Scientific Phenomena and Patterns in Data

Pascal Ströing

München 2018

Scientific Phenomena and Patterns in Data

Pascal Ströing

Inauguraldissertation
zur Erlangung des Doktorgrades der Philosophie an der
Ludwig-Maximilians-Universität
München

vorgelegt von
Pascal Ströing
aus Gronau (Westfalen)

25. Juli 2018

Erstgutachter: Prof. Dr. Stephan Hartmann

Zweitgutachter: Prof. DDr. Hannes Leitgeb

Tag der mündlichen Prüfung: 10. Juli 2018

Abstract

Scientific theories can generally be understood to predict and explain claims about phenomena. If empirical data is available to support a theory, then phenomena are often identified with patterns in these data. Data is idiosyncratic to the particular contexts of its origin, but phenomena are usually described as not being so. Therefore, a direct identification of phenomena with patterns is intricate. More specifically, phenomena are sometimes described as being *real* features of an observer-independent world, whereas patterns are purely syntactical and are chosen by an observer out of arbitrarily many possible options to decompose a set of data. In this thesis I focus on this explanatory gap and propose a solution by explicating the notions of data and of patterns in, according to my knowledge, a novel way. I explain the relations between data, patterns in them and phenomena. It turns out that these three kinds of objects or properties or utterances are distinctively related to each other without any pair being identical.

In conclusion, I defend the view that the notion of phenomena does not have to be a realistic one, as authors such as Bogen and Woodward suggest. My notion of phenomena is explicated by clarifying the criteria of phenomenon selection, which is the demarcation between features of the empirical world that are *interesting* to scientists on the one side and the other features of the empirical world on the other side. An *outside world* might be real or caused by something real or not real, but the notion of phenomena is not a realistic one. Instead, I defend the view that the notion of phenomena is best explained by reference to a complex body of background assumptions, as well as cognitive and sensory capabilities of the relevant agents in science. This view has, in different forms, a tradition from Kant to Kuhn, to name only some highly influential philosophers, but central aspects might also be traced back to thinkers of classical antiquity. I aim to provide an account of phenomena with both descriptive and normative implications.

To defend this view I have to present a sufficiently powerful epistemological framework that does not leave the door open for anything unexplained to happen in processes of phenomenon selection and that could be responsible for letting completely observer-independent criteria for phenomenon selection in favour of phenomenon realism creep in. Strategies for the defence of phenomenon realism are usually to deflate the notion of realism by elevating human consciousness or interests to the status of instantiating or creating reality (Dennett), or to stipulate that reality is best explained by notions that neatly fit to our human ways of describing the world (*e.g.* ontic structural realism). In my view, both of these defence strategies are unjustified.

I defend my view in three successive steps. We need to have a notion of patterns to clarify the relation between patterns and phenomena. Patterns occur or are detected in data. Therefore, a notion of data needs to be explicated first. As opposed to some notions from the literature (Hacking; Leonelli) but with similarities to others (Suppes) data is non-material and purely mathematical. Data itself does not play a representing role, due to the problem of relation without relata. In a second step, I follow Grenander's epistemological approach to define patterns by a genuinely *constructive* (with an idiosyncratic notion of constructivity) mathematical approach in opposition to, for this application in philosophy, more influential notions of information in data from information theory (Shannon; Kolmogorov). In a third step, I argue that not only data and patterns are mathematical, scientific inferences that lead to phenomena selection and theory formation can, in principle, be expressed in purely mathematical terms, too. This view has classical proponents (Russell) and can even be defended empirically with reference to recent developments in artificial intelligences that are employed to mind games (*e.g.* Go; poker).

Under this view of a mathematised epistemology, scientific reasoning is independent from having or not having a specific human *consciousness* and there is no reason to believe that human agency is necessary to accomplish cognitive tasks of even our most accomplished scientific reasoning, as some authors contrarily imply (Searle). The empirical world presents itself to agents of science by material causal interactions with sensory organs or measurement devices. What patterns in observation data appear as phenomena and what as uninteresting depends on the shared body of the agents' background assumptions, as well as the agents' sensory and cognitive capabilities. This distinction is misunderstood by some due to the extreme complexity of human cognitive processes and not due to the real fabric of the world or the importance of consciousness for scientific reasoning.

Contents

List of Figures

Deutschsprachige Zusammenfassung

Ziel wissenschaftlicher Theorienbildung ist die empirisch zugängliche Welt zu beschreiben. Unabhängig davon, welche Kriterien für *gute* Theoriebildung maßgebend sind, stellt sich die grundsätzliche Frage, welcher Teil oder Aspekt der überbordend komplexen empirisch zugänglichen Welt durch Theorien beschrieben werden soll. Wissenschaftler würden antworten, dass sie sich für *Phänomene* interessieren und könnten Beispiele nennen: Änderungen von Aggregatzuständen in der Physik, Albinismus in der Biologie oder Hitlers Verhältnis zu Juden in der Geschichtsforschung. Die zentrale Frage, die ich in dieser Arbeit zu beantworten versuche, ist, ob man für die Erklärung der Unterscheidung zwischen Phänomenen und die für Wissenschaftler weniger interessanten Eigenschaften der empirischen Welt auf irgend eine Form von ontologischen Realismus zurückgreifen muss, oder ob es eine anti-realistische Erklärung auf erkenntnistheoretischer Grundlage für die Phänomenauswahl gibt. Mein Ergebnis wird sein, dass zweiteres der Fall ist.

In einer ersten Annäherung kann man sagen, dass Phänomene Muster in Daten sind. Der kürzliche Nachweis von Gravitationswellen durch Abbott u. a. (2016) als ganz bestimmte Kurve in einer langen Zahlenreihe, die Messergebnisse aus einer komplizierten kausal-mechanischen Apparatur darstellt, ist ein anschauliches Beispiel. Allerdings sind Muster zu ein und dem selben Phänomen in jedem Datensatz zumindest etwas anders. Es gibt sogar völlig unterschiedliche Muster zu ein und dem selben Phänomen, wie etwa bei Albinismus der Phänotyp (d. h. weißes Haar und helle Haut) und der Genotyp (d. h. eine bestimmte Genmutation). Das heißt, man kann die beobachtbaren Eigenschaften eines Phänomens keinesfalls einfach mit einem bestimmten Muster identifizieren, auch wenn Autoren wie Bogen und Woodward das in ihren einflussreichen Arbeiten allem Anschein nach implizieren. Darüber hinaus sind *Muster* etwas mathematisches, wobei die Phänomene selbst etwas echtes in der Welt sind. Ihr stabiles Auftreten in der von uns erforschbaren Welt legt nahe, dass sie durch irgendeinen bewusstseinsunabhängigen Prozess außer uns kausal verursacht werden. All dies macht eine Identifizierung von Phänomenen mit Mustern sogar ontologisch hochgradig prob-

lematisch. Die Klärung der Beziehung zwischen diesen beiden Konzepten, oder Klassen von Dingen oder Eigenschaften ist das zentrale Thema dieser Doktorarbeit.

Anstatt sich uns durch die tatsächliche Konstitution der von uns unabhängigen, aber empirisch erforschbaren Welt unmittelbar aufzudrängen, wählen wir die Phänomene unserer Wissenschaften, wie ich glaube, nach Kriterien, die an uns selbst liegen, aus. Genauer gesagt ist unsere Phänomenauswahl geprägt durch die Gesamtheit der wissenschaftlichen Annahmen, die wir als sehr gesichert für ausreichend adäquat akzeptieren, sowie unseren ganz menschlichen, natürlichen Voraussetzungen von sensorischen und kognitiven Fähigkeiten. Beides ist aus (vermutlich) biologischen Gründen eigenartig und in seiner Leistung beschränkt.—Man vergleiche die Situation nur einmal mit den kürzlich auf prominenter Bühne vorgeführten Denksportpartien zwischen menschlichen Go-Meistern und selbstlernenden Computerprogrammen, die ihre Gegner mit nicht direkt nachvollziehbaren Strategien schlagen konnten. Die Idee, dass wir die empirische Welt bereits auf der grundlegendsten Ebene der Beobachtung vor-ordnen ist nicht neu und wird bei Kant schon als Begriff der *Anschauung* diskutiert. Man kann sogar über Verwandtschaften mit dem antiken Begriffen von *Ideen* und *Substanzen* diskutieren, wenn man diese eher epistemisch als ontologisch begreift.

Die Klärung des Phänomen-Begriffs in dieser Arbeit ist deskriptiv in dem Sinne, dass das Ergebnis mit dem übereinstimmen soll, was Wissenschaftler im alltäglichen Sprachgebrauch als ‘Phänomen’ bezeichnen. Sie ist normativ in dem Sinne, dass der Begriff vollständig erklärt und in ein umfassendes erkenntnistheoretisches Rahmenwerk eingeordnet werden soll.

Um meine Position zu stärken, muss ich eine ausreichend überzeugende Erkenntnistheorie anbieten, die die Phänomenauswahl so gut erklärt, dass keine bedeutsamen Stellen unklar bleiben, die dafür verantwortlich gemacht werden könnten, dass sich realistische, das heißt von der menschlichen Beobachtung völlig unabhängige, aber für die Phänomenauswahl bedeutsame Kriterien einschleichen. Eine Strategie zur Verteidigung eines Phänomen-Realismus besteht üblicherweise darin, den Realismusbegriff im ausreichenden Maße zu deflationieren, indem man das menschliche Bewusstsein und menschliche Interessen in den Stand erhebt, *Realität* auf irgend eine Weise zu instantiieren oder zu kreieren (Dennett). Eine andere Strategie besteht darin zu behaupten, dass die tatsächliche, von uns unabhängige Außenwelt durch einen glücklichen Zufall genau so konstituiert ist, wie unsere besten Beschreibungen es nahelegen (z. B. ontischer Strukturrealismus). Aus meiner Sicht sind diese beiden Ansätze nicht zu vertreten.

Ich lege meine Position in drei inhaltlichen Schritten dar. Um die Beziehung zwischen Mustern in Daten und Phänomenen näher zu beleuchten, müssen wir explizieren, was genau Muster sind. Da Muster in Daten *erkannt* werden, muss zuerst geklärt werden, was genau Daten sind. Im Gegensatz zu anderen explizierenden Ansätzen in der Literatur (Hacking, Leonelli), allerdings mit Verwandtschaften zu anderen (Suppes), schlage ich vor, Daten als nicht-materielle und rein mathematische Objekte oder Aussagen oder Äußerungen zu verstehen. Daten können keine repräsentative Rolle für das nicht direkt Beobachtbare spielen, weil es dann ein Problem mit einer Beziehung ohne Bezugsobjekt (engl. *relation without relata*) gäbe, weil nichts durch etwas repräsentiert werden kann, wenn über das vermeintlich repräsentierte nicht mehr gewusst werden kann, als die unterstellte Repräsentation offenbart.

Im zweiten Schritt folge ich Grenanders erkenntnistheoretischer Idee Muster als genuin *konstruktive* (in einem ganz bestimmten Sinne) mathematische Objekte zu verstehen. Dies steht betont im Gegensatz zu weit verbreiteten Ansätzen der Explikation von *Informationen* in Daten aus der Informatik (Shannon, Kolmogorov).

Drittens stelle ich dar, dass nicht nur alle Fälle von Daten und Mustern in wissenschaftlichen Schlussvorgängen im Prinzip mathematisch expliziert werden können, sondern auch die Schlussvorgänge selbst, die zur Phänomenauswahl und zur Theorienbildung führen. Eine solche Ansicht ist keinesfalls neu und hat bereits klassische, einflussreiche Vertreter, wie etwa Russell, und kann sogar auf Basis neuerlicher Ergebnisse aus der Forschung an künstlicher Intelligenz empirisch untermauert werden, weil alle auch noch so komplexen Computerprogramme durch Dekompilierung immer auf rein mathematische Funktionen zurückgeführt werden können.

Auf Basis dieser Erkenntnistheorie prinzipieller mathematischer Explizierbarkeit erscheint auch das wissenschaftliche Schließen nicht als ein von einem menschlichen Bewusstsein notwendig abhängiger Prozess. Daraus folgt, dass es keinen Grund gibt anzunehmen, dass ein spezifisch menschlicher Geist nötig wäre, um große kognitive Leistungen, wie etwa unsere beeindruckendsten wissenschaftliche Argumentation zu vollbringen, wie es einige Autoren (Searle) nahelegen.

Die empirische Welt präsentiert sich uns durch materielle kausale Interaktion mittels Messgeräten und unseren Sinnen. Welche Muster in den Beobachtungsdaten uns als Phänomene erscheinen und was uninteressante Eigenschaften der Daten sind, hängt einzig von unseren theoretischen Annahmen über die Welt, sowie unseren sensorischen und kognitiven Voraussetzungen ab. Diese Einordnung wird oft scheinbar aufgrund der extremen Komplexität menschlicher Kognitionsprozesse missverstanden und ist in keinster Weise als durch die Konstitution der realen Welt selbst begründet nachweisbar oder hätte in irgendeiner substanziellen Weise mit menschlichem Bewusstsein zu tun.

Preface

After studying mathematics with a focus on stochastic processes and mathematics of finance I spent some years as a practising consultant in finance. Applications of mathematical methods of pattern detection on actual market data was a crucial part of my daily work environment. Expensive computer hardware and modern computer applications, such as neural networks, deep learning and GPU programming are intensely discussed opportunities by practioners of quantitative hedge funds.

Coincidentally, I was confronted with the philosophical debate on the explication of scientific phenomena and patterns in data. I was deeply puzzled about how prominent author's of the field came to conclusions that seemed very counterintuitive to me and could easily be challenged with the help of some good empirical examples. I surveyed more philosophical work on the topic, in particular Machamer's 2011 *Synthese* special issue, and could not find any account that was sufficiently similar to my intuitions at that time. These intuitions are namely that, firstly, no realism is necessary to explain phenomena selection, secondly, scientific reasoning does not depend on human consciousness, and, thirdly, valid scientific reasoning and in particular pattern detection procedures can in principle always be explicated in mathematical terms.

For these reasons I changed my originally planned topic for the dissertation thesis from an epistemological analysis of scientific methods in the field of quantitative finance to the topic of an explication of scientific phenomena and patterns in data. This topic is, of course, highly general and it is therefore a particularly risky endeavour to tackle it as a rather inexperienced philosopher. If one aims to provide a *big picture* in a very general manner, then he faces an increased risk to be challenged from many different angles. However, as a learning experience I deem it more worthy to work on risky and general questions of epistemology, than to look for an immature niche of an established branch that a young researcher could try to fill.

The topic of phenomena and their relation to patterns in data is obviously a very relevant one in the field of epistemology, if the related notions do not fully deflate into aspects from other accounts from the general philosophy of science, which I believe they do not. That is why I am overall surprised about the comparatively little attention that the notion of phenomena received by philosophers, given the extensive use of the term ‘phenomenon’ in science and the philosophy of science.

Reading of the Document

These are some editorial comments regarding stylistic considerations of this text.

To increase readability, chapters are written in a manner that allows them to be read in isolation. This means that I prefer to tolerate slight redundancies over too many cross-references between very different text passages. Chapters and sections in them are numbered for referential purposes rather than a strict order of arguments, and both come with an abstract except for the introduction, the conclusion and some rare cases of very short sections. I use various forms of illustrations more often than common in the relevant branches of philosophy, because, in my view, illustrations are epistemically very advantageous for the discussion. Many footnotes were added to, firstly, provide broader philosophical backgrounds, secondly, mention further literature, and, thirdly, provide further scientific information about mentioned examples from empirical sciences. In the appendix a more thorough view is provided on some mathematical and scientific examples that play a non-negligible role for some philosophical arguments in the text.

For the PDF version of the text all links to chapters, sections, pages, figures and also referred to literature in the reference list are made clickable. I added a subject index and a person index, which should significantly increase the readability (and possible critical scrutiny) of the text, in particular if read as PDF version on a computer screen.

Citations of other authors are given in double quotation marks (“...”), whereas references to terms, strings or signs are given in single quotation marks (‘...’). *Italic* words are meant to be specifically emphasised and were chosen with extra care. I use British English and widely used standard LaTeX libraries with only slight modifications, and in particular the 2002 version of LMU’s template for dissertations from R. Dahlke and S. Stintzing.

Acknowledgements

I would like to thank the following colleagues for comments on the topics of this thesis on the occasions of personal conversations or conference discussions:

Erik Curiel, Radin Dardashti, Stephan Hartmann, Philipp Haueis, Mario Hubert, Johannes Leonhard Keller, Maximilian Kistler, Hans-Christoph Kotzsch, Hannes Leitgeb, Sebastian Lutz, Thomas Meier, Arthur Merin, Nikil Mukerji, Simon Scheller, Christina Schneider, Markus Schrenk, Irina Spiegel, Jan Sprenger, Klaus Staudacher, Inken Titz, Alexander Reutlinger, Thomas Roth, William Wimsatt, James Woodward, three anonymous reviewers at *Philosophy of Science*, one anonymous reviewer at *The British Journal for the Philosophy of Science* and my family.

My work on this thesis was generously supported by a dissertation scholarship from the Studienstiftung des deutschen Volkes for which I am very grateful. The Munich Center for Mathematical Philosophy provided an efficient infrastructure, travel opportunities and a network of highly skilled scholars in branches that are very relevant to the topic of this thesis. Furthermore, the Siemens Stiftung and LMU's university library made it possible to acquire every single book that I deemed interesting for my work but that are hard to come by.

Obituary

Ulf Grenander (23 July 1923 – 12 May 2016) was a professor of applied mathematics at Brown University and Stockholm University. The Swedish born statistician was a member of the Royal Academy of Science of Sweden and the US National Academy of Science.

He devoted decades of research to *general pattern theory* from a monograph that foreshadows his interest in 1963 to a comprehensive survey with Michael Miller (2007). In this thesis, I conclude that his works provide the best approach to constructively explicate the notion of a pattern in science.

Ulf Grenander passed away recently and I was just a little too late to offer him my philosophical comment on his mathematical work. I hope my analysis helps to promote his outstanding contribution in explicative mathematical philosophy to not only researchers in his area of applied mathematics, but to philosophers, too. Geman (2016) provides a thorough obituary in the name of the Royal Statistical Society.

Chapter 1

Introduction

A classical position to explain the aim of scientific theories is the view of logical empiricism[1] as held by Nagel (1961) and Carnap (1966). It implies, among many other things, as an essential aspect that theories are formulated by scientists to explain and predict facts about *observables*. Observables are properties or relations that are "directly" perceivable by human senses or with the help of only "relatively simple" (Carnap, ch. 23) auxiliary equipment. The notion of observables and observability is intricate on its own, but I take it as roughly understandable at this point for introductory purposes.[2] Figure 1.1 shows a schematic illustration of this essential view in its simplest form.

Bogen and Woodward (1988), and Woodward (1989; 2000; 2011)—in the following abbreviated as '(Bogen and) Woodward'—criticise this logical empiricists' account in this regard. They propose an alternative approach with realistic implications to explain the motivation for the formulation of theories in science. They highlight that in science observations are processed in the form of data, which are records of observations. They point out, using examples, that this data is often edited and modified. (cf. 1988, p. 308-310 and 315-316) And this modification and editing is a crucial aspect of the scientific inference from observation to theory. If a set of measurement data shows only noise, then it gets discarded, or it may occur that only a few out of thousands of pictures are considered to show *interesting* results.[3]

[1]I make no distinction between the historical labels 'logical empiricism' and 'logical positivism' whatsoever, even if some historical proponents from either the *Vienna Circle* or the *Berlin Society of Empirical Philosophy* (later renamed to *Berlin Society of Scientific Philosophy*) may stress differences, which are not relevant for my purposes. For a more detailed historical survey see Creath (2011).

