&];
point={zielset[[i,f,h,1,1]],zielset[[i,f,h,2,2]]};downsetalgoGeneral2;
downsetdown=Sort[Complement[downsetprov,downset],#1[[1]]>#2[[1]]&];

(*downsetdown CORRECTED FOR ELEMENTS WHICH ARE ALREADY IN
downsetdown OF f=1*)

downsetdown=Sort[Complement[downsetdown,downsetprov7],#1[[1]]>#2[[1]]&];

If[downsetdown\not={},zielset=ReplacePart[zielset,
downsetdown[[1,1]],{i,f,h,1,1}],];

laengedown=Length[downsetdown];

laengedownset[h]=laengeleft+laengedown;

(*COMPUTE MAX-VALUE FOR THE LENGTH OF ANY DOWNSET FOR THE h-
```

```
SOLUTIONS: YIELDS laengemax*)

If[h==1,laengemax=laengedownset[h],];
If[laengedownset[h]>laengedownset[h-1],laengemax=laengedownset[h],];

];

(*ELIMINATION OF SUB-OPTIMAL AREAS IN i,f, AS EARLIER FOR f=1. THE CHOICE
OF THE h-AREA IN f=1 TO SUPPLY THE REFERENCE SET IS THE LAST OF THE h's,
SINCE ALL OTHER REFERENCE SETS ARE SUBSETS OF THIS REFERENCE SETS.*)

For[t=1,t<(Length[zielset[[i,f]]]+1),t++,
zahlt=(Length[zielset[[i,f]]+1t]);point=zielset[[i,f,zahlt,2]];
downsetalgoGeneral2;downset=Complement[downset,downsetprov7];

If[Length[downset]<laengemax,zielset=Delete[zielset,{i,f,zahlt}],]];

(*DELETE NOT MEANINGNFUL AREAS FOR f=2, DUE TO BEING SMALLER IN X -
DIRECTION THAN THE AREAS IN f=1, WHICH IS NOT POSSSIBLE*)

(*FIRST THE NON - MEANINGFUL AREAS ARE DELETED, THEN IN A SECOND STEP THE
AREA IS ADJUSTED. FOR i2 I HAVE INTERCHANGED THE SEQUENCE, SINCE ELSE
THE CODE--AFTER DELETION OF A NON - MEANINGFUL SOLUTION--LOOKS FOR A
SOLUTION TO ADJUST, WHICH IS NO LONGER EXISTING. HERE, HOWEVER, I HAVE NOT
MADE THE CHANGE YET.*)


For[r=1,r<(Length[zielset[[i,f]]]+1),r++,
zahlr= (Length[zielset[[i,f]]]+1-r);
If[zielset[[i,f,zahlr,2,1]]<zielset[[i,f-1,1,1,1]],zielset=
Delete[zielset, {i, f, zahlr}],];If[zielset[[i,f,zahlr,1,1]]<
zielset[[i,f-1,1,1,1]],zielset=ReplacePart[zielset,
zielset[[i,f-1,1,1,1]],{i,f,zahlr,1,1}],]]


(*COMPUTING OVERALL LENGTH, I.E. OF f=1 AND f=2. AND INSERT THIS
```

```
CHARACTERISTIC NUMBER*)

laengetotal=laengeoverall+laengemax;

zielset=ReplacePart[zielset,Append[zielset[[i,3]],laengetotal],{i,3}];

(*AFTER ADJUSTMENT IN THE SEQUENCE (f=1,f=2), NOW ADJUSTMENTS FOR
THE SEQUENCE (f=2,f=1). TO THAT END, THE VARIABLE i IS CHANGED TO i2.*)

i2=i+zahlI;

f=2;

(*ADJUSTMENT TO THE LEFT OF THE AREA areasetd[[i2,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA h SEPARATELY. AREAS TO THE LEFT
IS OBTAINED AS THE COMPLEMENT OF TWO DOWNSETS. ALSO, LENGTH OF
DOWNSET THAT IS CORRECTLY CLASSIFIED IS CALCULATED laengeleft*);

For[h=1,h<(Length[zielset[[i2,f]]]+1),h++,
point=zielset[[i2,f,h,2]];downsetalgoGeneral2;downsetprov=
Sort[downset,#1[[2]]>#2[[2]]&];
point={zielset[[i2,f,h,2,1]],zielset[[i2,f,h,1,2]]};downsetalgoGeneral2;
downsetleft=Sort[Complement[downsetprov,downset],#1[[2]]>#2[[2]]&];

If[downsetleft\not={},zielset=ReplacePart[zielset,
downsetleft[[1,2]],{i2,f,h,1,2}],];

laengeleft=Length[downsetleft];

(*ADJUSTMENT FOR THE AREA BELOW AREA areasetd[[i2,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA SEPARATELY. AREAS BELOW IS OBTAINED
AS THE COMPLEMENT OF TWO DOWNSETS. ALSO, LENGTH OF
DOWNSET THAT IS CORRECTLY CLASSIFIED IS CALCULATED; laengedown.
ALSO; LENGTH OF OVERALL DOWNSET*)

point=zielset[[i2,f,h,2]];downsetalgoGeneral2;downsetprov=
```

```
Sort[downset,#1[[2]]>#2[[2]]&];
point={zielset[[i2,f,h,1,1]],zielset[[i2,f,h,2,2]]};downsetalgoGeneral2;
downsetdown=Sort[Complement[downsetprov,downset],#1[[1]]>#2[[1]]&];

If[downsetdown\not={},zielset=ReplacePart[zielset,
downsetdown[[1,1]],{i2,f,h,1,1}],];

laengedown=Length[downsetdown];

laengedownset[h]=laengeleft+laengedown;

(*COMPUTE MAX-VALUE FOR THE LENGTH OF ANY DOWNSET FOR THE h-SOLUTIONS:
YIELDS laengemax*)

If[h==1,laengemax=laengedownset[h],];
If[laengedownset[h]>laengedownset[h-1],laengemax=laengedownset[h],];
point=zielset[[i2,f,h,2]];downsetalgoGeneral;

];