[2]Observables play a central role for various non-realistic approaches about scientific theories and theoretical objects. Concerning its importance for van Fraassen's (1980; 1985; 2001) constructive empiricism, Churchland (1985) and Hacking (1985) criticise the notion of observables and Teller (2001) wants to extend it.

[3]Suppes (1962) offers a descriptive account for the inference from direct observation or raw data to theory. However, his influential account of *models of data* does not provide an epistemo-

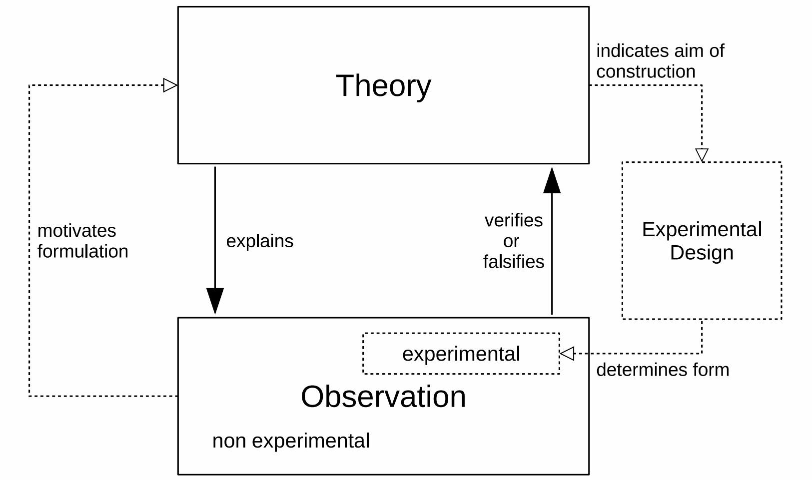


Figure 1.1: Schematic view of the role of scientific theories and observation in accordance with classical logical empiricism. Boxes indicate common notions from the description of scientific practice. Full arrows and lines indicate relations between the concepts. Dotted lines and lights arrows indicate concepts and relations with minor importance to the explanation of scientific practice. Note that, prima facie, observations from very specifically designed experiments (*e.g.* tracks of weak neutral currents in bubble chambers) provide more relevant examples for possible theory-ladenness than more incidental observations (*e.g.* photograph of a black raven).

A classical defence stresses the so-called *theory-ladenness* of observation. At this introductory point of the thesis, I mention this defence as a prior notice to which (Bogen and) Wooward's criticism is a reaction to. As advocated by Hanson (1958), and also subject of discussion in classical accounts from Feyerabend (1962) and Kuhn (1962), this defence states that to a certain scientist, observations can be (and for epistemic reasons are) biased by a theory under test. Bogen and Woodward argue that theory-ladenness may have an effect on some observations, but it does not sufficiently explain the aim of this bias, that is, why and how the observation is biased. (cf. 1988, p. 346)

If the only reason for this bias is that data is modified to verify a theory that was already proposed based on empirical knowledge, then the chain of scientific reasoning, which is, first, theory formation, then experimenting, and then theory verification, corroboration or falsification, seems empty prima facie. Bogen and Woodward's example to illustrate this aspect are the detection of weak neutral currents by bubble chamber photographs. In experiments from the 1970s, 290'000 of these photographs were taken, but only approximately 100 of them were considered to show *interesting* information with regard to testing the theory that introduces weak neutral currents. In their view, this practice is broadly inconsistent with classical logical empiricism and can, as well, not be explained by theory-ladenness,

logical explanation or any metaphysical implications about this inference. I investigate Suppes' and further, more contemporary accounts of data in 2.1

because if only the theory tells us, what data we have to *accept*, then it seems impossible to falsify it. That is why theory-ladenness cannot be the answer to describe this scientific practice.

The introduction of *phenomena* is a strategy to overcome these alleged problems with the logical empiricists' explanation of the aim of scientific theories and the introduction of the concept of theory-ladenness. Phenomena are real properties or relations of the part of the world under investigation by science. They play a crucial role in the scientific progress of theory construction and testing, but they are neither a *part* of theories (*i.e.* syntactical) nor a part of observation (*i.e.* observer dependent). Phenomena are best described by the role they play in science:[1] scientific theories aim to explain claims about phenomena, and these can be detected by patterns in data. Or, to put it another way, without a phenomenon data from some empirical part of the world under investigation has no information that is useful for scientific theory construction.

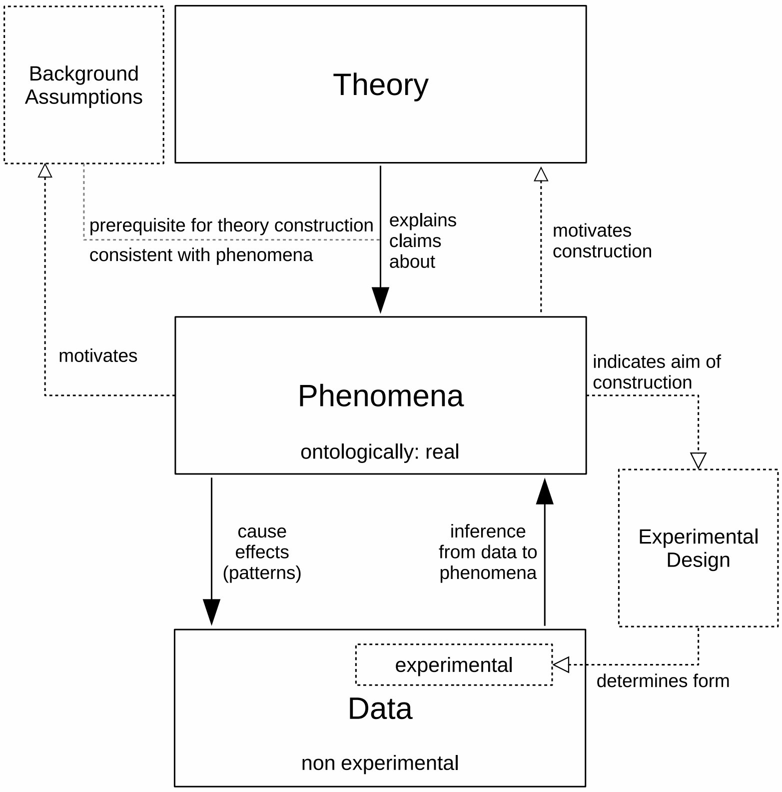


Figure 1.2: Schematic view of (Bogen and) Woodward's perspective on the role of scientific theories, phenomena and data. Box and line styles are used with the same meaning than in figure 1.1. I adapt this schematic view to my explication of phenomena in the conclusion of this thesis (figure 7.1 on page 191).

Despite the intuitively conclusive description of phenomena by its role in scientific processes of inference, an exact philosophical explanation is intricate. Wood-

[1] I do not introduce this terminology ('play', 'role') as an indication for a strong *fictionalism* about science. I want to give only a brief outline of the terminology used in this text.

ward provides the vivid description that "[d]etecting a phenomenon is like looking for a needle in a haystack or (...) like fiddling with a malfunctioning radio until one's favourite station finally comes through clearly" (1989, p. 438, also cited by Massimi 2011). Phenomena are characterized as

(i) not idiosyncratic to the different ways that are used to detect them,

(ii) real in an ontological sense, and

(iii) identifiable via patterns in data.

Figure 1.2 provides a schematic view in direct comparison to the logical empiricists' version as shown by figure 1.1. In this thesis I want to focus on a philosophical explication of phenomena by challenging these three properties, which are central to a characterization of phenomena.

As it turns out, a mere identification of phenomena with patterns is problematic for more than one reason. Firstly, even if patterns in data are a common and well explicated notion in the field of *statistical* analysis, it is hard to explicate it more generally to a suitable extent according to the use of the term 'pattern' in the everyday language of science from the various different fields. Secondly, the same phenomenon can be detected by different patterns in different data. Apparently, a phenomenon must, depending on the specific understanding of the concept of a pattern, rather be identified with a group of patterns and not with a single pattern, as will be shown.

(Bogen and) Woodward are philosophically motivated to highlight the important distinction between data and phenomena. They aim to bury the logical positivists' idea of science following an, in their view, overly simplified systematic, which is that the empirical world is observable and theories are formulated to describe or explain the observation:

> Our general thesis, then, is that we need to distinguish what theories explain (phenomena or facts about phenomena) from what is uncontroversially observable (data). Traditional accounts of the role of observation in science blur this distinction and, because of this, neglect or misdescribe the details of the procedures by which scientists move from claims about data to claims about phenomena. In doing so, such accounts also overlook a number of considerations which bear on the reliability of scientific knowledge. (Bogen and Woodward 1988, p. 314)

The notion of *reliability* of data and its role for the explication of scientific phenomena will be discussed in the context of Massimi's (2011) comments on (Bogen and) Woodward in 6.3. What Bogen and Woodward do not aim for is to explain the relation between phenomena and patterns in more detail. Woodward gives a description how patterns come into play in this picture:

> [S]cientific investigation is typically carried out in a noisy environment; an environment in which the data we confront reflect the operation of many different causal factors, a number of which are due to the local, idiosyncratic features of the instruments we employ (including our senses) or the particular background situation in which we find ourselves. The problem of detecting a phenomenon is the problem of detecting a signal in this sea of noise, of *identifying* a relatively stable and invariant pattern of some simplicity and generality with recurrent features – a pattern which is not just an artifact of the particular detection techniques we employ or the local environment in which we operate. (Woodward 1989, p. 396–7, my emphasis)

This wording indicates that Woodward's notion of phenomena and of patterns are at least very closely tied to each other. I cannot find any substantially more revealing information on the relation between phenomena and patterns in their texts. An identification, including all metaphysical and epistemic consequences, may even be intended. Brown summarises his understanding of relation between phenomena, data and patters as follows:

> Phenomena are to be distinguished from data, the stuff of observation and experience. They are relatively abstract, but have a strongly visual character. They are constructed out of data, but not just any construction will do. Phenomena are natural kinds (or patterns) that we can picture. (Brown 1994, p. 141)

If phenomena are "constructed out of data" and "natural kinds (or patterns)", then I want to raise the question whether an identification of phenomena and patterns in data is intended. I claim that a mere identification is problematic due to several reasons and aim to explain the relation more thoroughly. My, in my view, most substantial objection is that phenomena and patterns in data must have a very different ontological status; according to my findings that I present in this thesis phenomena are empirical features of the world and patterns are mathematical objects (or descriptions).

Phenomena are not idiosyncratic. They can recur in different experimental setups in the form of different patterns in the data. This fact is strongly related to their alleged reality, since it would be hard to explain the occurrence of a phenomenon as patterns in different data without it being real. Mathematical methods, including statistical data analysis, are used to detect patterns and can to some extent be used to describe them. As a simple, illustrative example I introduce the biological phenomenon of albinism in animals and use it to make the basic ideas of this thesis more tangible.

Not only does a working conceptual identification of phenomena with patterns demand further explanation, the concept of a pattern in data itself does to. Data can for instance be given as a list of numbers, as photographs or as text reports.

Due to this variety of formats of data, an explication of the notion of patterns must be very flexible to be applicable to all kinds of data relevant for all scientific fields. An explication of patterns in data and the relation between patterns and phenomena is helpful to achieve a thorough understanding of phenomena in science. This understanding helps to philosophically describe science.

Do I aim to provide a *normative* or a *descriptive* account of phenomena and patterns? My approach is descriptive in the sense that I make a lot of use of examples from various scientific fields and I aim to provide adaptable descriptions for these examples. My approach is normative in the sense that I provide explications for *good* scientific reasoning in very general. In my view, descriptively inadequate normative approaches are are of not much philosophical interest because they seem to explicate something different than what they are named after. Descriptively adequate but normatively unacceptable approaches (*e.g.* logically inconsistent) hint to either problems with proper reasoning among a significant proportion of scientists or an philosopher's inadequate set of background assumptions for the normative framework (*e.g.* *where* consistency matters). However, I believe that the problem of philosophically explicating phenomena is one that demands a rather wide-angled view on science with a significant amount of descriptive considerations, than a to a lesser amount empirically inferred normative drill-down based on axioms of reasoning.

Defending his methodological approach towards phenomena implicitly, Bogen describes why, in his view, explications with "rigor [and] precisions" are not appropriate for "universally applicable accounts of scientific reasoning":

> It's plausible that philosophers who value the kind of rigor, precision, and generality to which logical positivists, logical empiricists, and other exact philosophers aspired could do better by examining and developing techniques and results from logic, probability theory, statistics, machine learning, and computer modelling, etc. than by trying to construct highly general theories of observation and its role in science. Logic and the rest seem unable to deliver satisfactory, universally applicable accounts of scientific reasoning. But they have illuminating local applications, some of which can be of use to scientists as well as philosophers. (Bogen 2017)

The important question regarding "highly general theories of observation [or phenomena] and [their] role in science" in the context of logical and mathematical explications is whether these explications can either exemplify how generally powerful this philosophical methodology is by means of concrete examples, or whether they are general enough to capture everything in science that falls under these notions. As I elaborate on in chapters 2-4, classics such as Russell, Frege and Hempel had, according to my interpretation, such a high degree of generality for formal explications in mind, and I follow their approach by discussing a formal framework for patterns with maximum generality. I conclude in this thesis that, given relevant arguments from epistemology and philosophy of mind, an anti-realistic notion of phenomena and a mathematical notion of patterns are the most adequate explications.

In my view, my conclusive account of phenomena is able to provide sufficient answers to all the problems that I read about in the mentioned literature and that were put forward after my talks and in personal discussions.

This thesis is organised as follows. Throughout the text I comment on problems from recent and classical literature on phenomena and related epistemological and metaphysical topics. In chapter 2 I discuss available ideas regarding the explication of data and I suggest an explication of it that implies that sets of data are mathematical objects, which can and often are represented in processes of scientific inferences by human agents. In chapter 3 I defend the central role of mathematics in my epistemology by showing its expressive strength, its extensive application via computers and the empirical epistemic strength of these computer applications. Furthermore, in this chapter I provide an explicative notion of mathematics for any reader without an indepth knowledge of philosophy of mathematics. In chapter 4 I explicate patterns and discuss the epistemological as well as the metaphysical implications of my view. The notions of *concrete* and of *general* patterns prepare the ground for the explanation of the relation between patterns and phenomena, which is provided in chapter 5. Here I also comment on some problems with the notion of phenomena as they are discussed in the literature. My results provide a descriptively adequate notion of phenomena, whereas earlier chapters prepared the ground for the view that this descriptive notion substantially differs from a normative notion of *good* phenomenon selection. More precisely, scientist's actual selection of phenomena via patterns in data are not epistemically grounded to a sufficient degree. In the following chapter 6 I comment on some articles that specifically criticise (Bogen and) Woodward's work. Chapter 7 contains a conclusive discussion of the, in my view, most central questions and concerns regarding the relation between patterns and phenomena with reference to the relevant chapters and sections throughout this thesis. I added some comments from the perspective of mathematical statistics and its applications to the appendix, because these insights aid the philosophical discussion as an exemplary case of mathematically explicated detections of patterns and the discussion of related phenomena. In the remainder of this introductory chapter I discuss some historical aspects and central terminology for the topic of this thesis.

1.1 Historical Remarks on 'Phenomenon' in Philosophy

Section Abstract

Bogen and Woodward (1988) raised awareness for the necessity of a philosophical explication of the notion of phenomena. They focus on the use of the notion from science direcly rather than from the history of philosophy. For Suppes and Carnap the notion is unproblematic. Van Fraassen defends an anti-realistic notion of phenomena. The notion of phenomena in science and philosophy of science has some commonalities with the notion from philosophical phenomenology, but has, all things considered, a substantially different meaning.

The term 'phenomenon' is used in various different contexts in philosophy and is also widely used among scientists. Its meaning varies substantially. Some historical remarks maybe useful for the reader.[1]

The semantic connection between Bogen and Woodward's (1988) introduction of the term on the one side and the historical use in most parts of philosophy on the other side is very loose, but cannot be completely denied. More importantly, Bogen and Woodward adopt 'phenomenon' from its use in the empirical sciences (in particular physics) rather than from its occurrences in the history of philosophy. For the discourse in philosophy of science, van Fraassen (1980) uses 'phenomenon' with a distinct anti-realistic meaning in chapter 3, which is titled 'To Save the Phenomena':

> When Newton wrote his *Mathematical Principles of Natural Philosophy* and *System of the World*, he carefully distinguished the phenomena to be saved from the reality to be postulated. He distinguished the "absolute magnitudes" which appear in his axioms from their "sensible measures" which are determined experimentally. He discussed carefully the ways in which, and extent to which, "the true motions of particular bodies may be determined from the apparent", via the assertion that "the apparent motions ... are the differences of true motions". (p. 44, quotes in quote from Cajori 1960, p. 12)

Later in the text he writes:

> Electrified and magnetic bodies appear to set each other in motion although they are some distance apart. Early in the nineteenth century mathematical theories were developed treating these phenomena in analogy with gravitation, as cases of action at a distance, by means of forces which such bodies exert on each other. (p. 48)

[1] This introduction seems very noteworthy to *me*, since *I* was confused about the use of the term at my first contact with the topic.

For van Fraassen phenomena are an unproblematic notion and are whatever appearance an observing scientists in the context of a theory might have interest in without any further stipulation about reality, which is what we should expect from a *constructive empiricist*.

Suppes (1969) is philosophically more agnostic and uses 'phenomenon' in the same pragmatic and ambiguous way as scientists do[1] (*i.e.* without ontological or epistemological commitments). Carnap (1966) uses 'phenomenon' similar and does not mention it in his four pages long subject index. All in all, it seems fair to conclude that (Bogen and) Woodward were the first to influentially raise awareness for the philosophical problems with the notion of phenomena to a larger audience with a focus on philosophy of science.—Of course, I may not be aware of other authors' significant contributions to the topic.

Adopted from the ancient greek word 'φαινόμενον' ('*phainómenon*'), which means *that which appears* (cf. Preus 2007, p. 298), 'phenomenon' became a widely used term in philosophy describing, as Smith (2008) puts it, "appearances of things, or things as they appear in our experience, or the ways we experience things, thus the meanings things have in our experience". This notion is anti-realistic towards phenomena and I tend to claim that 'things' should more appropriately be replaced with '(parts of) the empirical world' according to my elaborations in 1.3.

Lewis (1929, 1929, ch. VI) pools positions that accentuate the unknown difference between the actual world and its appearance to us under 'phenomenalism'. Accordingly, Husserl[2] introduced *phenomenology* as a philosophical discipline focussing not on how things may be, but on how they are perceived, or what the structure of various types of experiences is. In classical Husserlian phenomenology our experiences are directed by *intentions* and therefore, similar to Kant (1787), the reality is in most cases[3] considered to be epistemically inaccessible. (cf. Smith 2008)[4]

Due to the different use of 'phenomenon' in the philosophy of science, I avoid the use of the term in the meaning as described by Smith. The substantial difference is that for science and philosophers of science phenomena are not only features of the empirical world how they appear to us. They need to fulfil a further important criterion. As I conclude in this thesis (chapter 5), scientific phenomena are a

[1] He uses the term in various of the papers from the 1950s and 1960s that are collected in the mentioned volume.

[2] Due to Husserl's style of writing and publishing it is improper to refer to some single text explaining *phenomenology*. A more extensive survey would be needed. Nevertheless, for our purpose I refer to the study by Moran (2000) and the encyclopaedia entry by Smith (2008). The original source, which may be considered to be the most important one, is Husserl (1913).

[3] An exception is, if ideal or abstract objects, for instance mathematical objects, are considered to be real.