(*SO FAR CORRECT BOUNDS HAVE BEEN IDENTIFIED FOR EACH OF THE h-AREAS.
THE LAST SECTION ALSO COMPUTED THE MAXIMAL LENGTH OF ANY h-AREA, I.E.
THE BEST PERFORMANCE ATTAINED BY ANY OF THE AREAS. THE AREA WHICH
PERFORMS THIS TASK OPTIMALLY, HOWEVER, IS NOT IDENTIFIED YET. IN THE NEXT
SECTION, THE LENGTH OF THE CORRESPONDING ELEMENT OF Hset IS COMPUTED.
THE h-AREAS FOR WHICH THIS NUMBER IS NOT MAXIMAL, ARE ELIMINATED. ALSO,
THE REFERENCE SET downsetprov7 IS COMPUTED, WHICH WILL BE USED IN f=2 TO
COMPUTE THE SUBSET OF Hset WHICH IS NON-REDUNDANT.*)

(*ELIMINATION OF SUB-OPTIMAL AREAS IN i2,f*)

For[t=1,t<(Length[zielset[[i2,f]]]+1),t++,zahlt=(Length[zielset[[i2,f]]+1-t]);
point=zielset[[i2,f,zahlt,2]];downsetalgoGeneral2;If[Length[downset]<
laengemax,zielset=Delete[zielset,{i2,f,zahlt}],];];

(*COMPUTE SET OF RELEVANCE WITH RESPECT TO REDUNDANCY FOR f=2*)
```

```
point=zielset[[i2,f,(Length[zielset[[i2,f]]]),2]];downsetalgoGeneral2;
downsetprov7=downset;

(*NUMBER OF ELEMENTS IN Hset WHICH ARE CORRECTLY CLASSIFIED AFTER
ADJUSTMENT OF f=1 IS RECORDED AS laengeoverall FOR THE*)

laengeoverall=laengemax;

f=1;

(*ADJUSTMENT TO THE LEFT OF THE AREA areasetd[[i2,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA h SEPARATELY. AREAS TO THE LEFT
IS OBTAINED AS THE COMPLEMENT OF TWO DOWNSETS. ALSO, LENGTH OF
DOWNSET THAT IS CORRECTLY CLASSIFIED IS CALCULATED laengeleft*);

For[h=1,h<(Length[zielset[[i2,f]]]+1),h++,
point=zielset[[i2,f,h,2]];downsetalgoGeneral2;downsetprov=
Sort[downset,#1[[2]]>#2[[2]]&];
point={zielset[[i2,f,h,2,1]],zielset[[i2,f,h,1,2]]};downsetalgoGeneral2;
downsetleft=Sort[Complement[downsetprov,downset],#1[[2]]>#2[[2]]&];

(*downsetleft IS CORRECTED FOR ELEMENTS WHICH ARE ALREADY IN downsetdown OF f=1*)

downsetleft=Sort[Complement[downsetleft,downsetprov7],#1[[2]]>#2[[2]]&];

If[downsetleft\not={},zielset=ReplacePart[zielset,
downsetleft[[1,2]],{i2,f,h,1,2}],];

laengeleft=Length[downsetleft];

(*ADJUSTMENT FOR THE AREA BELOW AREA areasetd[[i2,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA SEPARATELY. AREAS BELOW IS OBTAINED
AS THE COMPLEMENT OF TWO DOWNSETS. ALSO, LENGTH OF
DOWNSET THAT IS CORRECTLY CLASSIFIED IS CALCULATED; laengedown.
ALSO; LENGTH OF OVERALL DOWNSET*)
```

```
point=zielset[[i2,f,h,2]];downsetalgoGeneral2;downsetprov=
Sort[downset,#1[[2]]>#2[[2]]&];
point={zielset[[i2,f,h,1,1]],zielset[[i2,f,h,2,2]]};downsetalgoGeneral2;
downsetdown=Sort[Complement[downsetprov,downset],#1[[1]]>#2[[1]]&];

(*downsetdown CORRECTED FOR
ELEMENTS WHICH ARE ALREADY IN downsetdown OF f=1*)

downsetdown=Sort[Complement[downsetdown,downsetprov7],#1[[1]]>#2[[1]]&];

If[downsetdown\not={},zielset=ReplacePart[zielset,
downsetdown[[1,1]],{i2,f,h,1,1}],];

laengedown=Length[downsetdown];

laengedownset[h]=laengeleft+laengedown;

(*COMPUTE MAX-VALUE FOR THE LENGTH OF ANY DOWNSET FOR
THE h-SOLUTIONS: YIELDS laengemax*)

If[h==1,laengemax=laengedownset[h],];
If[laengedownset[h]>laengedownset[h-1],laengemax=laengedownset[h],];

];

(*ELIMINATION OF SUB-OPTIMAL AREAS IN i2,f, AS EARLIER FOR f=1. THE CHOICE
OF THE h-AREA IN f=1 TO SUPPLY THE REFERENCE SET IS THE LAST OF THE h's,
SINCE ALL OTHER REFERENCE SETS ARE SUBSETS OF THIS REFERENCE SETS.*)

For[t=1,t<(Length[zielset[[i2,f]]]+1),t++,
zahlt=(Length[zielset[[i2,f]]+1t]);point=zielset[[i2,f,zahlt,2]];
downsetalgoGeneral2;downset=Complement[downset,downsetprov7];

If[Length[downset]<laengemax,zielset=Delete[zielset,{i2,f,zahlt}],]];
```

```
(*DELETE NOT MEANINGNFUL AREAS FOR f=2, DUE TO BEING SMALLER IN X -
DIRECTION THAN THE AREAS IN f=1, WHICH IS NOT POSSSIBLE*)

zahlrlaenge=Length[zielset[[i2, f]]];
For[r=1,r<(zahlrlaenge+1),r++,zahlr=(zahlrlaenge+1-r);
If[zielset[[i2,f,zahlr,1,2]]<zielset[[i2,f+1,1,1,2]],zielset=
ReplacePart[zielset, zielset[[i2,f+1,1,1,2]], {i2,f,zahlr,1,2}],];
If[zielset[[i2,f,zahlr,2,2]]<zielset[[i2,f+1,1,1,2]],zielset=
Delete[zielset,{i2,f,zahlr}],]]

(*COMPUTING OVERALL LENGTH, I.E. OF f=1 AND f=2. AND INSERT THIS
CHARACTERISTIC NUMBER*)

laengetotal=laengeoverall+laengemax;

zielset=ReplacePart[zielset,Append[zielset[[i2,3]],laengetotal],{i2,3}];

];areasetb=zielset

(*BEST SOLUTIONS WITH RESPECT TO CRITERION a ARE SELECTED.*)

arbset=areasetb;position=3;extract;solutionextract;areasetb=zielset
```

J.5 Algorithm Criterion a

```
zielset:=areasetb;arbset:=areasetb;numset:=Rset;zahlI=Length[areasetb];

For[i=1,i<(zahlI+1),i++,

(*ADDITION OF ONE SOLUTION PER EACH i SOLUTION, TO ALLOW FOR
SEQUENCE OF ADJUSTMENT (f=1,f=2), AND AS WELL AS FOR SEQUENCE (f=2,f=1)*)

zielset=Append[zielset,zielset[[i]]];
```

```
(*ADJUSTMENTS FOR AREAS WITH RESPECT TO condition A. SEQUENCE OF
ADJUSTMENT (f=1,f=2)*)

f=1;

(*ADJUSTMENT TO THE RIGHT OF THE AREA areasetd[[i,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA h SEPARATELY. AREAS TO THE RIGHT
IS OBTAINED AS THE COMPLEMENT OF TWO UPSETS. ALSO, LENGTH OF
UPSET THAT IS CORRECTLY CLASSIFIED IS CALCULATED laengeleft*);

For[h=1,h<(Length[zielset[[i,f]]]+1),h++,point=zielset[[i,f,h,1]];
upsetalgoGeneral2;upsetprov=Sort[upset,#1[[2]]>#2[[2]]&];
point={zielset[[i,f,h,1,1]],zielset[[i,f,h,2,2]]};upsetalgoGeneral2;
upsetright=Sort[Complement[upsetprov,upset],#1[[2]]<#2[[2]]&];

If[upsetright\not={},zielset=ReplacePart[zielset,upsetright[[1,2]],
{i,f,h,2,2}],];

laengeright=Length[upsetright];

(*ADJUSTMENT FOR THE AREA ABOVE AREA areasetd[[i,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA SEPARATELY. AREAS ABOVE IS OBTAINED
AS THE COMPLEMENT OF TWO UPSETS. ALSO, LENGTH OF UPSET THAT IS
CORRECTLY CLASSIFIED IS CALCULATED; laengeup. ALSO, LENGTH OF
OVERALL UPSET: laengeupset[h]*)

point=zielset[[i,f,h,1]];upsetalgoGeneral2;upsetprov=Sort[upset,#1[[2]]>
#2[[2]]&];point={zielset[[i,f,h,2,1]],zielset[[i,f,h,1,2]]};upsetalgo-
General2;upsetup=Sort[Complement[upsetprov,upset],#1[[1]]<#2[[1]]&];

If[upsetup\not={},zielset=ReplacePart[zielset,upsetup[[1,1]],{i,f,h,2,1}],];

laengeup=Length[upsetup];

laengeupset[h]=laengeright+laengeup;
```

```
(*COMPUTE THE MAXIMAL VALUE OF THE LENGTH OF ANY UPSET
INSIDE THE THE f=1-AREA, I.E. THE LONGEST FOR ANY h.*)

(*IN COMPARISON TO THE CASE IN STEP b HERE IT APPEARS THAT laengemax IS
ONLY CALCULATED FOR h=1*)

If[h==1,laengemax=laengeupset[h],];
If[laengedownset[h]>laengedownset[h-1],laengemax=laengedownset[h],];

];

(*SO FAR CORRECT BORDERS HAVE BEEN FOUND FOR EACH OF THE h-AREAS
INSIDE f=1-AREA. THE MAXIMAL VALUE THAT CAN BE ATTAINED BY SOME h-AREA
WITH RESPECT TO CORRECTLY CLASSIFIED ELEMENTS FROM Rset HAS ALSO BEEN
COMPUTED. THE NEXT SECTION DELETES ALL NON-OPTIMAL AREAS. ALSO THE
REFERENCE SETS OF ELEMENTS OF Rset ALREADY CLASSIFIED, IS COMPUTED;
upsetprov9. THE DELETION TAKES PLACE BACKWARDS, SO THAT THE RUNNING
NUMBER t IS NOT INTERFERED WITH.*)

(*ELIMINATION OF SUB-OPTIMAL AREAS IN i,f*)

For[t=1,t<(Length[zielset[[i,f]]]),t++,
zahlt=(Length[zielset[[i,f]]+1-t]);point=zielset[[i,f,zahlt,1]];
upsetalgoGeneral2;If[Length[upset]<laengemax,zielset=Delete[zielset,
{i,f,zahlt}],]];

(*ADJUST BOUNDARIES IF NON - MEANINGFUL. WE ADJUST WITH RESPECT TO
h=1 IN THE REFERENCE SET, I.E. IN h=1 in f=2 WHEN STARTING IN f=1 AND
MOVING TO f=2 AND IN h=1 IN f=1 WHEN STARTING IN f=2 AND MOVING TO f=1.*)

For[r=1, r<(Length[zielset[[i, f]]]+1),r++,
zahlr=(Length[zielset[[i,f]]]+1-r);
If[zielset[[i,f-1,1,2,2]]<zielset[[i,f,zahlr,2,2]],
zielset=ReplacePart[zielset,zielset[[i,f-1,1,2,2]],$
{i,f,zahlr,2,2}],];
```

```
If[zielset[[i,f,zahlr,1,1]]<zielset[[i,f-1,1,1,1]],
zielset=ReplacePart[zielset,zielset[[i,f-1,1,1,1]],
{i,f,zahlr,1,1}],]];

(*COMPUTE SET OF RELEVANCE WITH RESPECT TO REDUNDANCY FOR f=2*)

point=zielset[[i,f,1,1]];upsetalgoGeneral2;upsetprov9=upset;

(*NUMBER OF ELEMENTS IN $Rset$ WHICH ARE CORRECTLY CLASSIFIED AFTER
ADJUSTMENT OF f=1 IS RECORDED AS laengeoverall FOR THE*)

laengeoverall=laengemax;

f=2;

(*ADJUSTMENT TO THE RIGHT OF THE AREA areasetd[[i,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA h SEPARATELY. AREAS TO THE RIGHT
IS OBTAINED AS THE COMPLEMENT OF TWO UPSETS. ALSO, LENGTH OF
UPSET THAT IS CORRECTLY CLASSIFIED IS CALCULATED laengeright*);

For[h=1,h<(Length[zielset[[i,f]]]+1),h++,point=zielset[[i,f,h,1]];
upsetalgoGeneral2;upsetprov=Sort[upset,#1[[2]]>#2[[2]]&];
point={zielset[[i,f,h,1,1]],zielset[[i,f,h,2,2]]};upsetalgoGeneral2;
upsetright=Sort[Complement[upsetprov,upset],#1[[2]]>#2[[2]]&];