[4] For extensive historical surveys on phenomenology see Spiegelberg (1994) or Moran (2000).

predominantly interesting target for scientific explanation on the background of a scientific agent's background assumptions and his epistemic capabilities. Phenomena in Smith's sense do not need to fulfil this criterion. Furthermore, scientific phenomena may only be observable with the help of technically aided scientific measurements (*e.g.* microscope). It is at least a non-trivial problem for Smith's account of phenomena where to draw the demarcation line between *apparent* and *non-apparent* with regard to all sorts of technical observation aids, according to his notion of *appearance*.

On the other hand, all philosophical notions of phenomena focus on how we perceive or theoretically describe parts of the empirical world in accentuated opposition to the constitution of a real empirical world itself. For this reason we can say that the philosophical notions of phenomena have a common semantic core. But due to the very different conceptual demands of philosophical phenomenology on the one hand, and a philosophical view on science on the other hand, the notions diverge. The philosophers of science adopted the scientist's notion of phenomena. I do the same throughout this thesis.

1.2 The Notion of Science in this Thesis

Section Abstract

Science is an activity, but it is hard to explicate what makes and act scientific. Scientific reasoning aims for explanations and predictions of parts of the empirical world. Scientific agency is not restricted to human beings. My notion of science is as broad as possible and includes also fields such as psychology, philology and historical studies. The defining criterion for my notion of science is the use of empirical data as evidence to corroborate or falsify hypotheses. Everyday reasoning of laymen that fulfils this criterion falls under my notion of science, as well. My approach to explicate phenomena is agnostic regarding a *syntactic*, *semantic* or *pragmatic* view towards scientific theories.

The meaning of the term 'science' is ambiguous, often used by speakers with various backgrounds (*e.g.* scientists; politicians; tabloid media), and subject of a substantial amount of works in philosophy of science. More specifically, it is a very hard task to explicate descriptively or normatively what qualifies an agent's *activity* as scientific or, more specifically, what scientific reasoning exactly is. Approaches to explain what science is can include arguments concerning the motivation for the activity (*e.g.* explain or predict empirical phenomena), its results (*e.g.* confirmation of a theory) and even arguments that aim for social or institutional circumstances (*e.g.* is discussed by academic professionals).

What can be said for certain at this point is that science is *done* by *agents*, because only the psychological or pragmatical motivations to gain knowledge and the epistemic restrictions that make knowledge hard to access make science necessary. In other words, god does not need science, because he has unbounded access to knowledge, and a stone or a horse does not do science because they are not motivated to gain knowledge.

However, in my view, we can, as a first approach, accept that the notion of science is very ambiguous, but it somehow includes the search for explanations and predictions of parts of the empirical world. Furthermore, there is no reason to restrict scientific activity to human agents; artificial intelligences or aliens can *do* science, as well. Various technical auxiliaries (*e.g.* pattern recognition with neural networks) already indicate that science is not a solely human activity. The argument for this is that these pattern recognition applications can come to conclusions that are impossible to reach with a human mind (*e.g.* due to complexity).—I elaborate on this argument in chapter 3.

Not independent of, but also not fully determined by the question what qualifies an activity as scientific, is the question what parts of the empirical world can be subject of scientific explanation or prediction. Science is often understood to explain *nature*. Nature certainly includes matter (physics; chemistry) or living organisms (biology). It is a philosophical question whether the human mind (psychology) can fully be understood as being a part of the nature; *free will*, *indeterminism* and *qualia* are keywords of philosophical positions against the interpretation of the human mind as a part of the nature in a very narrow sense. The philosophy of mind may be seen as a philosophical field in which the mind's relation to nature is investigated. (Empirical) psychology may usually be accepted as a science by traditional and for normative reasons, but what about social studies or academic fields that aim for the explanation of arts and human culture?

The aim of this thesis is to explain the relation between patterns and phenomena. Since Bogen and Woodward introduce phenomena simply and broadly as "features of the empirical world", there is no restriction to the empirical subject of study, in which these features appear, at all. On the other side of the relation that I aim to explain are patterns. I claim that patterns are mathematical objects (chapter 4), which can be detected in data and which are also mathematical objects (chapter 2). I imply that we understand the term 'pattern' in such a broad way, in which it is also used in everyday language and science. Data can be a series of numbers, as well as a facsimile of an ancient Roman law text. That is why, since my notion of phenomena, data and patterns in science is as broad as possible, the notion of *science* needs to be as broad as possible, as well.

I include into my notion of science all activities that involve data, phenomena and pattern recognition in this described sense. Examples of these activities are science of history, science of literature, physics, biology, history of art, psychology, linguistics and others. The defining criterion for my notion of science is the use of data as evidence to corroborate or falsify hypotheses that imply descriptions of *empirical* phenomena (chapter 5). Mathematics does therefore not count as a science according to my account, since there are no empirical phenomena in mathematics. Doing art does not count as doing science either, even if it is used to illustrate or highlight features of the empirical world that may have to do with phenomena; in arts data is not gathered or used to corroborate or falsify hypotheses. However, I emphasize the *activity* character of science, because designing or carrying out an experiment is undoubtedly a crucial part of science, but not necessarily fully based on conscious reasoning.

Can we say more to characterise science, in particular by explicating a *scientific method*? I believe it is very hard to justify an adequate *descriptive* account for *the* scientific method. Many attempts of various sorts have been made to provide *normative* accounts for scientific activity. Aristotle, Ockham, Descartes, Leibniz, Newton, Hume, Kant, the logical empiricists and Popper are only a short list of authors who provide normative accounts of *good* scientific reasoning throughout their philosophical works.[1]

But I want to point out that my notion of science is particularly broad. Many of the normative works in the literature focus on pragmatically convenient examples from, roughly, physics, astronomy or biology (even if our modern notions for these branches do not completely fit the more classical notions, such as *natural philosophy*). This plain environment helps to focus on specific normative aspects, but it implicitly oversimplifies relevant pragmatic issues for descriptive accounts. For instance, rules for *parsimony* of an explanation in psychology or social sciences are much more intricate than they are for common examples of heliocentrism in astronomy or evolutionary theory in biology. Another example is the impossibility to falsify a palaeontologic hypothesis about an extinct animal, if the only source of data is one single incomplete fossil without the possiblity to get another specimen.

Furthermore, if we do not include only the activities that aim for the evaluation of a hypothesis into our account of a scientific method but also all the psychological causes and the creativity—whatever that may be in detail—that lead to new ideas for hypotheses than it is even more apparent to reject the idea of *a* good scientific method. In my view, Feyerabend (1975) provides a good historical survey about how much methodological diversity, which include social and *irrational* factors, was involved in some of our most influential scientific revolutions (*e.g.* heliocentrism;

[1]Sober's (2015, ch. 1 and 2) historical account provides further details and references with a focus on parsimony criteria for theory selection.

general relativity). As Feyerabend, I am more interested in a *descriptive* notion of science how it is actually *done* by the agents than in normative accounts how *good* science *should normatively be*.[1] It is a further question, which is not in scope of this thesis, whether it is even possible to provide such an normative account that covers all branches of science.

What then is the difference between the allegedly distinct field of science in very general on the one hand, and everyday reasoning from observations to conclusions on the other hand? I may speculate why I got sick last week or why my favourite party did get only that few votes in the last election. My answer to this question is that there is no principle difference and that everyday reasoning entails a lot of reasoning about ad hoc working hypotheses. Science cannot be characterised as avoiding reasoning fallacies, since in this case, we need to exclude from our list of what counts as scientific a substantial part of our historical scientific heritage, as well as some modern scientific positions.—Time will hopefully tell which positions exactly those are. The agents of science receive some specific training at the universities, but these lessons are build upon pre-scientific reasoning on the individual level, as well as on the level of the scientific community. That is why everyday reasoning of laymen falls under my notion of science as well and, what follows from this, every claim about scientific reasoning in this thesis is also a claim about everyday reasoning.

In this thesis I suggest explicit notions of data, of patterns and of phenomena. The reason for this is that in my view to explicate phenomena we need a clear understanding of patterns and of data in which these patterns can occur. However, I do not elaborate on my view on *theories* or *hypotheses*. Common candidates to explicate what a theory is are the *syntactic view*, the *semantic view*[2] and the *pragmatic view*[3]. These views on theories were developed to further clarify how exactly theories should be linguistically framed (*i.e.* in a predicate logic or mathematically) and what content exactly is part of the theory and what not (*e.g.* models; intended applications). The only claim that I need to accept for this thesis is that theories are aimed to make claims about phenomena. Since I defend the view that phenomena do not depend on particular theories (cf. 5.4) I can remain agnostic regarding the particular view towards explicating theories.

[1]Throughout the book, Feyerabend criticises the logical empiricists and Popper in a very polemic tone for their allegedly erroneous *descriptive* understanding of science. In my view, their accounts are rather *normative* and the authors were well aware that the history of actual science is something very different.

[2]For a brief introduction and discussion of the syntactic and the semantic view see Lutz (2014).

[3]For a discussion of the pragmatic view see Winter (2016).

1.3 Empirical World, Parts of the Empirical World, Empirical Objects, Physical Systems and *Ding an sich*

Section Abstract

Classical ideas from Descartes, Kant, logical empiricists, Wittgenstein and Everett indicate that our observations cannot fully grasp what I call *the empirical world*. The empirical world is what *causes* our observations and we cannot know whether the empirical world is a play of Descarte's demon or whether it is composed of objects or a structure or something else. I avoid 'systems' and 'objects', because these terms imply too detailed ontological commitments. I avoid 'outside world', because of Descarte's possible demon and the unwanted exclusion of psychological phenomena.

The empirical (world)[1] and a proper notion of it is obviously a very relevant aspect for any theoretical understanding of scientific inferences. Some very influential ideas from the history of philosophy indicate how intricate the notion of the empirical (world) is. Descartes (1641, Meditation I) claims that we cannot be certain that there is such a thing as an outside world, and according to Kant (1787) the *Ding an sich* may be something very different from what we humans are able to observe and imagine in space and time. Consequently, logical empiricists describe our epistemic access to the outside world, if there is such a thing, as tied to the form of *observation sentences* (or *protocol sentences*)[2] to emphasise our epistemic restrictions to, firstly, sensorially gathered *observations* and, secondly, the grammatical form of sentences.

A concept of *the* empirical world is threatend by further classical ideas, if we leave the door sufficiently wide open for complaints. Everett's (1957) many-worlds interpretation of quantum mechanics suggests that we need to get rid of the 'the'. Wittgenstein (1953) implies that our human preconditions for observation are not even *stable* in the sense that *inner perception* (Kant) or sentence grammar (logical empiricists) are more or less observer independent. For him, our language

[1] I put 'world' in brackets to emphasise some general problems with the description of the empirical (world or universe). On the one hand, 'the empirical' has the advantage to remain indifferent about any further metaphysical characterisations of what we observe and may be an acceptable designation for proponents of all ontological positions concerning the distinguishability or indistinguishability of the empirical from something else. On the other hand, the description as a world or universe already indicates the realists' idea of alternative worlds, which I consider to be a pragmatically useful description.

[2] The most adequate notion for observation sentences was heavily discussed among Moritz Schlick (1934), Otto Neurath (1932) and Rudolf Carnap (1932; 1958), Karl Popper (1935) and W.V.O. Quine (1981). The best brief encyclopaedic overview over the discussion that I am aware of is the German Wikipedia entry *Beobachtungssatz* (2017).

adaption is much less driven by widely shared *grammatical* rules, and much more influenced by less organised rules of the *language game*. This may imply that linguistic descriptions of observations are crucially observer-dependent, which does not support a view of epistemic access to *an* empirical world.

The conceptional role of a notion of *the empirical world* is in my terminology, roughly speaking, to play the counterpart to descriptions or abstractions in the two sided conception to explain science or specific parts or aspects of it. In other words, the empirical world is what *causes* observations, even if those observations are constraint by certain sensory or human epistemic conditions. Kant's *Ding an sich* is the classical concept that is closest to my use of 'the empirical world'. However, my notion is open to Descarte's demon scenario or Bostrom's simulation scenario, which imply that we may *observe* only a demon's play or that we live in another civilization's computer simulation. In these cases the empirical world would be the demon with his acts and possible intentions or, respectively, the computer that causes our experiences.

In this sense, I use the notion of the empirical (world) in my epistemological framework in a manner of a *primitive* notion as explained by Tarski (1994) (for deductive sciences) and de Beauregard Robinson (1959) (for knowledge): it is an impression that we cannot further explain and that is necessary as a "root" of knowledge and can be used to "define" further, more complex terms.[1]

However, in the history of philosophy the description if the empirical (world) is

[1] Tarski defines *primitive* (or *undefined*) terms for deductive sciences, which are mathematics and logic:

> When we set out to construct a given discipline, we distinguish, first of all, a certain small group of expressions of this discipline that seem to us to be immediately understandable; we call the expressions of this group PRIMITIVE TERMS or UNDEFINED TERMS, and we employ them without explaining their meanings. At the same time we adopt the principle: not to use any other expression of the discipline under consideration, unless its meaning has first been determined, with the help of primitive terms and of such expressions of the discipline whose meanings were explained previously. The sentences which determine the meanings of terms in this way are called DEFINITIONS, and the expressions themselves whose meanings are thereby determined are accordingly known as DEFINED TERMS. (1994, p. 110)

I extend this idea of primitive notions to my philosophical explications, which are—of course—not one or a couple of theories of a deductive science. In the case of the notion of *the empirical world*, I deem the notion as unproblematic, even if I do not want to commit myself to Robinson's epistemic description that all knowledge is based on primitive notions (for which a thorough justification is a substantial philosophical endeavour by itself):

> To a non-mathematician it often comes as a surprise that it is impossible to define explicitly all the terms which are used. This is not a superficial problem but lies at the root of all knowledge; it is necessary to begin somewhere, and to make progress one must clearly state those elements and relations which are undefined and those properties which are taken for granted. (1959, p. 8)

not unambiguous. This specifically holds, if we do not only take into account the use of the terminology in philosophy, but also in the sciences themselves. In this section I want to further illuminate the problems of the terminology and support my preferred use of 'the empirical world', which plays in important role throughout this thesis. I want to introduce and further justify my use of the terminology and discuss the problems with it regarding the topic of an explication of the relation between pattern and phenomena in science.

I explain why I prefer 'the empirical world' over 'outside world' or the like. To ontologically classify thoughts, Frege (1918–1919) identifies three "realms" that we can briefly paraphrase as the following: the abstract, the mental and the empirical. The interesting point that I want to highlight is his criterion to separate these three ontological classes. He introduces the empirical as being objects in the "outside world" (Ger. *Dinge der Außenwelt*), as opposed to the mental, which are the "objects of imagination" (Ger. *Dinge der Vorstellung*). The abstract is defined as objects for which none of these two criteria apply. What is interesting about the characterisation of *the empirical* as objects in the *outside world*? Firstly, the empirical world is composed of objects, which is not an approach without alternatives, as we will see. Secondly, the exclusion of mental state from the empirical world indicates that for Frege the empirical objects are given without any conceptualisation. This makes the assumption that the empirical world is composed of objects even more radical.

To interpret the empirical world as something that is composed by empirical objects seems natural, if we base this interpretation on the use and the structure of the languages. This includes everyday languages, as well as specialised languages, such as used in science. When I refer to the empirical world, I mostly want to refer to what scientific experiments or measurements are applied to and to the aim of scientific theories, which is the explanation of phenomena. In this context, the talk of empirical objects may be misleading. This is because in many cases the objects that play a role in the experiment or measurement are described as theoretical objects. Electrons or genes are the empirical objects investigated by certain experiments. But the reality of such objects is an assumption. Therefore the terminological use of 'empirical objects' should always be considered very carefully in these contexts.

A common approach in physics to refer to the empirical (world) is the reference to *systems*. In many cases the aim of physical theories is to explain phenomena of these empirical systems. The introduction of systems is a technique of idealisation in scientific processes, since all parts of the empirical world and processes outside of the system, which can be subsumed as *environment*, are ignored in the scientific endeavour of describing the behaviour of the system.

The rules for the boundaries of empirical systems to its environment are manifold and differ among the several different applications of the physical concept of systematisation. In the simple examples from thermodynamics a system could be a closed and thermally insulated container, such as a refrigerator, which is also an *isolated* system in terms of the theory, due to the thermal insulation.[1] Another common way to define systems in physics are scaling criteria: a certain part of the empirical world is understood as a system, but only on a certain scale. The water in an ocean can be considered as a meteorological system on the macro scale, as a fluid mechanic system on an intermediate scale and as a molecular system on the atomic scale. The important point for us is that the describing theories in different systems describing the same parts of the empirical world cannot always be consistently transferred into one big theory for all systems; physical theories are not necessarily *scaling invariant*.

A classical debate about the consistency or inconsistency of explanations of phenomena on a micro and on the macro scale is the explanation of irreversible macroscopic laws that can or cannot be inferred from time-reversible laws of microscopic physics. Irreversible macroscopic laws show themselves in events, such as the breaking of a glass or the death of a human being, which can occur only in one direction of time, but not reversibly; there are no natural incidents of broken bits of glass recompose into the glass or the resurrection of a human being. A widely discussed irreversible law is the second law of thermodynamics[2], since according to it, the entropy of an isolated thermodynamic system increases to a maximum, but never decreases. As Goldstein (2001) and Lebowitz (1993; 1994; 1995) discuss, Ludwig Boltzmann showed, how the second law of thermodynamic originate in time-reversible microscopic laws. His Argumentation was famously, but unsuccessfully, criticised by contemporaries, including Ernst Zermelo and Josef Loschmidt.

How are we affected by the physical approach of describing the empirical world partitioned into systems? When writing about to the aim of scientific theory construction or pattern detection in scientific data, it is often inappropriate to refer to the behaviour of empirical objects, even if the talk of objects and its properties is a very established approach in the realists' ontology, as indicated by Frege. A more appropriate use of vocabulary is to talk about objects in a specific system, which is neither practical, nor does it provide any gain for our discussion.

But since I want to stay agnostic regarding the ontological reality of what is

[1]Strictly speaking, this is not an isolated system. The thermal insulation is not perfect and the heat exchanger transports heat out of the system.

[2]One simple expression of the second law of thermodynamics is:

Heat cannot spontaneously flow from a colder location to a hotter location.

observed, I avoid both ‘system’ and ‘object’ (and ‘system of objects’ or the like). Furthermore, I avoid ‘outside world’, because, firstly, we do not know whether there is an outside (*e.g.* Descarte’s demon), secondly, even if there is an outside, it may not be what causes our observation (*e.g.* Matrix movie), and thirdly, our mental states are subject to the scientific field of psychology, which is not properly described as *outside*, if we accept that we all are driven by psychological phenomena, which is obviously the case. That is why I choose ‘the empirical world’ and a *part* of the empirical world can be any kind of part of it—in many actual cases of picking or branching questions in science the partition criteria are hard to explicate in detail.

However, what can be said for certain is that the empirical world or a specific part of it *causes* in one way or another all observations of human agents or our superior artificial intelligences. *How* it does that exactly, may be a question that cannot be answered scientifically due to the agents’ inherently restricted epistemic access to observation.

Chapter 2

Data under an *Ante Rem* View on Mathematics

Chapter Abstract

I narrow down how the term 'data' is used in this survey and provide a descriptive explication of data in science. A brief survey about the available philosophical literature on data is given, which includes Suppes' concept of *models of data*, Hacking's idea of data as material externalisations of human acts and Leonelli's functional characterisation of data as portable material objects. I conclude, opposing to some of the views from the literature, that data is non-material and mathematical. Data, in the most general sense, is the class of all the mathematical *representations* of the actual data sets that are used for actual scientific inferences.

The aim of this thesis is to explain the relation between phenomena and patterns in data. Therefore, we need to provide a sufficiently detailed account of these two concepts. I start this elaboration with a discussion on data. It will turn out that sets of data are *abstract* objects. I understand abstractness as fulfilling the simple criterion of not having a specific position in time and space. Data can be *represented* by physical objects or mathematically and in particular, if computer processing is applied, by finite mathematical objects. The following explication of data in this chapter is the basis for the philosophical analysis of patterns in data that will follow in chapter 5 and which is necessary to describe the relation between patterns and phenomena.