(*upsetleft CORRECTED FOR ELEMENTS WHICH ARE ALREADY IN upsetup OR
upsetdown OF f=1*)

upsetright=Sort[Complement[upsetright,upsetprov9],#1[[2]]<#2[[2]]&];

If[upsetright\not={},zielset=ReplacePart[zielset,upsetright[[1,2]],
{i,f,h,2,2}],];

laengeright=Length[upsetright];
```

```
(*ADJUSTMENT FOR THE AREA ABOVE AREA areasetd[[i,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA SEPARATELY. AREAS ABOVE IS OBTAINED AS THE
COMPLEMENT OF TWO UPSETS. ALSO, LENGTH OF UPSET THAT IS CORRECTLY
CLASSIFIED IS CALCULATED; laengeup. ALSO; LENGTH OF OVERALL UPSET*)

point=zielset[[i,f,h,1]];upsetalgoGeneral2;upsetprov=Sort[upset,
#1[[2]]>#2[[2]]&];point={zielset[[i,f,h,2,1]],zielset[[i,f,h,1,2]]};upsetalgo-
General2;upsetup=Sort[Complement[upsetprov,upset],#1[[1]]>#2[[1]]&];

(*upsetleft CORRECTED FOR ELEMENTS WHICH ARE ALREADY IN upsetup OR
upsetdown OF f=1*)

upsetup=Sort[Complement[upsetup,upsetprov9],#1[[1]]<#2[[1]]&];

If[upsetup\not={},zielset=ReplacePart[zielset,upsetup[[1,1]],
{i,f,h,2,1}],];

laengeup=Length[upsetup];

laengeupset[h]=laengeright+laengeup;


(*COMPUTE MAXIMAL VALUE FOR THE LENGTH OF ANY UPSET FOR ANY h-AREA.*)

If[h==1,laengemax=laengeupset[h],];
If[laengeupset[h]>laengeupset[h-1],laengemax=laengeupset[h],];

];

(*ELIMINATIONAS FOR f=1: ANY h-AREA WHICH DOESN'T YIELD laengemax IS
ELIMINATED.*)

For[t=1,t<Length[zielset[[i,f]]+1],t++,
zahlt=(Length[zielset[[i,f]]+1-t]);point=zielset[[i,f,zahlt,1]];
upsetalgoGeneral2;
```

```
upset=Complement[upset,upsetprov9];

If[Length[upset]<laengemax,zielset=Delete[zielset,{i,f,zahlt}],]];

(*COMPUTING OVERALL LENGTH, I.E. OF f=1 AND f=2. AND INSERT THIS
CHARACTERISTIC NUMBER*)

laengetotal=laengeoverall+laengemax;

zielset=ReplacePart[zielset,Append[zielset[[i,3]],laengetotal],{i,3}];

(*AFTER ADJUSTMENT IN THE SEQUENCE (f=1,f=2), NOW ADJUSTMENTS FOR
THE SEQUENCE (f=2,f=1). TO THAT END, THE VARIABLE i IS CHANGED TO i2.*)

i2=i+zahlI;

f=2;

(*ADJUSTMENT TO THE RIGHT OF THE AREA areasetd[[i,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA h SEPARATELY. AREAS TO THE RIGHT
IS OBTAINED AS THE COMPLEMENT OF TWO UPSETS. ALSO, LENGTH OF
UPSET THAT IS CORRECTLY CLASSIFIED IS CALCULATED laengeleft*);

For[h=1,h<(Length[zielset[[i,f]]]+1),h++,point=zielset[[i,f,h,1]];
upsetalgoGeneral2;upsetprov=Sort[upset,#1[[2]]>#2[[2]]&];
point={zielset[[i,f,h,1,1]],zielset[[i,f,h,2,2]]};upsetalgoGeneral2;
upsetright=Sort[Complement[upsetprov,upset],#1[[2]]<#2[[2]]&];

If[upsetright\not={},zielset=ReplacePart[zielset,upsetright[[1,2]],
{i,f,h,2,2}],];

laengeright=Length[upsetright];

(*ADJUSTMENT FOR THE AREA ABOVE AREA areasetd[[i,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA SEPARATELY. AREAS ABOVE IS OBTAINED
AS THE COMPLEMENT OF TWO UPSETS. ALSO, LENGTH OF UPSET THAT IS
```

```
CORRECTLY CLASSIFIED IS CALCULATED; laengeup. ALSO; LENGTH OF OVERALL
UPSET*)

point=zielset[[i,f,h,1]];upsetalgoGeneral2;upsetprov=Sort[upset,
#1[[2]]>#2[[2]]&];point={zielset[[i,f,h,2,1]],zielset[[i,f,h,1,2]]};upsetalgo-
General2;upsetup=Sort[Complement[upsetprov,upset],#1[[1]]<#2[[1]]&];

If[upsetup\not={},zielset=ReplacePart[zielset,upsetup[[1,1]],
{i,f,h,2,1}],];

laengeup=Length[upsetup];

laengeupset[h]=laengeright+laengeup;

(*COMPUTE THE MAXIMAL VALUE OF THE LENGTH OF ANY UPSET
INSIDE THE THE f=1-AREA, I.E. THE LONGEST FOR ANY h.*)

(*IN COMPARISON TO THE CASE IN STEP b HERE IT APPEARS THAT laengemax IS
ONLY CALCULATED FOR h=1*)

If[h==1,laengemax=laengeupset[h],];
If[laengeupset[h]>laengeupset[h-1],laengemax=laengeupset[h],];

];

(*SO FAR CORRECT BORDERS HAVE BEEN FOUND FOR EACH OF THE h-AREAS
INSIDE f=1-AREA. THE MAXIMAL VALUE THAT CAN BE ATTAINED BY SOME h-AREA
WITH RESPECT TO CORRECTLY CLASSIFIED ELEMENTS FROM Rset HAS ALSO BEEN
COMPUTED. THE NEXT SECTION DELETES ALL NON-OPTIMAL AREAS. ALSO THE
REFERENCE SETS OF ELEMENTS OF Rset ALREADY CLASSIFIED, IS COMPUTED;
upsetprov9. THE DELETION TAKES PLACE BACKWARDS, SO THAT THE RUNNING
NUMBER t IS NOT INTERFERED WITH.*)

(*ELIMINATION OF SUB-OPTIMAL AREAS IN i,f*)