The notion of data plays a central role for the methodology of the empirical sciences. The concept of data is frequently used in the specialists' literature of sciences, such as physics, biology and the social sciences. Examples include numerical measurement results for the melting point of lead, the result of a gene analysis, an X-ray photograph, a photograph of a black raven, Goebbel's Sportpalast speech[1]

[1] In this often cited speech he praises to the audience the "total war", which should be an

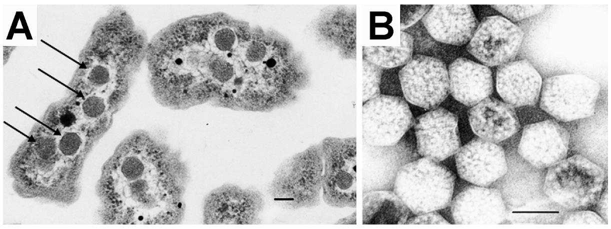
and last Sunday's survey about the German citizens' vote preferences for the Bundestag. I mention these examples to stress how broad this idea of data is and that I understand social and historical studies as scientific endeavours that do not differ from physics by any fundamental epistemological reasons regarding the treatment of data. That is why I do not stipulate any fundamental distinction between data from the different scientific branches. I aim to provide a justification for this epistemologically equal treatment of physics and non-physical sciences throughout the course of this thesis, since physics, according to my survey of the literature, seems to be the predominant example for investigations in the philosophy of science in general and regarding the treatment of data in particular.

The common translation of the Latin 'datum' (with its plural 'data') into 'what is given' does—obviously—not directly apply to the use of the notion in the context of scientific inferences, because actual scientific data is often not gathered from any form of *direct observation*, but by the use of highly specialised routines or laboratory equipment and translated into different formats. However, it is also not completely unrelated, since data plays the crucial role in scientific inferences to *record* empirical observations in a *language* that makes inferences between data and theories (by humans or computers) possible. For example: a modern theory about the detailed trajectories of planetary motion can best be confirmed by numerical measurement results of actual planetary motions and hardly by any non-quantitative observation of the empirical part of the world under investigation (*e.g.* a child's story about it seeing Venus in the night sky).

If googled today, the term 'data' refers in the majority of search results to binary stored information[1] in computer hardware. The starting point of our discussion of data is another: I use 'data' in the way (Bogen and) Woodward use it and how it is referred to in Illustration 1.2. Data is something that plays a role in scientific inferences and whereas the notion of observation seems to imply a certain role of an agent, the observer, data does not so much. Data is rather a somehow finished record from some measurement routine, that can be stored and transferred in relative isolation to any agent. In this sense, data is the intersubjective, observer independent counterpart to observation. One important result of my philosophical investigations will be the claim that a set of data is a mathematical object and this claim holds not only for (mathematical) physics, but for the empirical sciences in general. That is why the idea that data is some information stored in a computer system is a more adequate description than it might seem from the starting point of (Bogen and) Woodward.

absolute dedication to the war with the aim to end it quickly, devastating and victorious.

[1]My notion of *information* in this chapter and throughout the thesis is the one from information theory, which is a subfield of computer science. Information are linguistic signals (*e.g.* ASCII strings; English sentences).



time (seconds)	**strain** $*$ 1.e21
2.500000000000000000e-01	2.454791884395226415e-02
2.500610351562500000e-01	1.529268268197186628e-02
2.501220703125000000e-01	6.372337209458739903e-03
2.501831054687500000e-01	-2.075009545624776318e-03
⋮	⋮

Figure 2.1: Left: electronic micrograph images for a "structural analysis" of a bacillus from Tsai *et al.* (2007)[1]. Right: part of the *LIGO* data in which Abbott *et al.* (2016) found strong evidence for the existence of gravitational waves.

In the context of a specific scientific discussion, 'data' usually has a more or less unambiguous meaning or reference class, because scientists share a common understanding about the meaning of 'data' inside of their very specific and often strongly separated scientific communities, which implies a shared use of technical language. Borders between these communities lie not only between academic branches (*e.g.* physics; biology; economics), but also between different groups of experts on the investigation of specific scientific phenomena inside a scientific discipline (*e.g.* micro- and macroeconomics; scales above or below the quantum effect threshold in physics). There are even phenomena that are investigated by scientists from very different academic branches with very different basic concepts and strategies (*e.g.* biophysics; econophysics). In these cases scientists make use of the same or to some extent similar sets of data, but discuss them in the respective language of their specific fields.

To be more concrete, a cell biologist may regard microscope images of living tissue as data. An astrophysicist, who is interest in gravitational waves, expects a series of numbers (or respective wave patterns in plots) as result of some measurements from highly specialised equipment. Figure 2.1 shows examples for these two cases from modern scientific investigations. How do these two specific examples highlight the differences in the meaning of 'data'? To answer this question, we can simply discuss features of sets of data that are only necessary for the one side, but not for the other. The feature of a direct mathematical form: the gravitational physicist needs mathematical data, because his theoretical assumptions and hypotheses are formulated in mathematical terms and make extensive use of the mathematical calculus in a non-trivial way; the cell biologist, in this example, needs her data in a form that enables her to inspect it with her human visual capabilities, because she uses these capabilities to draw inferences in her field.

[1]The original caption reads:

> (A) A thin-section electron micrograph of H. neapolitanus cells with carboxysomes inside. In one of the cells shown, arrows highlight the visible carboxysomes.
>
> (B) A negatively stained image of intact carboxysomes isolated from H. neapolitanus. The features visualized arise from the distribution of stain around proteins

In this thesis I defend the view that there is in fact no substantial difference between these and other different forms in which data can be presented or circulated. But the use of the concept in the various very different scientific fields and its subfields makes it challenging to pin down one useful explication of data that applies to all scientific applications.

One quite simple route to answer the question to what 'data' in the different scientific fields refers to is a Wittgensteinian idea from the *Philosophical Investigations* (1953) about human speech acquisition. It states that there are no strict and specific primary concepts to which the nouns of our language refer to, but rather acquired habits of the word's use in our linguistic communities.[1] This idea implies that 'data' can refer to very different and unrelated things in different communities as long as the scientists receive their training and the acquisition of their field's terminology, which involves the reference to data, in strict separation from other scientific communities (which does not seem completely unrealistic, given the actual customs of our scientific education). However, I believe that there are general epistemic rules of scientific reasoning and a general epistemic framework for it. A notion of data is a substantial part of this framework and one can give a descriptive explication. The Wittgensteinian position implies that a universal notion of data for empirical sciences would be a fortunate coincidence or has to rest on some pre-specialised notion of data from our shared everyday language. The latter option cannot be easily disputed. In the following course of this chapter I focus on a descriptive account of data rather than on reasons why such an account exists, whereas Wittgenstein's account can be seen as an approach to consolidate my position that such an overall account of data is reasonable. I give philosophically less contestable arguments for my view by the actual elaboration of my account of data and by the reference to several examples from science with which I aim to show that the account applies in the suggested generality.

Concerning problems with the explication of a concept of data I distinguish the two following general aspects:

> **Problem of definition and demarcation:** A set of data is usually not scientifically discussed in strict isolation, but instead with additional infor-

---

forming the shell as well as around the RuBisCO molecules that fill the carboxysome interior. Scale bars indicate 100 nm.

[1]Wittgenstein's position about the reference or meaning of words in the *Philosophical Investigations* is best described by the following quote:

> For a *large* class of cases–though not for all–in which we employ the word 'meaning' it can be defined thus: the meaning of a word is its use in the language. And the *meaning* of a name is sometimes explained by pointing to its *bearer*. (Wittgenstein 1953, §43)

mation and assumptions about the part of the empirical world it was gathered from and the equipment that was used to gather it, as well as with theoretical descriptions of relevant methods and equipment. This set of information is—obviously—vastly complex. If we want to refer to a concept of data in a diligent philosophical way, then we need to clarify where exactly the border lies between a set of data and additional information *about* the data.

Ontological classification: What is data ontologically speaking? Some sets of data are given as mathematical objects, like a finite series of numbers, but others are given in a human language, as sounds, photographs or other forms that do, prima facie, not seem to be mathematical objects. In other influential debates from the general philosophy of science *evidence* from *observations* are explicated or modelled as logical propositions[1] or events in a probability space[2], which can be interpreted to imply that data belongs to one of these ontological classes, too.

The problem of definition and demarcation is, according to my knowledge, only rarely the subject of discussion in the philosophical literature. Most references to data, in particular in the discussion about phenomena that we scrutinize within this thesis, do not imply or presuppose *any* precise notion of data. The reason for these vague kinds of reference seems to be that, firstly, the articles and talks usually discuss very specific examples of scientific phenomena (*e.g.* the melting point of lead; weak neutral currents and bubble chamber photographs; the existence of an extrasolar planet) with very specific exemplary sets of data. In these examples, the specific sets of data are described in detail and the demarcation between the set of data and other information that play a role in the inferences from data to theories (*e.g.* description of the measurement routine) is more or less apparent. In other cases, examples are described without a specific demarcation between the data and further information about it. Secondly, given the very specific exemplary setups, a more precise concept of data does not seem to illuminate the examples more regarding the aspects of the inferences from data to phenomena or from data to theory or vice versa, which the examples are intended to highlight.

The ontological classification of data has to be answered in a larger metaphysical and also epistemological context. With this I mean that a philosopher's answer about how he ontologically classifies data (*e.g.* as real abstract objects; as syntactical propositions; as physical objects) reveals a lot about his philosophical beliefs regarding metaphysics and epistemology in general. The reason for this is that

[1] I refer to discussions that started with Hempel's accounts of scientific confirmation (1943; 1945a; 1945b) and explanation (with Oppenheim 1948).

[2] I refer to Bayesian accounts in general, for example the discussion on scientific confirmation; Fitelson (1999) provides an introduction and overview.

data plays a crucial conceptual mediating role between the empirical world and our human theorising.

I do not aim to provide a fully fleshed out ontological survey of the basic concepts from general philosophy of science (*e.g. observation*; *evidence*; *theory*). However, I want to provide a sufficient ontological account about *data*, *patterns* and *phenomena* to provide the answers to the main questions about phenomena that were introduced in this thesis.

If we restrict our survey about data to empirical investigations that are undoubtedly *scientific*, then reasoning without a sufficient amount of data support is usually accompanied with profound practical obstacles and the need for expensive solutions to gather useful data. Examples of such cases with the problem of underdetermination include many fields of social sciences and economics (*e.g.* political surveys; risk distributions for a stock positions). But also questions from physics on the extreme micro or astronomical level demand laborious ways to gather useful data (*e.g.* CERN Large Hadron Collider; extraterrestrial telescopes). The reason for these mentioned cases with the problem of underdetermination in social sciences is that social systems or markets are non-stationary phenomena that change quickly over time, are vasty complex and the amount of data that can possibly be gathered is simply restricted by the fact that there is only one or at maximum very few societies or markets. In the mentioned exemplary cases from physics engineering and financial limitations restrict our abilities to gather every set of data that we are interested in via appropriate measurements. However, we cannot infer any distinctions in principle between these sets of data from these distinguishing aspects regarding the explication of data and the more general framework of inference from data to phenomena, which I aim to provide.

What makes data *scientific*? The demarcation of science from non-science or *pseudoscience* is a subject of discussion in the philosophy of science. Concerning isolated empirical questions and the methods to answer these questions, which include the processing of observation in any way, the most influential approach to detect non-scientific theories is Popper's (1962) falsificationism.[1] I aim to avoid examples from science that may be declared as unscientific according to Popper's criterion for scientific data. I also focus on the often mentioned criterion of reproducibility[2] for observations or sets of data that we want to discuss. This criterion is relevant to many standard cases and all our discussed examples of scientific in-

[1]He states that an empirical theory is scientific, if it is "capable of conflicting with possible, or conceivable observation" (p. 39). A theory must be testable by a routine to gather data, such as a measurement, and would turn out to be false if the resulting data does not show a specific pattern that was predicted under the assumption of the theory being valid.

[2]We find this criterion explicitly stated in Vollmer (1992) and Merton (1973); Mahner (2007) provides a more thorough survey of criteria for science in the philosophical literature.

vestigation. But it is not suitable for a strict demarcation between scientific and non-scientific theories, due to the already mentioned lack of possible observations for decent scientific investigations; these include the explanation of the following phenomena: the Big Bang, the political success of Hitler and the sudden extinction of large dinosaurs. For none of these phenomena we are able to reproduce sufficiently useful repetitions of the observations.

The view that I lay out here does not imply that a set of data has to play any role in a scientific discovery. It also does not imply that data has to be gathered with the aim of being fruitful for any scientific discovery. Here, as will be shown, I disagree with Leonelli (2015, my discussion in 2.1) who claims that data is data due to its "prospective usefulness as evidence". Some data of anthropology and historical sciences (*e.g.* outline of the arrangement of the Giza pyramids; cataloguing of biological species) was not produced or gathered with the aim of being evidence for a specific hypothesis. Leonelli herself mentions DNA sequences to promote the related point that data is often gathered without the aim to validate a specific scientific claim, but rather to catalogue data that may become interesting to the field for whatever scientific reasons in detail. However, she does not include cataloguing data as an end in itself for her notion of data. Tycho Brahe's[1] gathering of astronomical observations, which was intensely and fruitfully used by later astrophysicists, is a more historical example that supports our claim that data does not have to play an inferential role as evidence (or may be used to defend other claims). A set of data can also turn out to be bad or useless due to, for example, a broken measurement device or empirically inadequate theoretical assumptions. But it still counts as data, according to my notion.

What does the title of this chapter mean? The distinction between abstract and concrete objects is a subject of intense metaphysical discussion. However, this distinction is relevant to the problem of ontological classification of data. To characterise abstractness, concepts of it for *mental* objects, *sensible* objects, *physical* objects and *causality* are employed.[2] For our purposes, the term is used in a very simplified way. Basically, I call an object *abstract*, if it is a mathematical object. I justify this terminology in 3.1. I motivate this terminology by the fact that the abstract objects that are discussed in this text are solely mathematical objects, but I want to highlight their property of being abstract objects. My philosophical position is that data in science are always mathematical objects.

[1]Tycho Brahe (1546–1601) was a Danish astronomer, who provided very precise records of the positions of fixed stars and planets without the aid of a telescope, which had not yet been invented. Brahe's data was of crucial importance for Johannes Kepler's claim that planets move on ellipses and not on circular trajectories around the sun. For a historical survey on Brahe see Ashworth's (1999) encyclopaedia entry.

[2]For an introduction to abstract objects see Rosen's (2012) encyclopaedia entry.

Given our very broad understanding of science (*e.g.* including social sciences and historical studies), this position seems quite radical prima facie, but I argue that the crucial role of data in science, which is to be analysed for patterns, cannot be explained without mathematics.

A notorious problem of abstract objects can still be highlighted regarding our account of it. Opponents may argue that a set of data, as an abstract object, is simply *not there* before it is gathered by the execution of an experiment or measurements. But this argument applies analogously to mathematical objects in general: a probability space, for example, is and always was something mathematical and abstract in our sense independent of Kolmogorov's (1933) first formulation. As elaborated on in 3.1, in my view mathematical objects are so-called *ante rem* structures or positions in those. This view implies that the ontological classification of mathematical objects or propositions does not depend on the fact whether a human has defined or imagined it, or whether they are unknown to us.

Why should we care about a more precise concept of data?—Other authors with a focus on scientific phenomena do not do this. The reason why I bother with explicating the notion of data is that I want to provide a very general explication of phenomena and their relation to patterns in data. Therefore, we need to make precise what patterns in data are, and the first step to achieve this is to explicate the concept of data itself. As it turns out, specific philosophical views on data and patterns in them are necessary to satisfyingly pin down a notion of phenomena. These views are often not explicated by previous authors due to the aim of their articles, limited printing space or the assumption that a sufficiently common agreement about these matters can be presupposed (which is not the case, in my opinion).

2.1 Available Literature on the Explication of Data

Section Abstract

I elaborate on some influential historical and recent philosophical positions concerning an explication of data in science. Suppes distinguishes between *data*, *models of data* and *theories of models of data* in a formalised framework. In this framework theories, which are something linguistic, are strictly separated from models, which are something set-theoretical. According to Hacking's view, data are *marks*, which are material externalisations of human acts. For Leonelli, data is material, too, and it is characterised by its functional role as evidence and its portability.

Some philosophical literature about what data is is available. I introduce positions that I deem noteworthy and helpful for our discussion.

Suppes: Data in a Hierarchy of Theories with Models

Patrick Suppes (1962) is one influential source for an attempt to explicate data. He introduces *models of data*, which are, roughly speaking, possible outcomes of data after extracting from *raw* observational data only the aspects that are relevant for a specific inference between data and hypothesis. This implies that a model of data does not depend only on the methods and equipments to gather the data, but also on the theory in question to confirm or to corroborate or to estimate parameters of.

Let us investigate his idea about data a little bit further; this investigation helps to shed more light on the distinction between data and other concepts that play a part in scientific reasoning and are not data (*e.g.* theories; phenomena). For Suppes "a theory is a linguistic entity consisting of a set of sentences" in, for instance, a logical language. In his terminology, theories can be *realised*, whereas "a possible realization of a theory is a set-theoretical entity" (1960, p. 5). His example is: the theory of algebraic *groups*[1] is the list of group axioms, and an actual group, *e.g.* $\{\mathbb{R}, +\}$ or $\{\mathbb{Q}\backslash\{0\}, \cdot\}$, is *a* realisation of it.[2]

This distinction between the linguistic theory, and its realisations as set-theoretic structures is an adaption from Tarski's (1953) logical model theory. Among many others, Ebbinghaus *et al.*'s (1994) introductory textbook on mathematical logic is one example of this influence. Here, logical theories, which are a finite list of

---

[1] A *group* is defined in purely logical terms by the following list of formulae with variables a, b, c, constant e, the unary function $\cdot^{-1}$ and the binary function $\cdot \circ \cdot$:

(G1) $a \circ (b \circ c) = (a \circ b) \circ c$

(G2) $a \circ e = a$

(G3) $a \circ a^{-1} = e$

The logical group axioms G1-G3 express the full logical structure about a group, but, however, only the formulation of a group with sets adds an additional property: the closure under $\cdot \circ \cdot$. This can be formulated as the property of a model

(G*4) All variables in G1-G3 refer to elements in a set $\mathcal{G}$ for which holds that

$$\forall\, a, b \in \mathcal{G}\, (a \circ b \in \mathcal{G})$$

Realisations (or *models*) of a group based on a set $\mathcal{G}$ and the binary function $\cdot \circ \cdot$ are written $(\mathcal{G},\circ)$.

[2] In fact, other authors, including Tarski (1953, ch. III), and Balzer *et. al.* (1987), use the same example due to its simplicity and clarity to exemplify related points. However, seen from a more critical angle, this repeated use of the same example also sheds light on the fact that other, non-algebraic but still convincing examples from scientific theorising and mathematics are not easy to find.

axioms, have *models*, which are set-theoretic objects that fulfil the axioms of the theory like the above example of group theory shows:

> A possible realization in which all valid sentences of a theory T are satisfied is called a *model* of T. (Tarski, 1953, p. 11)

What then is *data* in Suppes' account? Data *for* a scientific theory is a possible realisation of it that fulfils further constraints, so that not every possible realisation can be data. *Functionally*, data has to be gathered by some experiment or measurement. *Mathematically*, data has to be finite or, respectively, at least bounded, since, for example, "no actual experiment can include an infinite number of discrete trials" (1962, p. 254). This mathematical criterion hints at the fact that the gathering and recording of data has to be *feasible*, meaning that it must at least be possible to gather this set of data by an actual action of an apparatus with or without any human interaction.