For[t=1,t<(Length[zielset[[i,f]]]),t++,
```

```
zahlt=(Length[zielset[[i,f]]+1-t]);
point=zielset[[i,f,zahlt,1]];upsetalgoGeneral2;If[Length[upset]<
laengemax,zielset=Delete[zielset,{i,f,zahlt}],]];

(*COMPUTE SET OF RELEVANCE WITH RESPECT TO REDUNDANCY FOR f=2*)

point=zielset[[i,f,1,1]];upsetalgoGeneral2;upsetprov9=upset;

(*NUMBER OF ELEMENTS IN $Rset$ WHICH ARE CORRECTLY CLASSIFIED AFTER
ADJUSTMENT OF f=1 IS RECORDED AS laengeoverall FOR THE*)

laengeoverall=laengemax;

f=1;

(*ADJUSTMENT TO THE RIGHT OF THE AREA areasetd[[i,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA h SEPARATELY. AREAS TO THE RIGHT
IS OBTAINED AS THE COMPLEMENT OF TWO UPSETS. ALSO, LENGTH OF
UPSET THAT IS CORRECTLY CLASSIFIED IS CALCULATED laengeright*);

For[h=1,h<(Length[zielset[[i,f]]]+1),h++,point=zielset[[i,f,h,1]];
upsetalgoGeneral2;upsetprov=Sort[upset,#1[[2]]>#2[[2]]&];
point={zielset[[i,f,h,1,1]],zielset[[i,f,h,2,2]]};upsetalgoGeneral2;

upsetright=Sort[Complement[upsetprov,upset],#1[[2]]>#2[[2]]&];

(*upsetleft CORRECTED FOR ELEMENTS WHICH ARE ALREADY IN upsetup OR
upsetdown OF f=1*)

upsetright=Sort[Complement[upsetright,upsetprov9],#1[[2]]<#2[[2]]&];

If[upsetright\not={},zielset=ReplacePart[zielset,upsetright[[1,2]],
{i,f,h,2,2}],];

laengeright=Length[upsetright];
```

```
(*ADJUSTMENT FOR THE AREA ABOVE AREA areasetd[[i,f,h]]. ADJUSTMENT IS
PERFORMED FOR EACH SUB-AREA SEPARATELY. AREAS ABOVE IS OBTAINED AS THE
COMPLEMENT OF TWO UPSETS. ALSO, LENGTH OF UPSET THAT IS CORRECTLY
CLASSIFIED IS CALCULATED; laengeup. ALSO; LENGTH OF OVERALL UPSET*)

point=zielset[[i,f,h,1]];upsetalgoGeneral2;upsetprov=Sort[upset,#1[[2]]>
#2[[2]]&];point={zielset[[i,f,h,2,1]],zielset[[i,f,h,1,2]]};upsetalgoGeneral2;
upsetup=Sort[Complement[upsetprov,upset],#1[[1]]>#2[[1]]&];

(*upsetleft CORRECTED FOR ELEMENTS WHICH ARE ALREADY IN upsetup OR
upsetdown OF f=1*)

upsetup=Sort[Complement[upsetup,upsetprov9],#1[[1]]<#2[[1]]&];

If[upsetup\not={},zielset=ReplacePart[zielset,upsetup[[1,1]],{i,f,h,2,1}],];

laengeup=Length[upsetup];

laengeupset[h]=laengeright+laengeup;

(*COMPUTE MAXIMAL VALUE FOR THE LENGTH OF ANY UPSET FOR ANY h-AREA.*)

If[h==1,laengemax=laengeupset[h],];
If[laengeupset[h]>laengeupset[h-1],laengemax=laengeupset[h],];

];

(*ELIMINATIONAS FOR f=1: ANY h-AREA WHICH DOESN'T YIELD laengemax IS
ELIMINATED.*)

For[t=1,t<Length[zielset[[i,f]]+1],t++,
zahlt=(Length[zielset[[i,f]]+1-t]);point=zielset[[i,f,zahlt,1]];
upsetalgoGeneral2;upset=Complement[upset,upsetprov9];

If[Length[upset]<laengemax,zielset=Delete[zielset,{i,f,zahlt}],]];
```

```
(*ADJUST BOUNDARIES IF NON - MEANINGFUL. WE ADJUST WITH RESPECT
TO h=1 IN THE REFERENCE SET, I.E. IN h=1 in f=2 WHEN STARTING IN f=1
AND MOVING TO f=2 AND IN h=1 IN f=1 WHEN STARTING IN f=2 AND
MOVING TO f=1.*)

For[r=1,r<(Length[zielset[[i2,f+1]]]+1),r++,
zahlr=(Length[zielset[[i2,f+1]]]+1-r);
If[zielset[[i2,f,1,2,2]]<zielset[[i2,f+1,zahlr,2,2]],zielset=
ReplacePart[zielset,zielset[[i2,f,1,2,2]],{i2,f+1,zahlr,2,2}],];

If[zielset[[i2,f+1,zahlr,1,1]]<zielset[[i2,f,1,1,1]],zielset=
ReplacePart[zielset,zielset[[i2,f,1,1,1]],{i2,f+1,zahlr,1,1}],]];

(*COMPUTING OVERALL LENGTH, I.E. OF f=1 AND f=2. AND INSERT THIS
CHARACTERISTIC NUMBER*)

laengetotal=laengeoverall+laengemax;

zielset=ReplacePart[zielset,Append[zielset[[i,3]],laengetotal],{i,3}];

];areaseta=zielset
```

J.6 Algorithm Clean-Up and Graphics

```
areaseta

(*SOME SOLUTIONS CAN BE PRESENT IN MULTIPLE COPIES INSIDE OF THE
SOLUTION SET. TO ELIMINATE THIS REDUNDANCY, THESE SOLUTIONS ARE
DELETED FROM THE SOLUTION SET*)

areasetaMod=areaseta;

For[i=1,i<(Length[areaseta]+1),i++,laenge=Length[areasetaMod];
```

```
If[laenge<i,,For[k=i+1,k<(Length[areasetaMod]+1),k++,
If[areasetaMod[[i]]==areasetaMod[[k]],areasetaMod=ReplacePart[
areasetaMod,{},k]]]];

For[j=1,j<(Length[areasetaMod]+1),j++,If[areasetaMod[[j]]=={},
areasetaMod=Delete[areasetaMod,j]]

]];