What are *models of the data* for Suppes? A model of the data is a set-theoretic object, which is consistent with the elaborations on Suppes' concept of a model as given above. But according to this criterion, we need to have a *theory of models of the data* (sic!), which are the sentences that have to be satisfied by the model of the data. The theory of models of the data are sentences that describe what properties the data *should have* (*e.g.* be a series of length 10 of real numbers between 2 and 4). One can give a theory of models of the data only if one has sufficient theoretical knowledge about, on the one hand, the scientific hypothesis that motivates the scientist to gather the data and, on the other hand, the entire measurement procedure used to gather the information that will be transformed into data.

As opposed to data, a model of the data has not to be gathered by some experiment or measurement, but can be inferred from purely theoretical considerations. That is why a *theory of models of the data* has to be distinguished from a *theory of data*: the latter one may include claims about the data that are irrelevant to the inferential role of the data for the scientific hypothesis in question. Those irrelevant claims occur due to the rather mundane fact that experiments and measurements can often not be designed to produce only the results of observations in which the scientist is interested in the specific context of conducting them. A *model of the data* does not encode these irrelevant features of the data. But to be able to pin down a model of the data, theoretical knowledge of the experiment or measurement must be available as well. This knowledge can be vastly complex and may involve large social groups of collaborating expert agents who carry different pieces of the relevant knowledge. To sum up this view more generally, a scientific theory and the most basic descriptions of how the raw data is gathered are "two extremes between which a hierarchy of theories and their models is to be fitted in a detailed analysis" (1962, p. 255).

The data may have features that are not features of the model of the data, because they are not relevant for the inferential role of the data in the relevant scientific context. However, many inferences between data and hypothesis incorporate a treatment of *noise*, which is the distortion of the *idealised* data that would have been gathered without any theoretically disregarded or unknown or pragmatically unavoidable influences. This is my proposed notion of noise, which is particularly suitable for statistical physics. Noise may not be mistaken for uncertainties that are intentionally described by the theory in question (*e.g.* by the wave function in quantum mechanics). Where do we have to locate the noise in Suppes' distinction between data and models of data? The answer to this question helps to further clarify Suppes' notion of models of data. There is an important distinction between noise and the features of the data that are certainly not features of a model of this data; noise is—adopting some terminology from statistics—the true *residual* in the data after all the theoretical knowledge (about the measurement, the experiment, the empirical part of the world etc.) involved at the inferential process is taken account of; noise are the features of the data that remain unexplained by the theoretical knowledge. But the with the model of the data intentionally disregarded features of the data are the features that the scientist can disregard concerning the inferential role of the data by reference to his theoretical knowledge. That is why I understand under Suppes' notion of models of data not the exclusion of noise. It is very common in scientific inferences from data to hypothesis to formalise statistical tests that take an acceptable level of noise into consideration. That is why we have good reasons to simply add a stipulated noise term (*e.g.* a Gaußian random variable called 'ε') to the theory of the model of the data. This noise term marks a noticeable difference between a model of data (having noise) and an idealised outcome of data (being free of noise). However, in some cases a model of data may specifically exclude noise, for instance, the only two-valued answer to the question 'Are the lights on in this room?'.

What conceptual role, under closer scrutiny, does a model of the data play in the inference from observation to theory? It should, in the context of an already understood experimental design, exemplify all the information from the observation that is relevant to an affirmation or falsification of the hypothesis in question and strip away all the irrelevant information. This concept is most apparent in statistically formulated scientific setups:

> The characterization of models of data is not really determined, however, by relevant information about experimental design which can easily be formalized. (...)

> The central idea, corresponding well, I think, to a rough but generally clear distinction made by experimenters and statisticians, is to restrict models of the data to those aspects of the experiment which have a parametric analogue in the theory. A model of the data is designed to incorporate all the information about the experiment which can be used in statistical tests of the adequacy of the theory. (1962, p. 258)

These ideas from statistics illuminate Suppes' approach, but I want to point out that these ideas are by no means restricted to a statistical framework. I claim that we can find analogies in *all* frameworks of empirical scientific reasoning independent of the degree of mathematical explications involved. I see no reason why for Suppes' account the "parametric analogue in the theory" (*e.g.* the melting point of lead as a temperature in degree Kelvin), which is statistically testable in a mathematical framework, should not translate to important aspects in a framework of scientific inferences without any worked out mathematical tests and even without data from experiments (*e.g.* Hitler's 1939 perception of America's future influence on the war[1]).

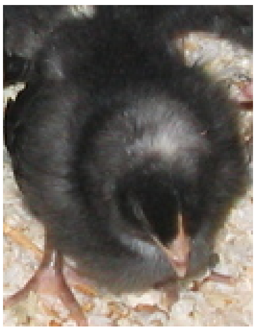


Figure 2.2: Data for inspection of a chick

The model of the data plays a middle level role in five levels of conceptual abstraction between the observational data in its *rawest* form and the model of the actual scientific theory in question. I further describe these five levels (plus the theory itself on top) by figure 2.3 with the help of a very simple example. This example is the distinction between females and males among one day old chicks by visual inspection.[2] Photographs of the chicks play the role of data in this example. Figure 2.2 shows the data of this example.

I intentionally selected an example of an everyday inference between data and theory, rather than a mathematically formalised one such as Suppes' case of *learning theory*. The reason is that, again, there is no principle reason to distinguish mathematically explicated inferences from ones that are performed by a human (*e.g.* identifying a black raven), an animal (*e.g.* finding a truffle) or an apparatus (*e.g.* detecting smoke in a building). My example fits our purpose much better than Suppes' example about learning theory, which is formulated in a formalised probabilistic setup. Furthermore, my example is easy to understand without any lengthy introduction or expertise in mathematical statistics.

Concerning figure 2.3, it may be debatable which claim has to be assigned

[1] Data in this example include original documents, testimonies, secondary sources and the like.

[2] Chickens have cloaca and can therefore not be distinguished as easily as mammals. For farming purposes distinction procedures were developed, including a breeding that aims to make chicks distinguishable by visual criteria like down colours.

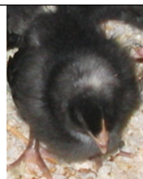
Theory of ("linguistic")	Exemplary claims of theory	**Model** of ("set-theoretic")
Chicken sexes	- There are two sexes - Only hens lay eggs - Only roosters crow	
Sex selection model	- White spot on the head shows male - Dark grey spot on the head make the sex selection fail	`m` ∈ { `female`, `male`, `undecidable` }
Models of experiment	- Image processing for better visibility (*e.g.* contrast enhancement) - Use more than one photograph per chick	
Models of data	- It must look like a chick - Head must be detectable - Spot on the head must be detectable - Every other datum is irrelevant	
Experimental Design	- Photography equipment - Distance and angle for photograph	
Ceteris paribus conditions	- Lighting conditions - One day old chicks	

Figure 2.3: Adaption of Suppes' (1962, p. 259) schematic view on the levels of conceptual abstraction between observation and theory for a much more mundane example. (Original photograph by `FreeImages.com/shelley Cunningham`)

to which level in detail; what part is of the *ceteris paribus* conditions, or of the experimental design, or of the experiment is in some cases only a matter of stipulation motivated by very practical concerns. Suppes emphasizes his aim to provide a framework in which the "[t]heory at one level is given empirical meaning by making formal connections with theory at a lower level" (1962, p. 260). As my choice of the example already hints at, I reject a fundamental epistemological or ontological distinction between formalised and non-formalised sciences, branches or specific theories. Suppes' example is formulated in statistical terms and the data is used to estimate parameters. This approach may help to clarify his ideas concerning their applicability to an already worked out formal framework. However, such a worked out formal framework is not necessary for his general concept of the many levels of theories.

The *models* in figure 2.3, right column, are mostly images and obviously not "set-theoretic" in any direct sense. However, the information given by an image can easily be stored in form of a purely mathematical structure (figure 2.4 on page 42 shows an example). Image storing and processing with computers is a simple example for this. I do not agree with Suppes predominant implication that mathematical structures have to be—or at least: are usually—expressed in set-theoretic terms. Mathematical structures are rather those abstract objects, that can be unambiguously expressed in principle, and therefore everything that can be transformed into a digital computer signal is mathematical by definition (but not vice versa due to infinities in mathematics).—For more details on this view and the mentioned examples see 3.2 of this thesis. Suppes', Tarski's and other's preferences for set-theoretically expressed models seem to be motivated by some *realistic* (in the metaphysical sense) implications concerning sets or mathematical objects in general that do not apply to logical propositions, which are in their view therefore something *linguistic* and not set-theoretical. However, if sets are *real* in some sense, since we have some imagination or intuition or *Anschauung* about them, then two-dimensional images have to be as well. A historical perspective without metaphysical implications is that Tarski's model theory is an approach to make predicate logic applicable to proofs and other forms of inference from the actual body of mathematics.[1] If these proofs or other inferences are noted in strictly set-theoretic terms, which was and still is an influential claim, then strictly set-theoretic models are a very natural approach for Tarski's goals. As I conclude a couple of paragraphs below more thoroughly, Suppes' focus on set-theory is too narrow and seems to have been influenced by Tarski's noticeably different goals

[1] Tarski (1953, p. 5) introduces the distinction between theories and models in this way with reference to Hilbert's and Bernays' (1934/1939) predicate logical approach to mathematical reasoning.

for introducing set-theoretic objects and call them 'models'.

To conclude, what do we learn about data from Suppes' ideas? He offers an approach to conceptualise theories and data in one and the same framework via models, which may seem counterintuitive, on the background of the classical logical empicists' and in particular Carnapian (1966) picture of a more or less strict epistemic distinction between *observables* and *theoretical terms*. This distinction is, more precisely, that for Suppes theories already come with models—whereas I do not care much about whether they *belong* to the theory not—and data comes in the light of a specific theory with a model of data. The inference from observation to theory happens via abstractions of the models from the different conceptual levels; Carnap, on the contrary, needs to stipulate "correspondence rules" to connect theory to an actual observation. For our aim to understand what data is, Suppes' view implies that theories are descriptions of possible data that enter several levels of conceptual *modelling*, rather than a vehicle to stipulate "observable or non-observable" entities, which is an important aspect of Carnap's view. In contrast to Carnap's notion of observation, Suppes' notion of data puts much more weight on the actual data gathering regarding the formulation of a theory, since the theory has models, which are, in the end, abstractions of gathered data.

Suppes explains why interesting aspects of data are explicable *before* actually gathering the data (because the data has to show the aspects according to a model of data). Furthermore, he introduces the idea that the hypothesis in question is not the only theory involved, because theories also occur at the level of data and experiment. Bogen and Woodward's (1988) (cf. 5.1) motivating example for the discussion about phenomena is that bubble chamber photographs for the detection of weak neutral currents are selected from a much larger collection of the experimental results from which most of the resulting photographs are ignored. This example can straightforwardly and successfully be explained with Suppes' concept about models of data and models of experiment.

I agree with the idea to offer a comprehensive formal approach for scientific inference from observation to theory via different conceptual levels of abstraction, and I also agree with the predominant role of mathematics for the models in these levels. But I disagree with Suppes' view that the models, which are always mathematical, have to be set-theoretic or at least set-theoretically expressed. Modern mathematical structuralism (see part 3.1) offers a too convincing fundamentally non-set-theoretic approach for a characterisation of mathematics to claim that mathematics is best described as a business of working on a set-theoretic fundament. In much later works, Suppes (2002) adapts his approach to the mathematical developments towards a non-set-theoretical vocabulary for the expression of mathematical structures. He admits "that a pluralistic attitude toward the concept of structure [= mathematical objects or propositions] can be taken. The modern mathematical theory of categories [(find a definition on page 74 of this thesis)] provides other arguments [against sets as the conceptual basis or fundamental language for mathematics] of a different sort." (p. 35) And he defends the choice of set-theory as the vocabulary for the expression of structures for scientific purposes as pragmatic:

> Total adherence to a reduction of all entities to sets, or sets and individuals, is not really crucial to the viewpoint adopted here. It is also natural and sometimes easier to use creative definitions of identity to make available new abstract entities that are not necessarily sets. Such definitions are called *creative* because new propositions can be proved with their introduction. Already in my book on set theory (1960) I departed from pure set theory (...) (p. 34)

After this passage he mentions examples of these abstract entities, including ordered pairs (which I also use and formally develop as an example for category theory at part 3.1). It may be fair to speculate that Suppes' often cited articles about models in science (1960) and models in data (1962) would imply less reference to set-theory, if he had published them much later, for instance after 2002.

Hacking: Data as Material Externalisations of Human Acts

Ian Hacking (1992) provides a description of data for the "laboratory sciences". His notion of data is embedded in a more general conceptual framework for these sciences, and this framework is "metaphysics and epistemology, a contribution to our radically changing vision of truth, being, logic, reason, meaning, knowledge and reality" (p. 29). Be that as it may, with his text he aims to provide support for the claim that laboratory sciences "tend to produce a sort of self-vindicating structure that keeps them stable" (p. 29–30) throughout the historical changes of our scientific knowledge. 'Laboratory sciences', in his terminology, is only loosely defined as studying "phenomena that seldom or never occur in a pure state before people have brought them under surveillance" with the use of "apparatus[es] in isolation to interfere with the course of nature that is under study" (p. 33). What are *data* in his account?

> *Data:* what a data generator produces. By data I mean uninterpreted inscriptions, graphs recording variation over time, photographs, tables, displays. These are covered by the first sense of my portmanteau word "mark." (p. 48)

I comment this quoted passage. *Data generators*, in his terminology, are objects or agents that produce data, for instance "people or teams who count", "micrographs" or "automatic printouts". (p. 48) He generally distinguishes between three general groups of "elements" for his description: *ideas*, *things* and *marks*. Things and ideas are very much what our common understanding of these concepts suggest. Marks, however, are a concept that is more unique to Hacking's account. He introduces *marks* as "outcomes of an experiment" or "manipulations of marks" (p. 44). Given that we want to know what data are and we read above that they are marks, this answer is not informative and almost circular. He further states that "marks are 'visible impressions', 'signs or symbols that distinguish something', 'written or printed signs or symbols', 'indications of some quality' and also 'goals' ". These things are hard to put into one class by any traditional philosophical approach (epistemological or ontological); however, Hacking later clarifies that "marks are things", but not all things are marks. (p. 44)

But how can *impressions* or *indications* be things? Hacking clarifies his views, implying that the data (things) are some sort of *externalised* human acts:

> Some will pleonastically call such marks [= data as introduced above] "raw data." Others will protest that all data are of their nature interpreted: to think that there are uninterpreted data, they will urge, is to indulge in "the myth of the given." I agree that in the laboratory nothing is just given. Measurements are taken, not given. Data are made, but as a good first approximation, the making and taking come before interpreting. (p. 48)

"The myth of the given" is a reference to the influential debate inaugurated by Wilfrid Sellars (1956), whose position is best described by Eric Watkins' (2008) summary that "empiricists are mistaken in thinking that what is given through sensibility could be sufficient for knowledge, since sensations do not have the structure, however it is characterized, that knowledge has".[1]

What can we take from this quote? It is important to note that the "making and taking" of data is not "interpreting" them, which is completely consistent with an everyday description of scientific work. The "making and taking" can best be classified as (human) acts in very general. No theoretical description of them is necessary, but only a sufficient *instruction* for the agent (*e.g.* experimental physicist; programmer). As an example, for LIGO's 2015 data to be a thing that is isolated from any mind and interpretation I understand the notion of a thing in this context in a way that this set of data is the actual storage statuses on the hard drive on LIGO's data server; it is the magnetic configurations. However, one can make the point that a mark can be a mathematical object (*e.g.* a series of numbers), too, given a sufficient degree of mathematical realism.

[1] A good starting point to further investigate the discussion is, in my opinion, Watkins article. Other influential philosophers, including Donald Davidson (1980; 2004), Robert Brandom (1994) and John McDowell (1994) responded to the discussion, whereas Kant (1787) provides an implicit influential position on the subject—without, of course, using Sellar's label for it and motivating the eponym *neo-Kantianism* for a modern adaption of his stance in this discussion.

The description of data as things, as opposed to set-theoretic objects in Suppe's account, accentuates that a set of data can be as bad scientifically understood as any other thing in our world. For Suppes, a "possible realisation", which a set of data is, is a possible realisation of a theory. That is why he needs to introduce theories for every level of conceptual abstraction. Hacking opens the door for data being some outcome of an act, which in itself does not have to be completely scientifically understood and described by theories. If, hypothetically, data is gathered to support a specific hypothesis, but the outcome is negative and the data shows only results that cannot be explained by any known theory, then, according to Suppes account, the model of the data and the actual set of data at hand do not have any important common features. That is why Suppes account neglects unexpected data and could be called *blind* to it. A further important argument against Suppes' account to explicate data in science in general is that in almost no realistic scientific episode of important inference between data and hypothesis are all the conceptual levels of theories clearly described or available; it would be a lot of work to actually reveal all the intermediate theories in a non-trivial but realistic example from, for example, gravitational physics. However, negative measurement results are data too and Hacking's account provides an open door for any kind of unexpected data.

Leonelli: Data as Portable Material Objects

Sabina Leonelli (2015) describes data as "first and foremost, material artifacts", that are "essentially fungible objects, which are defined by their *portability* and their *prospective usefulness as evidence*" (p. 811). Why is the materiality of data important? "The crucial role of portability is also what leads me [Leonelli] to characterize data as material artifacts, independently of whether they are circulated in a digital form or not." (p. 819) According to her account, strictly speaking, every material object can be data, if it is some "research output" and plays the role to "provide evidence for knowledge claims of interest to the researchers involved" (p. 811).

I criticise this view, because, roughly speaking, 'data' does not refer to material objects in any strict sense (*e.g.* X-ray photographs; sheets of paper with numbers on it; specific configurations of magnetic polarisations on a computer hard drive). A distinction between *material* and *abstract* is well established in ontology, fruitful and should not become indistinct here. Instead 'data' refers to the *information* that is illustrated or stored by the object. Information is something structural and abstract, and not something material. In my view, this information is structural and therefore mathematical. This difference is more important than it may seem at first sight, because in my view the inferential transition from something in the empirical world (*e.g.* a chair; a light beam; a scream) to an *object of our thought*—how Frege famously put it—has to be located in the process of producing or reading data and not at the interpretation of material objects. With this I mean that data is something that is already formulated in a humanly accessible and for humans unambiguous way (*e.g.* series of numbers; images); we would by no means consider an, for instance, invisible infrared light signal as data, but rather a somehow transformed visible image of it with the additional information that this image shows an infrared signal. My account is descriptively more adequate to what physicists and other scientists refer to by using the term 'data'.

The portability criterion has two important implications: firstly, data has to be genuinely non-idiosyncratic, meaning that it can be *read* by everyone with the proper background knowledge, which has to be accessible (unlike *e.g.* the golden plates that allegedly only Joseph Smith could read to write the book of Mormon). Secondly, it must be logistically possible to move data from one recipient to another. This second criterion is less trivial than it may seem at a first sight. Imagine an archaeologist, who finds some interesting excavation site of an ancient city. If we identify the old foundations of the city as interesting archaeological data in the context of the discussion of some hypotheses about ancient societies, then the portability criterion seems violated. We may save the portability criterion by claiming that everyone can just visit the excavation site. But this is not what Leonelli and also me have in mind. We would both not be willing to call the actual foundations the data, the actual data are the photographs or written measurement results that the archaeologists circulate in their working group. If an original photograph gets lost in a building fire, then the data is still available when copies were made; what the photograph shows is the data and not the photograph itself. However, my claim is the more general one that only the (structural) information on or in the circulated material objects (*e.g.* photographs; hard drives) is the data. That is why the portability criterion is well fulfilled in may account, too.