(*areasetaMod REPRESENTS THE SET CONTAINING ALL OPTIMAL SOLUTIONS OF
CALIBRATION UNDER THE SEQUENCE OF CRITERIA cdba. EACH SOLUTION
CONSISTS OF TWO AREAS: ONE PERMISSIBLE FOR THE POINT P_2, AND ANOTHER
FOR THE POINT P_4. EACH SUCH AREA IS GIVEN BY TWO POINTS: THE LOWER-LEFT
CORNER AND THE UPPER-RIGHT CORNER OF THE AREA.*)
```

Bibliography

[1] M.-J.-A.-N. de Caritat de Condorcet, Essai sur l'Application de l'Analyse a la Probabilite des Decisions Rendues a la Pluralite des Vois, Paris, **1785**

[2] A. Sen, Collective choice and Social Welfare, Holden Day Inc., **1970**

[3] D. Saari, Decisions and Elections: Explaining the Unexpected, Cambridge University Press, **2001**

[4] J. J. Hickey and D. W. Anderson, Chlorinated Hydrocarbons and Eggshell Changes in Raptorial and Fish-eating Birds, *Science*, 162, **1968**, 271 – 273

[5] Stoffliche Risiken (Chapter 5), in: Welt im Wandel, Strategien zur Bewältigung globaler Umweltrisiken, Jahresgutachten 1998, Wissenschaftlicher Beirag der Bundesregierung Globale Umweltveränderungen, Springer, **1998**

[6] E. Halfon and M. Reggiani, On Ranking Chemicals for Environmental Hazard, *Environ. Sci. Technol.*, 20, **1986**, 1173 – 1179

[7] R. Brüggemann and B. Münzer, A Graph-theoretical Tool for Priority Setting of Chemicals, *Chemosphere*, (27), 9, **1993**, 1729 – 1736

[8] E. Halfon, S. Galassi, R. Brügemann and A. Provini, Selection of Priority Properties to Asses Environmental Hazard of Pesticides, *Chemosphere*, (33), 8, **1996**, 1543 – 1562

[9] R. Brüggemann and H.-G. Bartel, A Theoretical Concept to Rank Environmentally Significant Chemicals, *J. Chem. Inf. Sci.*, 39, **1999**, 211 – 217

[10] R. Brüggemann, C. Bücherl, S. Pudenz and D. W. W. Steinberg, Application of the Concept of Partial Order on Comparative Evaluation of Environmental Chemicals, *Acta hydochim. hydrobiol.*, 27, **1999**, 170 – 178

[11] S. O. Hansson and P. Sandin, Persistence, Liability to Bioaccumulate, and the Precautionary Principle, Background Paper for the Policy Forum, Falkenberg, 3 – 4 April **2001**

[12] The final document can be found at http://www.sussex.ac.uk/spru/environment/precaupripdfs.html

[13] Council Directive 67/548/EEC ”on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. *Official Journal of the European Communities*, *L*196 (10) : 1 – 5, **1967**

[14] Council Directive 92/32/EEC ”amending for the seventh time Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances.” *Official Journal of the European Communities*, *L*154 (35) : 1 – 29, **1992**

[15] Council Directive 93/67/EEC ”laying down the principles for assessment to man and the environment of substances notified in accordance with Council Directive 67/93/*EEC*”. *Official Journal of the European Communities*, *L*227 (36) : 9 – 18, **1993a**.

[16] Commission Regulation 1488/94 laying down the principles for the assessment of risks to man and the environment of existing substances in accordances with Council Regulation (EEC) No 793/93. *Official Journal of the European Communities*, *L*161 (37) : 3 – 11, **1994**

[17] European Union System for the Evaluation of Substances (EUSES). Principles and Structure. *Chemosphere*, *34*, 8, 1823 – 1836

[18] Latest proposal on chemicals regulation.

[19] Council Directive 79/831/EEC ”amending for the sixth time Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances.” *Official Journal of the European Communities*, *L*259 (22) : 10 – 28, **1979**

[20] EINECS: European Inventory of Existing Chemical Substances. Access through the European Chemicals Bureau's official site: http://ecb.jrc.it/existing-chemicals/

[21] TGD: Technical Guidance Document. "Risk assessment of notified substances: Technical Guidance Document in support of Commission Directive 93/67/EEC on risk assessment for new notified substances and Commission Regulation EC *No* 1488/94 on risk assessment for existing substances". Office for Official Publications of the European Communities, Luxembourg, **1996**

[22] White Paper; Strategy for a Future Chemicals Policy, Commission of the European Communities.

[23] K. Fenner, *Transformation Products in Environmental Risk Assessment: Joint and Secondary Persistence as New Indicators for the Overall Hazard of Chemical Pollutants*, PhD Thesis, Diss. ETH Nr. 14400, ETH Zürich, Zürich, Switzerland **2001**

[24] Proposal for a "Regulation of the European Parliament and of the Council"; concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency and amending Directive 1999/45/EC and Regulation (EC) on Persistent Organic Pollutants, etc. COM2003 0644(02)

[25] H. Berg and M. Scheringer, Problems in environmental risk assessment and the need for proxy measure *Fresenius Environmental Bulletin*, **1994**, *3*, 487 – 492

[26] M. Scheringer, Persistence and Spatial Range of Environmental Chemicals; New Ethical and Scientific Concepts for Risk Assessment, Wiley-VCH, **2002**

[27] K. Mathes and G. Winter, Ecological Risk Assessment and the regulation of chemicals: III. Balancing risks and benefits. *The Science of the Total Environment*, Suppl. part 2, **1993**, 1679 – 1687

[28] U. Müeller-Herold and M. Scheringer, Zur Umweltgefährdungsbewertung von Schadstoffen und Schadstoffkombinationen durch Reichweiten- und Persistenzanalyse, *Graue Reihe, Europäische Akademie Bad Neuenahr-Ahrweiler*, 18, **1999**

[29] *Non-Hazardous Products Proposals for Implementation of New Guidelines on Chemicals Policy*, Swedish Committee on new Guidelines on Chemicals Policy, SOU 2000: 53, Swedish Government Official Reports, **2000**

[30] G. H. Williamson and H. Hulpke, Das Vorsorgeprinzip: Internationaler Vergleich, Möglichkeiten und Grenzen, Lösungsvorschläge *Z. Umweltchm. Ökotox.*, **2000**, *12(1)*, 27 – 39