Furthermore, Leonelli (2009) criticises (Bogen and) Woodward's description of data as being "idiosyncratic to particular experimental contexts, and typically cannot occur outside of those contexts" (1988, p. 317) in contrast to phenomena. She discusses the example of DNA sequences, which are in many cases not gathered to provide evidence for particular claims about a phenomenon, but rather to provide a more general database for several possible claims about genetics in very general.

2.2 Data are Mathematical Objects

Section Abstract

Sets of data are mathematical objects, which are not defined by any functional role in scientific reasoning. Data comes always with a description of how it was gathered without this description being part of the data itself. Manipulation of a set of data must be explained by theories of measurement and experiment; some aspects of Suppes' approach are convincing. Data does not *represent* anything empirical, but material objects can represent data. Computer technology to store data exemplifies how data is mathematical. Scientific data is mathematical, because patterns can be explicated only mathematically. More specifically, a set of data is an equivalence class over some minimal structure.

Before we dig deeper into the philosophical discussion about why data is something mathematical, let us have a look at how data is actually processed and stored in everyday scientific applications. This helps to sharpen the view on what aspects of data and data processing are without much doubt mathematically explicable and for what aspects further argumentative work has to be done.

An interesting aspect of data is that it does not only appear in various forms (*e.g.* lists of numbers; photographs; texts), but is also *epistemically approached* in very different ways. Statisticians developed an extensive mathematical body for the recognition of patterns and the construction of models of data in rather simple mathematical forms, such as time series of numbers. Another approach to detect patterns in data, which may be called *semi-statistical*, is to provide unambiguous detection rules for human agents, but no full mathematical explication. Examples of semi-statistical pattern detection are the detection of a lung tumour by an X-ray image or the astronomical classification of stars (*e.g.* blue giant; red dwarf) by telescope images. The phenomena involved are relatively clearly described and false reports occur only very rarely, if the scientist or expert is well trained. But with many other forms of scientific empirical investigations, not even semi-statistical techniques of pattern recognition are applied. In historical, psychological or social studies, texts or the behaviour of a test person are the subject of pattern detection, but the patterns or the detection procedure are usually not mathematically explicated. Clinical psychology is an exemplary field in which the phenomena (*e.g.* psychopathy) are often described only very vaguely and the explication of the corresponding patterns is a substantial aspect of further investigation.

Data can be stored digitally or non-digitally. The following discussion about techniques of digital data storage and pattern recognition via software helps us in the further discussion about data and pattern recognition on a general epistemic level. Computers need digital transformations of the data to store it. Physical, non-digital storage techniques may be used for photographs, videos, sound tracks or documents in physical archives as, for instance, negatives or magnetic tapes. If data is translated into digital signals, then it is basically always translated into a finite list of binary numbers that can be stored using different techniques (*e.g.* magnetic tape; optical storage). Note that for this translation into binary numbers a multi-level setup of software plays a determining role. But this software setup does not obscure the fact that there is nothing more to a set of data than what can be mathematically explicated. Images are approximated by translating them into discrete and bounded grids of pixels (see figure 2.4) that are further coded into binary signals. Sound tracks are approximated by fragmenting them into discrete and bounded spectra of waves[1]. These digitalisation techniques, if applied in science, are optimised to store the aspects of the physically represented data that are *interesting* to the scientists of a specific community or who are engaged in a certain topic. These image or sound track digitalisations are designed to apply intended data analysis techniques by either computers or human beings. Digitally stored X-ray photos, for instance, are intended to be as informative to the naked eye as the original physically stored photos.

Strictly speaking, digitally stored data sets are examples of models of data in Suppes' (1962) sense. But the rules that govern which aspect or information in the data is stripped away are comparatively rigid and rather driven by technical standards than by specialised needs for a specific set of data in question (*e.g.* the scanner at the MCMP can provide a 600 dpi scan of a photograph). Most often the rules for approximating the data by the digital transformation are chosen very conservatively concerning the possible loss of information. If we store an image, we can easily approximate it by an image file that shows us the same picture to the naked eye, but not necessarily to a more indepth inspection with the help of more advanced visual inspection techniques (*e.g.* a microscope). However, due to needs concerning storage capacities or software restrictions, the approximation for digitalisation can be more substantial than just stripping away what a human agent would neglect in every data analysis routine for sensory and epistemic reasons without auxiliary equipment anyway.

For the concept of data that we want to explicate it is not of crucial interest whether data is stored or processed in computer systems or not. The point of interest is, whether computational storing and pattern recognition is something that exemplifies data perception and pattern recognition in data in general. If data and pattern recognition can satisfactorily be reduced to mathematical objects and algorithms applied to these objects, then these concepts can be explained more

[1]A translation of a signal into a spectrum of waves is mathematically explicated by the *Fourier-transformation*.

simply on the basis of our knowledge about mathematics. This idea leads to a program of a specific mathematical formalisation of data, patterns and pattern recognition. I will follow this approach under certain assumptions in chapter 4. I further explain my claim that the modern development of computer technology should strengthen the belief that there is nothing more to human pattern recognition in science than what can be mathematically explicated, because it can be implemented—see my elaborations in 3.3.

An aspect of interest concerning the *ontological classification* of data (cf. p. 23) is: what are sets of data that are originally not mathematically represented, such as the information on an X-ray photographs, ontologically speaking? If data were not mathematical, then pattern recognition techniques directly applied to them cannot be purely mathematical either and further explanation is necessary to clarify the concepts of pattern and pattern recognition. As mentioned, digital translation of data always results, basically, in a finite list of binary numbers. Therefore, some digitalised photograph is information stored as some digital signal and, in principal, nothing other than a finite mathematical object, such as the number four or a triangle. On the other hand, objects that *show* data such as photographs or sound tracks are not always directly analysed by mathematical algorithms. These are often analysed via the sensory impressions they cause in human beings. I state that in all of these exemplary cases the set of data is a mathematical object, and opposing intuitions are misleading. Rather naive arguments against my claim, but also some strong ones will be presented throughout this section and rebutted.

Further arguments can be put forward to strengthen the position that not all data is reducible to mathematical objects without the loss of important information. These are that some data are still stored physically for actual scientific applications (*e.g.* non-digital audio tapes; X-ray photographs), digitalisation is often conducted only for pragmatic storage needs (*e.g.* hundreds of thousands of night sky photographs) and translations into digital formats are never unique (*e.g.* different choices of resolutions for raster images). A further argument is that common scientific analyses of images or sound tracks *seem* very non-mathematical; recall the example of animal noises or historical speeches as data in science.

When we refer to data in scientific contexts, we do not refer to some physical object. We refer to what it shows, what sound is recorded on it or what the results of an analysis procedure applied to it are. Sets of data are abstract objects in the sense that they do not have a determinate position in time and space. Therefore, a physical object (*e.g.* piece of paper; computer hard drive) can only *represent* some data (with a notion of representation that is, admittedly, idiosyncratic to this thesis but thoroughly introduced in 2.3). The time-spatial position of a physical X-ray photograph is, of course, irrelevant for the data it shows, that is, the data it represents.

However, I claim that data are mathematical objects and patterns in data are mathematical properties of these mathematical objects, which seems more surprising in the cases of social and historical studies than for physics. As opposed to Hacking and Leonelli (see 2.1), I do not believe that data is material in any sense. A set of data is rather an abstract or linguistic object—this distinction is, in my view, not relevant, since I see no ground for a substantial metaphysical distinction between a logical proposition on the one side, and a mathematical or abstract object on the other side.[1] My account is agnostic about realism or nominalism regarding mathematical objects or propositions. When I claim that data are mathematical objects, I do not refer only to any model of data in Suppes' sense, but rather to the raw data itself (*e.g.* the actual measurement results) without any form of modification.

I need to strengthen my position on data being purely mathematical. I formulate my replies to two substantial challenges to my position, that I received so far.

Objection against Data as Mathematical Objects: Mental States are not Mathematical or Reducible to Mathematical Objects

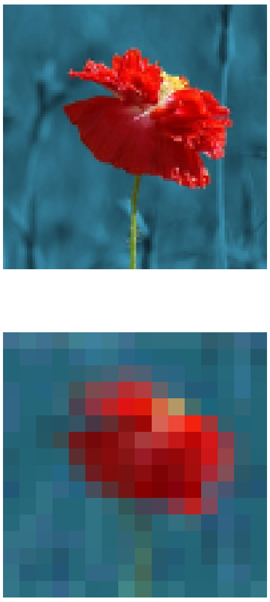
A lot has been written about the philosophical nature of *mental states*,[2] which are experienced feelings like a headache or the sensation of the colour red for humans. I do not intend to provide a thorough discussion of the field. But an objection to my claim from proponents of irreducibility of mental states to anything independent of a human mind (*i.e.* physical brain processes) would read like the following: a photograph or political speech may be transferable into a mathematical object and digitally stored as such, but the colours or the impression of a voice is by no means mathematical, though they are substantial parts of the data.

Figure 2.4, top left, shows an example of some observational data that is simply a photograph of a flower. In case you doubt this being a proper example of scientific data, imagine a botanisc who shows it to some audience as a newly discovered plant. In the figure, bottom left, we find an outline how the photograph might be stored as a raster graphics in a computer. Due to presentational reasons, the raster graphics shows only 16×16 pixels and is reduced to $4'096 = 16^3$ colours. We could chose more pixels and colours, but there are obvious practical obstacles

[1] As further elaborated on in 2.1, this distinction originates in the application of Tarskian model theory, which was invented to bridge the gap between predicate logic and actual mathematics, to philosophy of science by Suppes.

[2] I use 'mental state' in accordance with Putnam's (1967) classical paper.

with the complete presentation that I aim to give with the help of this figure. On the right side of the figure you can find the image transferred to a matrix with a common mathematical notation.



$$D = \begin{pmatrix}
\mathtt{368} & \mathtt{378} & \mathtt{368} & \mathtt{367} & \mathtt{378} & \mathtt{379} & \mathtt{378} & \mathtt{368} & \mathtt{267} & \mathtt{267} & \mathtt{267} & \mathtt{267} & \mathtt{256} & \mathtt{256} & \mathtt{368} & \mathtt{257} \\
\mathtt{267} & \mathtt{368} & \mathtt{368} & \mathtt{267} & \mathtt{368} & \mathtt{379} & \mathtt{378} & \mathtt{268} & \mathtt{268} & \mathtt{267} & \mathtt{267} & \mathtt{267} & \mathtt{256} & \mathtt{267} & \mathtt{378} & \mathtt{256} \\
\mathtt{267} & \mathtt{267} & \mathtt{267} & \mathtt{267} & \mathtt{267} & \mathtt{267} & \mathtt{367} & \mathtt{567} & \mathtt{467} & \mathtt{268} & \mathtt{368} & \mathtt{267} & \mathtt{267} & \mathtt{267} & \mathtt{379} & \mathtt{267} \\
\mathtt{368} & \mathtt{267} & \mathtt{268} & \mathtt{367} & \mathtt{455} & \mathtt{944} & \mathtt{845} & \mathtt{C22} & \mathtt{C33} & \mathtt{666} & \mathtt{268} & \mathtt{257} & \mathtt{257} & \mathtt{267} & \mathtt{379} & \mathtt{267} \\
\mathtt{368} & \mathtt{267} & \mathtt{267} & \mathtt{367} & \mathtt{445} & \mathtt{834} & \mathtt{B22} & \mathtt{D11} & \mathtt{D10} & \mathtt{D64} & \mathtt{CA6} & \mathtt{466} & \mathtt{267} & \mathtt{267} & \mathtt{367} & \mathtt{267} \\
\mathtt{368} & \mathtt{267} & \mathtt{256} & \mathtt{256} & \mathtt{356} & \mathtt{434} & \mathtt{633} & \mathtt{811} & \mathtt{900} & \mathtt{D76} & \mathtt{E43} & \mathtt{E21} & \mathtt{A33} & \mathtt{267} & \mathtt{267} & \mathtt{267} \\
\mathtt{267} & \mathtt{267} & \mathtt{256} & \mathtt{444} & \mathtt{711} & \mathtt{700} & \mathtt{800} & \mathtt{922} & \mathtt{A44} & \mathtt{B33} & \mathtt{C11} & \mathtt{B00} & \mathtt{D11} & \mathtt{934} & \mathtt{267} & \mathtt{267} \\
\mathtt{256} & \mathtt{267} & \mathtt{267} & \mathtt{267} & \mathtt{533} & \mathtt{711} & \mathtt{811} & \mathtt{811} & \mathtt{912} & \mathtt{801} & \mathtt{800} & \mathtt{A11} & \mathtt{C11} & \mathtt{A44} & \mathtt{267} & \mathtt{267} \\
\mathtt{267} & \mathtt{267} & \mathtt{267} & \mathtt{268} & \mathtt{356} & \mathtt{711} & \mathtt{800} & \mathtt{711} & \mathtt{811} & \mathtt{700} & \mathtt{811} & \mathtt{A11} & \mathtt{A23} & \mathtt{955} & \mathtt{156} & \mathtt{256} \\
\mathtt{257} & \mathtt{267} & \mathtt{378} & \mathtt{368} & \mathtt{268} & \mathtt{456} & \mathtt{533} & \mathtt{711} & \mathtt{800} & \mathtt{700} & \mathtt{801} & \mathtt{900} & \mathtt{656} & \mathtt{457} & \mathtt{256} & \mathtt{257} \\
\mathtt{267} & \mathtt{267} & \mathtt{378} & \mathtt{267} & \mathtt{267} & \mathtt{378} & \mathtt{367} & \mathtt{355} & \mathtt{445} & \mathtt{445} & \mathtt{834} & \mathtt{B22} & \mathtt{845} & \mathtt{268} & \mathtt{267} & \mathtt{267} \\
\mathtt{267} & \mathtt{378} & \mathtt{368} & \mathtt{267} & \mathtt{267} & \mathtt{378} & \mathtt{267} & \mathtt{366} & \mathtt{355} & \mathtt{267} & \mathtt{378} & \mathtt{367} & \mathtt{257} & \mathtt{368} & \mathtt{257} & \mathtt{267} \\
\mathtt{378} & \mathtt{378} & \mathtt{368} & \mathtt{368} & \mathtt{368} & \mathtt{378} & \mathtt{267} & \mathtt{367} & \mathtt{366} & \mathtt{267} & \mathtt{267} & \mathtt{267} & \mathtt{256} & \mathtt{368} & \mathtt{256} & \mathtt{267} \\
\mathtt{368} & \mathtt{367} & \mathtt{267} & \mathtt{256} & \mathtt{267} & \mathtt{378} & \mathtt{267} & \mathtt{367} & \mathtt{476} & \mathtt{256} & \mathtt{257} & \mathtt{267} & \mathtt{256} & \mathtt{378} & \mathtt{256} & \mathtt{267} \\
\mathtt{378} & \mathtt{257} & \mathtt{267} & \mathtt{257} & \mathtt{267} & \mathtt{368} & \mathtt{368} & \mathtt{377} & \mathtt{465} & \mathtt{257} & \mathtt{267} & \mathtt{267} & \mathtt{267} & \mathtt{379} & \mathtt{368} & \mathtt{267} \\
\mathtt{378} & \mathtt{267} & \mathtt{378} & \mathtt{368} & \mathtt{378} & \mathtt{267} & \mathtt{267} & \mathtt{488} & \mathtt{476} & \mathtt{267} & \mathtt{378} & \mathtt{267} & \mathtt{267} & \mathtt{378} & \mathtt{267} & \mathtt{378}
\end{pmatrix}$$

Figure 2.4: Top left: cut from an observation, bottom left: data that is sufficient for the inference, right: *one* of many possible mathematical explication of the data (a 16×16 matrix, the positions i and j for $d_{ij} \in D$ refer to the position of the pixels in the image, the three digit hexadecimal number per pixel refers to the RGB colour code).

My opponent may have read Frank Jackson's (1982; 1986) widely discussed knowledge argument[1], which I use as a paradigmatic example in the discussion of irreducibility of mental states. My opponent would stress the point that the impression of red, which the left images in the figure provide, is not provided by the mathematical object, which is described on the right in the figure. For Mary from the thought experiment, who never experienced seeing red, the left image and the matrix carry equivalent information, but for everyone else (without colour-blindness) the unique, non-physical sensation of red gets lost in the translation into a mathematical object.

Here is my answer: our human sensation of red has nothing to do with the *information* (or: datum) that there is red colour. Regarding the role of data in scientific inferences the proposition 'the flower is red' is as useful as a picture that shows only that the flower is red (and nothing more, which is—obviously—hard

[1]Jackson presents a thought experiment to argue in favour of the claim that not all *knowledge* is *physical*: the brilliant scientist Mary has a perfect physical and biological knowledge about the colour red, why objects can be red (*i.e.* wavelengths of light) and how we are able to perceive red with our human senses (*i.e.* how eyes and brain work). However, she never experienced seeing red, since she always worked and lived in rooms without red objects. Jackson claims that with the first experience of a red colour sensation, she learns something new, which is not physical knowledge (which in my approach is mathematically explicable knowledge).

to make up). There is no inference between data and theory that could be done only with only one of these two options and not with the other. That is why the matrix in the figure carries all the *information* that the photograph does and *is* therefore the data.—And that is why Mary, who never experienced seeing red, can still be a brilliant scientist with flawless and full-featured scientific reasoning: the sensation of red does not have any value for scientific inferences, it *explains* nothing and it does not have to be explained. In other words, mere sensations cannot enter any process of scientific inference, since they are *by definition* not externalisable, whereas externalisability is a predominant criterion of data, as Hacking and Leonelli point out, too. And this arguments still holds, if we remain agnostic about the irreducibility of mental states.

This point is related to the points I raise against Dennett in 4.2. Dennett claims that for humans detectable patterns in data (*e.g.* glider in Game of Life) have another metaphysical status than the ones to which we humans are blind to (*e.g.* AlphaGo's analysis results for a match of Go). I reject any relevance of this distinction for the purpose of scientific inferences. That is why Jackson's Mary can be a brilliant scientist.

Leonelli emphasizes that an important criterion of data is portability. Without even having to dig into computational theory, data portation is a topic of digital information theory[1]. Data has a mathematical form due to its crucial role of being portable. But do mental states, like the sensation of red, influence *creative* processes in science (*e.g.* designing experiments; finding new patterns)? The next objection deals with this question.

Objection against Data as Mathematical Objects: There is no Convincing Theory of *Good Data* or of a *Good Experiment*

A common objection against the claim that data and the inferences between data and hypotheses are mathematical objects and procedures is that, very roughly speaking, *doing* science is an *art* (whatever that means in detail). A more concrete aspect put into words by the same intuition is: designing a successful experiment that produces fruitful data in a difficult empirical context is a highly *creative act* (whatever that means in detail) that can by no means be understood as being logically or mathematically inferred by a set of mathematically explicated rules.

At 3.3 I elaborate in a more general context on the view why I believe that scientific inferences and the information and abstract objects involved can in principle always be mathematically explicated. In the following, I outline my answer

[1] Claude Shannon's (1948a; 1948b) ideas became influential, because he provides an approach to transfer digital data over long distances via conducting wires or electromagnetic waves without any loss of information. This is at the core of the theory of computer networks.

for the specific case of data that is useful for a scientific inference, and the methods of gathering it.

The above outlined objection is based on the assumption that a human being acts or thinks in some regard differently than any machine could do. I want to remind the reader—with reference to 3.1—that 'mathematical' does not mean only actually *explicated by an algorithm.* Something is mathematical, if it can be explicated in mathematical terms *in principle.* That is why the inferential work that a computer can provide (*e.g.* pattern recognition by an algorithm or neural net) is always mathematical. This includes the execution of vastly CPU-expensive, but terminating algorithms or executions in accord with rules that were established via machine learning and that may never be understood by a human being (e.g *AlphaGo* playing Go).

Given the fact that our body, and our brain in particular, is a biological cellular automaton, the view that a human being acts or thinks in some regard differently than any machine could do needs a convincing foundation. To be more specific, the argument for such a view should not only explain why a human being may have mental states and a machine does not; it has to explain why a human being can make inferences or act in a way that a machine can not. Some arguments for this claim were developed in the context of the philosophy of mind and have their roots in Descartes' *Meditationes* (1641).

A candidate to frame this issue is the discussion about *mental causation*, which can be regarded as being "at the heart of the mind-body problem" (Shoemaker 2001, p. 74). The *mind-body problem* refers to the philosophical tradition of claiming a general ontological distinction between *bodies*, which are spatiotemporal, and the *unextended* minds, which are traditionally related to a terminology of *souls* (Thomas Aquinas[1]) or *monads* (Leibniz 1714). Since these are two distinct realms, any sort of mutual influence between them has to be explained by, so to speak, *bridging principles* (in an ontological sense). A *non-reductive* stance about mental causation implies that mental events (*e.g.* having a thought) can cause bodily actions (*e.g.* tasting the water), but these actions are not caused by the physical brain functions that correspond to the mental event. To be clear, the difference between *free will* and mental causation is, according to this terminology, that free will implies the much stronger claim that human actions are not guided by fully deterministic rules (neither mental, nor physical).

By this terminology, electromagnetic processes in a computer are bodily processes, and human thoughts are mental processes. The critical view of my opponent at this point would be that inferences between data and hypotheses performed

[1] Aquinas' (1225–1274) main notes on souls coined this notion and are described in his writings *De spiritualibus Creaturis* and *Summa theologica*, which is one of his major works.

by a computer may be mathematical, but this tells us nothing about inferences, thoughts and decisions performed by a human mind, since a computer and also any form of mathematical explication does not carry any mental consciousness.

I agree with Shoemaker's assessment and want to reintepret it with respect to my topic by claiming that mental causation is at the heart of the problem whether data, data gathering and inferences with data are in principle purely mathematical or not.—What else than mental causation can draw a line between the capabilities of a computer and the epistemic capabilities of a human being? Without the philosophical introduction of mental consciousness, the human brain is a mere cellular automaton and it is hard to believe that we could not just classify it as a computer (with a very different architecture than our electromagnetism based computers, of course). That is why I focus on mental causation at this point. In the earlier objection I explained why mental states are irrelevant to a concept of data; a further step is this claim that mental causation does not play a relevant role for the scientific practices that involve data.

The literature on mental causation is extensive[1] and I do not aim to provide a thorough discussion. However, my philosophical conviction is that there is, overall, no convincing argument for mental causation in general and for our purpose in particular. In this regard I generally agree with the *epiphenomenological* camp in this discussion. Traditionally, the relevant discussion on epiphenomenalism starts with Huxley (1874), but newer empirical based investigations from the psychologist's direction—like the ones from Wegner (2002; 2004)[2]—strengthen my negative stance. No convincing *operational* criterion for the existence of mental causation has been proposed. Such an operational criterion would have to show how mental causation influences human actions for which a purely physical explanation is in principle insufficient.

As elaborated on in 3.3, recent software developments show, roughly speaking, that computer programs can accomplish tasks that some may claim to necessitate specific human *creativity* (which presupposes consciousness according to my understanding of the notion). Examples includes IBM Deep Blues' victory in a game of chess against reigning world champion Gary Kasparov in 1996, IBM Watson's victory against *Jeopardy!* champions in 2011 and Google AlphaGo's victory in Go over Fan Hui in 2015. Many examples of remarkable performances by visual pattern recognition software are available, too—see 3.3.

At 3.2 I also explain why the Church–Turing thesis aids my general conclusion that an actual implemented software *proves* mathematical explicability. Computational logic provides a carefully developed ground about what is computational

[1] See Robb and Heil (2013) for an introduction and overview.
[2] Find more on this at Robb and Heil, 1.2.

in general and what is not. However, my empirical arguments about the recent software developments rest on the constructive successes rather than fundamental boundaries.

One important criterion of a mathematical explication or mathematical proof or computational function is its finiteness or, respectively, termination. Is it possible that a *creative* human act cannot be finitely explicated in mathematical terms? If this was the case, then we have a simple argument at hand to oppose the view that a creative human act can always be mathematically explicated. However, for this to be the case we need some plausible explanation where the mathematically inexplicable infinities may occur. But epistemology and logic provide good arguments to believe that human scientific reasoning is best explicated by finite propositions and human scientific inferences are best explicated by the use of a logical language. And our scientific knowledge in biology, chemistry and physics provides good arguments to believe that the cellular automaton in our body works in accordance with deterministic rules (even if they are very complex). That is why there is not much room left to believe that mathematically inexplicable infinities are possible.

I introduced the opponent's view with the note that there is no *theory* or even any form of loose *guide* about designing an experiment or gathering data for the cutting-edge scientific developments. My opponent may claim that, such as with many other *practices* and *arts*, it takes years of experience to master these fields of science and every scientist has a very unique personal mindset and *style* about what he is doing. My answer to this opponent is that all of this is true, but it is no argument against my claim. The human brain and cognition are vastly complex (by any adequate notion of complexity) and so are almost all parts of the empirical world under investigation (*e.g.* physical; biological; human psych) by the different scientific disciplines. Empirical studies about the neural complexity and mere amount of neurons in a brain support this description. I do not see any reason why we should not believe that a groundbreaking experimental design can be mathematically inferred by a set of rules; the important point is that this set of rules is vastly complex and incorporates a lot of idiosyncratic knowledge and assumptions of the one experimental physicist. It was designed by a biological cellular *computer* that works with a very different architecture and magnitude of performance than our laptops or smartphones can provide.

Let us go on with the discussion about the claim that data is purely mathematical. The most compelling argument in favour of this claim is, in my view, that inferences from data to hypothesis are in many worked out scientific examples straight-forwardly mathematically explicated. Statistical parameter estimating and testing of null hypotheses, as well as *pattern detection* in general are applied mathematical algorithms. Bogen and Woodward (1988) introduce *phenomena*, which are the important intermediate layer between data and theory, as "patterns in data" on the background of this intuition: data are mathematical objects to which we apply routines to detect patterns. In 4.4 I discuss how Ulf Grenander's *general pattern theory* is a sufficient and purely mathematical approach to explicate patterns with sufficient generality. In chapter 5 I discuss how my notion of *phenomena* illuminates the inferences between data and theory, which are, in my view, in principle an application of a mathematical routine.

Most sets of data can serve only in a role in scientific reasoning with proper additional information about the experiment or measurement or its origin in more general. A series of tuples can only serve as evidence for the existence of black holes with the additional information that it shows the observed trajectories of a certain distant star captured by a specific telescope directed with a specific angle into space. But this information is not part of the data. I use the term 'data' as close as possible to the everyday use by scientists. This implies that a set of data has to be gathered via an observation or measurement routine. With 'observation or measurement routine' I do not refer to any elaborated concept of observation or measurement; this terminology refers only to the mere fact that data has to come with a description (in the most general sense of 'description') of how it was gathered by a human agent and/or a technical device.

What then classifies a mathematical object as data, if we do not imply any degree of actual use or usefulness for scientific reasoning at our concept of data? Data always has to come with a description about how it was gathered. The data does not *represent*[1] anything empirical, but to be data scientists have to know how it was gathered. One could claim that under this view, data has to be a tuple of the data itself in the narrow sense and the information how it was gathered, *i.e.* a description of the observation or measurement routine. But I reject such a definition, because when scientists refer to data, they do not refer to such a tuple, but to the data itself in the narrow sense: if a gravitational physicist asks another one for the data, he does not receive a series of numbers with a comprehensive survey about the interferometer with all its technical details and the body of relevant theory; he would receive the later as an answer for the question what the data *means* or what it *shows*.

Mathematical objects are afflicted with metaphysical problems themselves, which are widely discussed in the philosophy of mathematics.[2] Are mathematical

[1] I give some remarks on the recent discussion on the notion of representation in science in 2.3.

[2] Shapiro (2007) offers a survey of the field that serves well as an introduction.

objects real or should we take a nominalistic approach? Can mathematical objects be described in isolation, or should mathematics be described using the structuralist's approach? For an explication of data the ontological classification of data as mathematical objects is sufficiently meaningful, since a thorough metaphysical discussion of these objects is available within the philosophy of mathematics. We can clearly distinguish a mathematical object from a non-mathematical object, even if the ontological classification of a mathematical object is a subject of discussion in the philosophy of mathematics. The assumption that a set of data is a mathematical object prepares the ground to provide a more concrete explication of patterns in data, as will be given in 4.2. Since, as will be argued in 4.2 about the explication of patterns, there is no way around explicating patterns as mathematical. That is why data has to be mathematical, as well.

Under my suggested view, problems for an explication of data arise in scientific contexts where data is usually *analysed* by not only purely algorithmic methods, but rather genuine human "creativity", whatever that may be. In chapter 4, we discuss non-algorithmic pattern recognition procedures conducted by human agents. Data with respect to such an analysis include images, texts and sound tracks. Examples of images as data are photos of the night sky in astronomy, X-ray and ultrasonic photographs in biology and medicine, bubble chamber photographs in physics and patients' drawings in clinical psychology. Examples of texts as data include interviews in psychology and social sciences, poetry and prose in literary studies, and historical documents in historical sciences. Examples of sound tracks as data include voice recordings in linguistics, music for musicology and animal noises in biology.

Translations of data into mathematical objects are not unique; recall the several digital encoding formats for images or sound tracks with computers.[1] Analogous to the case of physically stored data, scientists do not refer to a specific mathematical representation of a set of data when they refer to data. As already mentioned above, the relation of identity or representation between two mathematical objects is not a trivial subject of matter in the first place.

A typical example from the field of quantitative finance are historical asset returns that are used to model the statistical behaviour of certain financial asset prices in financial markets. Imagine a simple example of prices of a liquidly traded stock. For reasons of applicability we are interested in one price listings per day for the last ten years and choose the end-of-business-day price quotes delivered by a certain quote provider. This information is commonly already given as a time series, but not in a unique mathematical format. The data can be given in the form of

[1] For images: BMP, JPEG *etc.*, for sound tracks WAV, MP3, OGG *etc.*

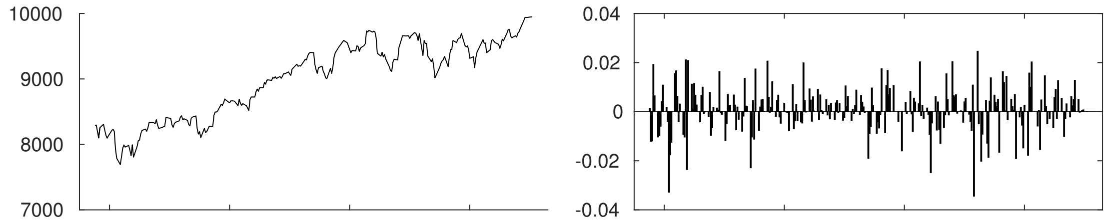


Figure 2.5: Illustration of one year daily DAX performance. Left: time series of index points. Right: time series of simple net returns.[1]

- *prices*, which is a vector of positive real numbers $(S_t)_{t\in\{0,1,\ldots,T\}}$,
- *simple net returns*, which is a vector of real numbers in $(-1,\infty)$ $(R_t)_{t\in\{1,2,\ldots,T\}} = \left(\frac{S_t}{S_{t-1}} - 1\right)_{t\in\{1,2,\ldots,T\}}$, or
- *log-returns*, which is a vector of real numbers in $(-\infty,\infty)$ $(r_t)_{t\in\{1,2,\ldots,T\}} = \left(\ln\left(\frac{S_t}{S_{t-1}}\right)\right)_{t\in\{1,2,\ldots,T\}}$

with some $T \in \mathbb{N}$. 0 denotes the first data point of the time series, which is usually associated with a date as, for instance, June 3$^{\text{st}}$ 2013 and T denotes the last date of the time series as, for instance, June 2$^{\text{nd}}$ 2014. This example is illustrated for the DAX (German top 30 industrial index) by figure 2.5.

This example of asset price or return time series illustrates a very common treatment of data not only in finance, but in science in general. Photographs in cell biology or astrophysics get manipulated by altering contrasts or brightness, signs in classical texts change for new prints after language reforms, and measurements of temperatures can be given in degree Celsius or degree Kelvin. These alterations clearly do not have to be located at the *model of data*-level in Suppes' (1962) approach. According to his terminology they should rather be located on the experimental level.

If they do not refer to physical objects, nor to specific mathematical representations, then to what do scientists refer when they talk about data? If some scientific analysis of financial market data is conducted, then the scientist will choose the data in or transform the data to the most convenient format for her specific purposes. Prices show the absolute level of investment gain in one asset. Simple net returns show the relative returns in percentage. Log-returns show more symmetric empirical distributions than simple net returns and are therefore analytically more convenient to model. If the scientist is interested in the volatility, then prices, simple returns or log-returns are suitable mathematical representations of the data. If the scientist needs to know whether the asset price broke through a certain threshold, then simple returns or log-returns are not suitable mathematical representations of the data.

[1]To produce this plot I adapted a script that I wrote for my diploma thesis (2012).

However, when scientists talk about data, they rather refer to the abstract (mathematical) object that is represented by all of its mathematical translations, which's use in science depends on the scientist's specific interests of application. More specifically, the data is the class of all mathematical representations of it and therefore also a mathematical object. Next, I give a more thorough view on the notion of representation of data.

2.3 Representations of Data

Section Abstract

Data are abstract, mathematical objects. We epistemically access data via a mathematical or physical representation of it. A mathematical representation of data is an element of the class, which is the data. On the other hand, data is most often treated as the direct (measurement) results from an experiment, which is a specific representation of it; in many simple cases all the information given by the data can be given by one of its representations. But as in the exemplary cases of stock returns or photographs, translations of the data for unique pragmatic reasons are common in scientific practice. Therefore, data are abstract, and more precisely: data are mathematical classes and we epistemically access them via mathematical or physical representations of them.

I distinguish between physically *represented* data and data in the direct form of an (abstract) *structure*, that is: a mathematical object. I will elaborate on this distinction further in this section. In many cases, representations of data can be translated between these two forms and remain useful for scientific reasoning. A historical time series of earth quakes that shows numbers according to the Richter scale is useful written down in an old book as well as given as a digital file in a computer. Here, simple translations from one into the other form can be applied. However, ancient accounts of volcanic eruptions in written texts or drawings are not straightforwardly translatable into numerical records. Aspects that are not in the focus of a translation can be important for future scientific use of this data.[1]

Representation of one object by another, and more specifically: scientific representation, is a topic of discussion recently focused on by, amongst others, van Fraassen (2008) and Suárez (1999; 2003; 2004). The representation of data by mathematical objects, as we want to discuss it, should not be mistaken with the "*mathematical imaginary*", as discussed by van Fraassen (2008, p. 39–49) Under mathematical imaging he understands the representation by selective resemblance

[1]Volcanic eruptions coincide with certain kinds of earthquakes. Anyhow, even if detailed records of the earthquakes are available, then not all information on volcanic eruptions are given.

of genuinely non-mathematical objects or parts of the nature by mathematical *models*[1] of them. I restrict my survey about representation to mathematical or physical objects, which are called 'data' in everyday scientific contexts (*e.g.* a computer file; a piece of paper with ink on it), and state that these actually represent something abstract that is referred to and should more appropriately bear the name 'data'. Note that I do not use the concept of representation to the extent that van Fraassen and Suárez have in mind. My motivation for the introduction of this notion of representation is mainly pragmatic and its denotation 'representation' hints to its similarities to other notions of representation, but this denotation is to some extent chosen due to the absence of better alternatives (at least in English). My philosophical motivation for the introduction of representation in this description is to bridge the ontological gap between physical objects that seemingly play the role of data and the mathematical objects that are the data. This is how my account of data deviates from Hacking's and Leonelli's: data are not material artefacts, data is rather represented by these artefacts.

As van Fraassen points out, the available mathematical representations of nature are—scrutinised in detail—generally inaccurate descriptions for pragmatic reasons; a mathematical model that is formulated to describe an empirical phenomenon or part of the nature is usually formulated on the basis of idealising assumptions to achieve a tangible simplicity of the mathematical model. How is this fact related to representations of data? The physical or mathematical representations of data that are used for scientific inferences are not distorted or idealised in any substantial way in comparison to the abstract data itself and regarding their usefulness for the intended scientific inferences. In fact, these representations, or at least a high proportion of them, *manifest* the data they represent. Even if the data that is shown by an X-ray photograph is something abstract, it cannot have any properties that are not shown by the photograph. There are naturally—that means: by definition—no properties of data that cannot be represented by one of its representations. An illuminating example is a developed (non-digital) photograph: it shows something, but it cannot represent more data than the most informative representation of the data, which is the original negative the photograph was developed from. The abstract object that they refer to as data can inherit only the properties that at least one of their physical or mathematical representations show. Any notion of idealisation for representations of data would therefore be misleading.

In van Fraassen's terminology a physical object, the nature, can be represented by a mathematical object, a scientific model. In my use of the terminology the

[1]Here, 'models' has the meaning that it most often has in science: a *model* exemplifies certain aspects of a phenomenon.

mathematical object, the data, is represented by a mathematical or by a physical object, for instance: computer files or photographs. But more strictly speaking (and more complex), such a representing physical object represents a mathematical object (a structure), which is a representation of the data. Since a set of data can be translated into various mathematical forms (see the example from figure 2.5 at page 49), the data itself is the class of all possible mathematical translations. This class is then, of course, mathematical, too.

A sound is something physical, and not mathematical; human beings or animals obviously do not hear a mathematical object, when they hear something. But the data of a sound, irrelevant of it being stored digitally or on a magnet tape as an analogue signal, is abstract and even mathematical. This distinction is not only a distinction of mere wording; the physical phenomenon of sound waves and the biological (or *mental*[1]) phenomenon of hearing are involved in sensorially perceiving the data. But, as will be shown later on, by stating only this property of data as being mathematical makes it possible to explain patterns, pattern recognition and therefore phenomena detection in data. Aspects of human minds do not influence the ontological classification of data and patterns in them—I substantiate this claim regarding patterns in the discussion about Daniel Dennett's (1991) "real patterns" (see 4.2), which is in some regard an opposing position to my views.

Not only material objects can represent data. A mathematical object can represent another mathematical object, which is some data. I sympathise with category theory and *ante rem* structuralism (see 3.1) regarding an answer to the question what a mathematical object metaphysically is. However, structural equivalence in these terms may make two mathematical objects be identical, even if they are formulated very differently (*e.g.* one in set theory and the other in category theory—see the example of an ordered pair at page 75). There is room for arguments in this direction for cases in which one set of data is isomorphic to another set; those cases can be seen as mutual representations or as identities. Be that as it may, more convincing cases of representation of a set of data by another mathematical object are those in which approximations or data compressions for very pragmatic reasons are involved. If you scan an X-ray image to a computer raster graphic, it loses information due to the restricted amount of pixels. Compression methods for computer files (*e.g.* JPEG) often decrease the amount of information they

[1]Taking the specific field of the philosophy of mind into consideration, we should at this point in the discussion not fully omit the fact that a sensation may be a different kind of phenomenon than other (complex) biological phenomena, such as digestion. But the main point that I want to raise here is the general ontological difference between phenomena and data in which patterns that correspond to the phenomena may be detected. It is a further aim to fully describe all the metaphysical varieties and aspects of phenomena. As it turns out (see chapter 5), phenomena selection strongly depends on human sensory and cognitive capabilities.

provide. In such cases the original data is represented by another mathematical object without isomorphism, but by a mapping that is surjective and not injective.