[31] Communication from the Comission on The Precautionary Principle, Commission of the European Communities, **2000**

[32] T. O'Riordan and J. Cameron, The History and Contemporary Significance of the Precautionary Principle, *in* Interpreting the Precautionary Principle, T. O'Riordan and J. Cameron (Editors), Earthscan, **1994**

[33] S. Böhmer-Christiansen, The Precautionary Principle in Germany - Enabling Government, *in* Interpreting the Precautionary Principle (T. O'Riordan and J. Cameron: Editors), Earthscan, **1994**, 31 – 60

[34] J. Hunt, article, in Interpreting the Precautionary Principle, T. O'Rordan and J. Cameron (Editors), Earthscan, **1994**

[35] M. A. Stoto, The precautionary principle and emerging biological risks: lessons from swine flue and HIV in blood products, *Public Health Rep.*, 117(6), 546−52, **2002**

[36] S. Böschen, Risikogenese - Prozesse gesellschaftlicher Gefahrenwahrnehmung: FCKW, DDT, Dioxin und Ökologische Chemie, Leske + Budrich, Opladen, **2000**

[37] M. F. Rogers, J. A. Sinden and T. De Lacy, The Precauationary Principle for Environmental Management: A Defensive-expenditure Application, *Journal of Environmental Mangement*, 51, **1997**, 343 – 360

[38] F. H. Knight, Risk, Uncertainty and Profit, London School of Economics and Political Science, **1990**

[39] Paris Convention: To Prevent and Eliminate Pollution and to Protect the Maritime Area Against the Adverse Effects of Human Activities. Cf. Council Decision 98/249/EC on the Conclusion of the Convention for the Protection of the Marine Environment of the North-Est Atlantic, **1992**

[40] E. Hey, Implications of the Revisions of the Oslo and Paris Conventions, *Marine Policy*, July **1991**, 244 – 254

[41] Rio Declaration on Environment and Development, **1992**

[42] Rethinking Risk and the Precautionary Principle, J. Morris (Editor), Butterworth Heinemann, **2000**

[43] K. R. Foster, P. Vecchia and M. H. Repacholi, Science and the Precautionary Principle, *Science*, 288, (5468), **2000**, 979 – 981

[44] D. Bodansky, Scientific Uncertainty and the Precautionary Principle, *Environment*, 33, 4 – 5, **1992**, 43 – 44

[45] P. Sandin, M. Peterson, S. O. Hansson, C. Rudèn and A. Juthe, Science and the Precautionary Principle, *Journal of Risk Research*, 5, (4), **2002**, 287 – 299

[46] Bagehot, Genetically Modified Government, *The Economist*, 29. May **1999**

[47] S. Holm and J. Harris, Precautionary Principle Stifles Discovery, *Nature*, 400, **1999**, 398

[48] G. Charnley, Annual Meeting: Past President's message: Risk Analysis Under Fire, *RISK Newsletter*, 1st quarter, **2000**

[49] C. Starr, The Precautionary Principle Versus Risk Analysis, *Risk Analysis*, (23), 1, **2003**, 1 – 3

[50] J. S. Gray and J. M. Bewers, Towards a Scientific Definition of the Precautionary Principle, *Marine Pollution Bulletin*, (32), 11, **1996**, 768 – 771

[51] B. Durodiè, Plastics Panics and the Perils of the Precautionary Principle, *The Economist*, **1999**

[52] Plastic panics: European Risk Regulation in the aftermath of BSE, in Rethinking Risk and the Precautionary Principle, J. Morris (Editor), Butterworth Heinemann, **2000**

[53] B. Durodiè, The True Cost of Precautionary Chemicals Regulation, *Risk Analysis*, (23), 2, **2003**, 389 – 398

[54] J. von Moltke, Commentary: The Precautionary Principle, *Environment*, (34), 3, **1992**, 3 – 4

[55] B. Hileman, Precautionary Principle, *Chemical and Engineering News*, February 9, 16 – 18, **1998**

[56] P. Sandin, Dimensions of the Precautionary Principle, *Human and Ecological Risk Assessment*, (5), 5, **1999**, 889 – 907

[57] M. Scheringer and M. Berg, Spatial and Temporal Range as Measures of Environmental Threat, *Fresenius Environmental bulletin*, 3, **1994**, 493 – 498

[58] M. S. Stephenson, An Approach to the Identification of Organic Compounds Hazardous to the Environment and Human Health, *Toxicology and Environmental Safety*, 1, **1977**, 39 – 48

[59] K. Ballschmiter, Globale Verteilung von Umweltchemikalien, *Nachrichten aus Chemie, Technik und Laboratorium*, 33, **1985**, 206 – 208

[60] K. Ballschmiter, Global Distribution of Organic Compounds, *Environmental Carcinogenesis and Ecotoxicological Reviews C*, 9, **1991**, 1 – 46

[61] K. Ballschmiter, Persistent, Ecotoxic and Bioaccumulating Compounds, *Pure & Applied Chemistry*, 68, **1998**, 1771 – 1780

[62] M. Scheringer, M. Berg, U. Müller-Herold, "Jenseits der Schadensfrage. Umweltschutz durch Gefährdungsbegrenzung" in: Was ist ein Schaden? Zur normativen Dimension des Schadensbegriffs in der Risikowissenschaft, vdf, ETH Zürich, **1994**

[63] M. Scheringer, Persistence and Spatial Range as Endpoints of an Exposure-based Assessment of Organic Chemicals, *Environmental Science and Technology*, 30, **1996**, 1652 – 1659

[64] M. Scheringer, Räumliche und Zeitliche Reichweite als Indikatoren zur Bewertung von Umweltchemikalien, PhD thesis, ETH Zürich, Zürich, Switzerland, **1996**

[65] M. Scheringer, Characterisation of the Environmental Distribution Behavior of Organic Chemicals by Means of Persistence and Spatial Range, *Environmental Science and Technology*, 31, **1997**, 2891 – 2897

[66] U. Müller-Herold, A Simple General Limiting Law for the Overall Decay of Organic Compounds with Global Pollution Potential, *Environmental Science and Technology*, 30, **1996**, 586 – 591

[67] R. Quartier, *Secondary Spatial Range of Transformation Products: A New Proxy Measure for the Spatial Extent of the Overall Chemical Impact of a Pollutant*, PhD Thesis, Diss. ETH Nr. 13754, ETH Zürich, Zürich, Switzerland**2000**

[68] U. Müller-Herold, Measures of Endangerment, *The Geneva Papers on Risk and Insurance*, (21), 80, **1996**, 383 – 392

[69] M. Scheringer, K. Hungerbühler and M. Matthies, The Spatial Scale of Organic Chemicals in Multimedia Fate Modelling, *Environmental Science and Pollution Research*, 8, **2001**, 150 – 155

[70] H. Kunreuther, P. R. Kleindorfer and O. Renn, Forward: Risk Management (Special Issue), *The Geneva Papers on Risk and Insurance*, (21), 80, **1996**, 307

[71] M. Scheringer, K. Mathes, G. Weidemann and G. Winter, Für einen Paradigmenwechsel bei der Bewertung ökologischer Risiken durch Chemikalien im Rahmen der staatlichen Chemikalienregulierung, *Zeitschrift für angewandte Umweltforschung*, 11, **1998**, 227 – 233

[72] O. Renn and A. Klinke, Prometheus Unbound - Challenges of Risk Evaluation, Risk Classification, and Risk Management, Working Report No. 153, Center of Technology Assessment in Baden-W"urttemberg; Stuttgart **1999**

[73] U. Müller-Herold and G. Nickel, A Closed Analytical Formula for the Characteristic Spatial Range of Persistent Organic Pollutants, *Ecological Modelling*, 126, **2000**, 191 – 200

[74] D. Connell, Bioaccumulation of Xenobiotic Compounds, CRC Pres, Boca Raton, **1990**

[75] M. Peterson, Transformative Decision Rules, *Erkenntnis*, (58), **2003**, 71 – 85

[76] M. Peterson, The Limits of Catastrophe Aversion, *Risk Analysis*, (22), 3, **2002**, 527 – 538

[77] Personal Communication with M. Peterson, June **2003**

[78] S. O. Hansson, The Limits of Precaution, *Foundations of Science*, 2, **1997**, 293 – 306

[79] Chem. and Eng. News, June 24, 38, **1996**; at http://pubs.acs.org/hotartcl/cenear/960624/prod.html

[80] B. A. Davey and H. A. Priestley, Introduction to Lattices and Order, Cambridge University Press 1990

[81] G. Birkhoff, Lattice Theory, American Mathematical Society; colloquium publications, volume 25, **1967**

[82] A. Rényi, Wahrscheinlichkeitsrechnung, Vierte Auflage, VEB Deutscher Verlag der Wissenschaften, **1973**

[83] D. G. Saari, Geometry of Voting, Springer-Verlag, **1994**

[84] D. G. Saari, Basic Geometry of Voting, Springer-Verlag, **1995**

[85] D. G. Saari, Chaotic Elections! A Mathematician Looks at Voting, American Mathematical Society, **2001**

[86] P. Fishburn, Interval Orders and Interval Graphs, A study of Partially Ordered Sets, Wiley-Interscience Series in Discrete Mathematics, John Wiley&Sons, **1985**

[87] K. J. Arrow, Social Choice and Individual Values, 2*nd* Edition, Yale University Press, **1951**

[88] H. Markovitz and L. A. Goodman, Social Welfare Functions Based on Individual Rankings, *The American Journal of Sociology*, 58, 3, 257 – 262, **1952**

[89] C. W. Doble, J.-P. Doignon, J.-C. Falmagne and P. Fishburn, Almost Connected Orders, *Order*, 18, 295 – 311, **2001**

[90] J.-L. Marichal, Aggregation Operators for Multicriteria Decision Aid, PhD Thesis, Liege University, **1999**

[91] R. P. Schwarzenbach, P. M. Geschwend and D. G. Imboden, Environmental Organic Chemistry, John Wiley & Sons, **1993**

[92] A. Reller, M. Braungart, J. Soth, O. von Uexkll, GAIA 2000, 9, 1324.

Acknowledgements

First and foremost, I wish to express my sincere gratitude to Prof. Ulrich Müller-Herold who not only inspired this thesis with his previous work, but also encouraged me to explore unexpected and often unchartered territories and patiently endured many hours of discussions. His deep understanding of formal problems as well as his eye for clarity and simplicity helped me recognise and navigate the more treacherous parts of my thesis. I also came to greatly appreciate his generosity which allowed me to conduct activities outside of my PhD studies, both academically and not.

I would also like to thank Prof. Nils-Eric Sahlin and Prof. Konrad Hungerbühler for accepting to be my co-examiners and for the time they invested in reading and discussing my manuscript.

During the four years of my thesis I had the opportunity to discuss my work and related topics with many people, who all contributed significantly to this thesis. My special thanks go to Dr. Martin Scheringer, Dr. Rainer Brüggemann, my former co-worker Dr. Robin Quartier and my co-worker Dr. Marco Morosini. My thanks also go to Simon Schwarz for studying my algorithm and for diligently reading my manuscript. Dr. Pierre Funck and Dr. Dieter Hennig generously and repeatedly offered their invaluable help with LaTeX.

I also would like to express my gratitude to ”Stiftung Forscher für die Umwelt” and the ”Stipendiendienst der ETH” for financing my thesis at the outset.

Most importantly, I would like to thank Tina, my parents, as well as my brother and my sister for their love and support over the last years.

CURRICULUM VITAE

PERSONAL DATA

Name: Olivier Schucht
Date of Birth: January 6, 1974
Nationalities: Swiss, French

EDUCATION

1986–1992	Freies Gymnasium Zürich, Zürich, Switzerland. Matura Type B
1993–1994	Studies in Mathematics, Philosophy, and Economics at Integral, Paris, France
1994–1999	Studies in Chemistry at the Federal Institute of Technology Zürich, Zürich, Switzerland and at the Department of Chemistry, Cambridge University, Cambridge, England Diploma as Dipl. natw. ETH Zürich Diploma Thesis: *Synthesis of a Novel Polymer-bound Hypervalent Iodine Reagents: Application to the Construction of a* 1,3-*Thiazole Library and the Synthesis of the Alkaloid* (±)-*Epimaritidine*
1999–2004	Ph.D. Studies in Ecological Risk Management at the Department of Environmental Sciences with Prof. U. Müller-Herold, ETH Zürich, Switzerland. Ph.D. Thesis: *Partial Order Theory in the Assessment of Environmental Chemicals: Formal Aspects of a Precautionary Assessment Approach*

PROFESSIONAL EXPERIENCE

summer 2000	Internship in the Investment Banking Division of Goldman Sachs, London, England
2001–2003	Teaching assistant, ETH Zürich, Switzerland
2002–2003	Part-time Associate, Translink Corporate Finance, Zürich, Switzerland
2003–2004	Consulting Agreement, Vallotech SA, Vallorbe, Switzerland