The fact that data are abstract objects, which have representations does—of course—not include that data are some kind of a *perfect ideal* object that carries all the information which can be inferred from the experiment. A repeated performance of a certain experiment under different circumstances certainly produces different data. A phenomenon, as a certain pattern, may not be detectable in the data from the first measurement, but in the data from the second one (*e.g.* signs of weak neutral currents in bubble chamber photographs). But if a pattern can be detected in data, then this pattern should not depend on the quite arbitrary representation of the data it is detected in. The scientist's choice of a specific representation is driven by merely pragmatic considerations. Similar to data being a class of all of its mathematical representations, a pattern is a class of all *concrete* patterns that can be detected in suitable representations of the data. Concrete patterns are mathematical objects as well and will be discussed in 4.1.

This theoretical framework may look overcomplicated at the first sight. But, as we will see, this detailed construction is necessary to explain the relation between patterns and phenomena. We need this clear picture about data and representations of it to explain what patterns in data are. Patterns are something that can or cannot be recognised in a specific set of data. The representations of data are used to apply the pattern recognition procedures (mathematically explicated or not) to them. Different representations of data may be chosen by a scientist due to pragmatic reasons, which are, in parts, guided by the available pattern detection routines.

2.4 Data and the Information about its Origin

Section Abstract

Information about the part of the empirical world under investigation, about the experiments or the measurement routines is not part of the data, which is purely mathematical. The reasons for this are, firstly, how scientists use the term 'data' and, secondly, Hacking's criterion of externality of data that also demands investigator-independence of data. Different claims about scientific phenomena from different parts of the empirical world can be made, defended or argued against on basis of the same data, since scientific *inferences* from data can imply investigator-dependent background knowledge.

Data are mathematical objects in the ontological sense. For scientific fields, in which mathematical vocabulary is often used to express theories, propositions and sets of data, such as in some parts of physics, this approach seems natural.

For other fields in science I aimed to consolidate this approach earlier in this chapter. Disregarding the mentioned problems with an explication of data as purely mathematical (*i.e.* data and pattern recognition are in many cases in science done without the explicit application of defined mathematical algorithms, but by human agents) I believe there is a bigger threat to my position and I want to discuss it in the following paragraphs. In a nutshell, it is this: is a set of data sufficiently described, if we identify it with the mathematical measurement result and neglect the experimental setup and further information about the empirical part of the world under investigation? Or do we need to incorporate these descriptions into our account of data?

I describe this problem more thoroughly. If we restrict a set of data to be the mathematical object to that scientists usually refer to with the term 'data', then all information about the empirical object or system it describes properties of are completely excluded. Think of an X-ray image showing a very specific part of a human lung without us knowing that it shows this part of a lung. Therefore—being rigorous—it is theoretically irrelevant to the data itself from which part of the empirical world and how it was gathered. At another part of this text, I introduce the example of long-range dependence, a statistical property that describes a pattern that can be found in a variety of very different sets of data. (see pages 115 ff.) These sets of data may be gathered from some meteorological measurements, or from internet traffic data or from other very different parts of the empirical world. The long-range dependence that is found in sets of data with different empirical origin may be exactly the same pattern. More specifically, the parameters describing the long-range dependence in the model and the magnitude of approximation of the pattern of long-range dependence in the data may be identical in both cases.

The question whether these two sets of data, which were gathered from very different empirical parts of the world are identical, modulo different noise, or not is related to the question whether the same phenomenon is detected in these two sets of data. Statisticians would agree that they can detect the *statistical* phenomenon of long-range dependence in both sets of data. As elaborated on in 4.2, the example of long-range dependence has the peculiarity that its *general* pattern and the *concrete* pattern (see 4.1 for the explications), which is only one unique concrete patter in this case, are identical, since there is no further description of the phenomenon apart from the statistical definition. Recall that for (Bogen and) and Woodward a phenomenon "is" a pattern in data. Therefore, long-range dependence is an extreme case of a phenomenon that occurs in various empirical parts (or systems) of the world, but may still be considered to be one and the same phenomenon in every case of its occurrence due to its sole statistical nature.

By their structuralist's account in the philosophy of science Balzer *et al.* (1987)

and Sneed (1971) aim for an explication of scientific theories, but a similar conceptual problem to the one that I introduced about data and the information about it occurs to them. It does not suffice to describe an empirical theory solely as a set of formalised mathematical propositions. They introduce "*intended applications*" to restrain the empirical realm of application for the theory by informal descriptions of the intended part of the empirical world (*e.g.* motion of matter on certain scale for Newtonian mechanic; electromagnetic phenomena for Maxwell's equations). The question whether theories can be described directly in purely mathematical terms is therefore denied by the structuralists.

Is there any common ground between the ontological status of scientific theories and the one of scientific data that could lead to a related conclusion regarding the ontological classification of data as mathematical objects? Theories and data have in common that they can both be seen as propositions in a language that provides in some way descriptions of parts of the empirical world. Theories are formulated to explain and predict claims about a part of the empirical world, or more precisely: about its phenomena. Data is gathered to detect patterns in it. The relation between a pattern and a phenomenon is a surjective assignment (*i.e.* every phenomenon can be detected by at least one pattern, but there can be more than one pattern that correspond to a specific phenomenon). But to really define this assignment, that is to describe which pattern exactly has to be detected to validate the occurrence of a certain phenomenon or not, the empirical background information about the set of data is necessary. To detect, for instance, albinism in a human being we need to know that the data we have at hand is the resulting record of a gene analysis (more on this example at 4.2). Theories and data have in common that they can play only their role in processes of scientific inference, if appropriate empirical information is included at this inference. It is important to note that this information plays a role in the process of inference, which is something different than the data.

A simple example can lead us to a decision about the open question how the additional information *about* the data is related to the data. Let us assume one and the same pattern can be detected in two sets of data that were gathered from very different parts of the empirical world. The one set of data describes the amount of bacteria in an artificial colony in a nutrient solution over five days. The other set of data describes the average prices for a litre of beer in Munich's bars in Euro over the last 50 years[1]. In both sets of data the pattern of exponential growth can clearly be detected. The phenomenon of a constant reproduction rate per individual

[1]Prices in Deutsche Mark until 2001 are simply converted by the foreign exchange rate as fixed in 1999, which is the common conversion rule for price statistics including prices in Deutsche Mark and in Euro.

bacterium is assigned to this pattern in the data set that shows the growth of the bacterial colony, but is obviously not a reasonable phenomenon of beer prices in Munich. Here, the pattern of exponential growth may be assigned to a phenomenon concerning conditions of wealth and rental prices in Munich, inflation rates, the inhabitant's drinking habits and so on. Various other examples can easily be found,[1] including the already mentioned case of long-range dependence.

If only the purely mathematical set of data would guide the inference from data to the phenomenon, then it could never be decided whether it validates the occurrence of a certain *empirical* phenomenon or not. It may even be possible to construct artificial examples in which one and the same part of the empirical world is investigated by two very different experiments. Patterns may be detected in the two resulting sets of data, but one and the same pattern in one set of data may validate the occurrence of an empirical phenomenon. But it does not so in the other set of data, due to the different experimental setup and measurement routines. That is why the additional empirical information that has to be added to the purely mathematical set of data does not only have to give information about the empirical part of the world under investigation, but also a detailed account of experimental setup and the applied measurement routines. That is why Suppes' (1962) proposes a conceptually extensive five level hierarchy of models in his account (see 2.1).

In my view, the solution to the riddle is that, even with data being mathematical objects that imply no information about the experiments or measurement routines, the *inference* from data to phenomena is a processes that is strongly guided by further background knowledge. For the detection of gravitational waves with LIGO by Abbott *et al.* (2016), the inference from the detected pattern to the phenomenon could be drawn only with regard to a vast amount of background knowledge about LIGO (*e.g.* the exact position of the lasers and receivers; the exact locations on the earth's surface).

Why is this additional background information not a part of the data itself? There are two main reasons for this. Firstly, when scientist refer to data, they refer to the highly standardised and limited record of some experiment and measurement and not to any description of these experiments and measurements themselves. Secondly, if we incorporate this background information into the concept of data, then the data itself becomes something very investigator-dependent and even ambiguous. Experiments in many scientific fields are often highly collaborative and no involved agent has the epistemic capabilities to cognitively process all

[1]The standard tools of statistical data analysis include many very simple patterns, which can be detected in time series or signals from various sources. See Brockwell and Davis (1991) for an introductory survey. This criterion of a wide range of applicability is rather a criterion that the most basic statistical methods *have to* fulfil to be regarded as most basic.

the relevant information about the measurement and experiment (*e.g.* a detailed description of the entire LIGO architecture). Additionally, there is no report or an in any form organised way of a sufficient description of all the involved background information. Therefore, there would not even be a practical way to really *externalise* the data with all the background information. But the externalisation is substantial criteria of data in Hacking's (1992) approach with which I agree in this regard. In the case of the LIGO example, we *trust* the experts that the measurement works in the way that was only briefly outlined to us. There is no trust in data, but only in scientific agents, who make scientific inferences that involve not explicated background assumptions.

In other words, in many processes of scientific inference a vastly complex set of more or less idiosyncratic background assumptions are stipulated and necessary. This set of background assumptions provides a more or less idiosyncratic *semantics* (in a more mundane sense of the term, without having much model-theoretic implications) to the data and to the pattern recognition routines. And the information about the data's origin is part of this set of background assumptions, which are necessary to make inferences from data to theory or phenomenon.

If the data is mathematical and all the information about the part of the empirical world under investigation, about the experiments and measurement routines is not part of the data, then what are these ontologically speaking? I believe that these assumptions are mathematically explicable, too, but in a way that is much more complex and may differ from human agent to human agent, who are the *believers*. I further elaborate on this non-trivial matter, which is fundamental to the validity of many of my claims in this thesis, in chapter 3. However, Russell (1927) provides an account that expresses exactly this basic epistemological claim that I want to revitalise:

> [W]herever we infer from perceptions, it is only structure that we can validly infer; and structure is what can be expressed by mathematical logic (...) The only legitimate attitude about the physical world seems to be one of complete agnosticism as regards all but its mathematical properties. (p. 254 and 270–1)

A more specific elaboration on the perception of colours and sounds:

> Colours and sounds can be arranged in an order with respect to several characteristics; we have a right to assume that their stimuli can be arranged in an order with respect to corresponding characteristics, but this, by itself, determines only certain logical properties of the stimuli. This applies to all varieties of percepts, and accounts for the fact that our knowledge of physics is mathematical: it is mathematical because no non-mathematical properties of the physical world can be inferred from perception. (p. 253)

The concept of time is derived as a "logical or mathematical property", too:

> [A]ll that we perceive is in the present, and the time-order of the original events is inferred from the relations among the simultaneous events which constitute our present recollection. Thus the conclusion seems to be: Psychological time may be identified with physical time, because because neither is a datum, but each is derived from data by inferences of the sort we have found elsewhere, namely, inferences which allow us to know only the logical or mathematical properties of what we infer. (p. 254)

This stance of Russell also fits well to my explication of data being mathematical objects (section 2.2).

Why is this distinction between data and inferences that are drawn from the data important? I introduced a justified descriptive taxonomy that we use and that is most adequate for the use of the terminology in actual scientific practice. In later chapters of this thesis much of the argumentation is based on a thoroughly explicated conceptual basis of which the concept of data is an important part. Patterns in data are scrutinised in chapter 4 and their relation to phenomena in chapter 5.

I do not believe that a further formalisation of this stated structure of the description of data is useful for the cause of this thesis. The reason for this is simply the vast variety of mathematical formats in which data in science occur.

2.5 Conclusion

At the beginning of this chapter I introduced the problems of *definition and demarcation*, and of *ontological classification* for which an explication of data has to imply a satisfactory solution. The first problem is solved by the conceptual separation of data as purely mathematical objects from all further background assumptions and knowledge involved, whereas this separation is guided by the actual scientist's use of the term 'data'. The latter problem is addressed by the fact that there is a well established field in metaphysics that deals with the issue what a mathematical proposition or object ontologically is. I prefer an ante rem version structuralism, but my explication of data also works well under alternative views about mathematical propositions or objects.

My account of data is *descriptive* in the sense that it corresponds well with what scientists actually refer to with 'data'. It is *normative* in the sense that it ontologically and epistemologically classifies data strictly to the extent that Hacking's and Leonelli's *material* notion of data is rejected. As will be elaborated on in chapter 4, *patterns* can be described only mathematically, which implies that the data's role in scientific inferences is the one of having or not having a certain mathematical property; this gives further strong support for my claim that there is ontologically nothing more to a set of data than being a mathematical object.

According to a newly and pragmatically introduced notion of *representation*, data can be represented by finite or infinite mathematical translations, or by physical objects, such as photographs. The set of data itself can be characterised by all of its mathematical representations and is therefore identified with the class of all of its mathematical representations.

Chapter 3

Mathematics, Mathematical Agents and Computers

Chapter Abstract

I commit myself to an *ante rem* version of mathematical structuralism and mention category theory as our best current approach to explicate such a view mathematically. Agents of mathematics are not restricted to human beings and include artificial intelligences, aliens and the like. Therefore, mathematics is not restricted to human cognitive capabilities. Since mathematics is expressively very powerful, and some parts of scientific inferences are already explicated in mathematical terms (*e.g.* statistics), mathematics is pragmatically very suitable to explicate scientific inferences and theoretical objects. Computer processes can be explicated mathematically via decompiling. Given the recent success of artificial intelligences, following an *optimistic induction*, we should assume that artificially intelligent agents play a more important role for scientific hypothesising at some future point in time.

Having a sufficient grasp of a philosophically adequate notion of mathematics is necessary to follow (or criticise) a large part of the discussion in this thesis. Given the extensive literature and history of the philosophy of mathematics with the aim to explicate what mathematics is, this question is intricate. Since my arguments rest heavily on the expressive power of mathematics, this chapter is aimed to clarify my notion of mathematics to the degree necessary to understand to what I refer to with 'mathematics'. Computer agents are relevant to this study, because in later stages of it I imply that *good* science does not need to be a science of solely human agents. Overall, in the thesis in general and in this chapter in particular I aim to strengthen the view that mathematics is *the* study of structure. That is why I claim that classes of things or propositions like sets of data, patterns or background assumptions are something mathematical.

A general objection against such a view may be to stress the point that only a very small fraction of these things or propositions are actually mathematically explicated. In today's science, we have mathematical models for, for instance, classical mechanics, dividend payments or the determination of dates for palaeontological fossils with chemical methods. However, the vast majority of what we consider as empirical knowledge is not explicated in any mathematical way. The reason for this is, I believe, very pragmatic: the biological cellular computer, which is our brain, remembers and communicates on different levels of abstractness. And the level we are explicitly *aware* of and can use *intentionally* is one after the application of a massive amount unintentional preprocessing. An example for a rough comparison with worked-out mathematical theory is to compress matrices to its determinants.

It is important to note that mathematics does not comprise only the body of already defined mathematical objects and proven theorems from our libraries. It also comprises those objects and theorems that *can be* defined in mathematical terms. Mathematics does not comprise objects that cannot be defined in mathematical terms in principle or theorems that cannot be proven in mathematical terms. Gödel (1931)[1] showed with his incompleteness theorems that some theorems in some weak logical calculi cannot be proven to be false or true.

To say it more mundanely, mathematics is not only about numbers, and not about an academic branch. It is rather an ontological and epistemological concept. This approach is completely consistent with the results about foundations of mathematics at the *Grundlagenstreit* and later developments. These results include, among others, Cantor's introduction of set theory as the vocabulary of mathematics, Hilbert's efforts the establish the need of a characterisation of mathematics by axioms and Bourbaki's comprehensive work to show that the actual body of known mathematics can be formulated on such foundations. The most important later development for the aims of this thesis is the composition of category theory. I elaborate on my view about how to explicate and justify the notion of mathematics from category theory in section 3.1.

Why do I focus so much on mathematics instead of *logic* for the explication of concepts like data or assumptions? The vast majority of philosophical explicative formal work in epistemology and ontology makes use of propositional or predicate logic. Further candidates are modal logic, probabilistic logic (*i.e.* probability theory over logical propositions) and fuzzy logic. The reason for their wide use in philosophy seems to be that these logics are the *sparsest* formal approach to tackle philosophical problems. The development and foundation of mathematics, however, is to a substantial amount driven by problems of application like geometry or number theory. Why then do I want to focus on mathematics? The reason

[1] Incompleteness has two versions and depends on the axiomatic formal system under investigation. See Raatikainen's (2015) encyclopaedia entry for a survey.

is that my epistemological aim is to cover actual scientific inferences in general and not only a conveniently facilitated version of it that ignores the complexity of actual data, scientific data processing and other scientific inferences. I elaborate further on these substantial differences between these two views on scientific inference and data in another article (2018). The central point is that scientific data analysis often makes use of mathematical methods, and newer developments regarding artificial intelligences exemplify how human-like inferences are made by a machine that offers us the complex mathematical structure of this inference via decompiling. I elaborate on the close relation between computability and mathematical explicability in section 3.2 and on the regarding implications of artificial intelligence in section 3.3.

3.1 What a Mathematical Object is

Section Abstract

Throughout the thesis I pragmatically use the term ‘mathematical object’, but commit myself to an *ante rem* version of mathematical structuralism. Mathematical objects can be identically imagined by different agents, they can be fully described with relatively few signs, they can be epistemically fully accessed, they are composed of less complex mathematical objects, and they allow for infinities. Category theory is our most adequate approach to explicate structures and therefore mathematical objects. Following classical notions from Russell and Frege, abstract objects are mathematical objects. Since my notion of epistemology is not restricted to human agents, I do not have to distinguish between epistemology and ontology regarding mathematical objects.

A central point regarding our explanation of the relation between patterns and phenomena is the ontological and epistemological classification of patterns and of data as mathematical objects. Despite all controversies about the specific ontic nature of mathematics in the philosophy of mathematics, the reference of the term ‘mathematical object’, that is, what counts as a mathematical object and what does not, is rather unproblematic. There is prima facie little doubt about the fact that a four or a Hilbert space are mathematical, but a tiger is not (except for the most extreme version of ontic structural realism as discussed in section 4.3).

However, for the sake of a complete explanation of the philosophical ideas in this thesis, I introduce my use of the term ‘mathematical object’. At first, I clarify the very pragmatic motivation for the use of the terminology, which is the one of a *realist*, such as Frege (1884; 1918–1919) according to most historical interpretations[1], Gödel (1944), Maddy (1990) or Putnam (1971), who are also

[1]Frege (1918–1919, p. 69) presents an unambiguous account of his interpretation of *thoughts*

called 'Platonists' (named after Plato's ontological concept of a realm of *ideal* objects) about mathematical objects. The realists' doctrine about mathematics is, as Dummett (1991) puts it, "that mathematical theories relate to systems of abstract objects, existing independently of us, and that the statements of those theories are determinately true or false independently of our knowledge." (p. 301) The main alternative and rivalling approach to ontologically describe the references of mathematical terms is the complete denial of referenced (abstract) entities and to explain mathematics as a mere system of syntactic rules; this approach is the one of a *nominalist*, such as Hilbert (1926)[1], Goodman and Quine (1947), Field (1980), Hellman (1989) and Azzouni (2004). At second, I discuss properties and historical controversies concerning mathematical objects.

When using the term 'mathematical object', as done throughout the entire text of the thesis, I seem philosophically committed to a realism about mathematical objects. A nominalist cannot talk about *mathematical objects*, but rather about *mathematical terms*. However, I choose this terminology mainly for pragmatic reasons. It is applied to distinguish between terms or objects that everyone should be willing to accept as mathematical and those that are not mathematical rather than to really emphasise a strong commitment about the ontology of mathematics.[2]

When talking about *abstract objects* and *mathematical objects* I intend to avoid a more thorough discussion on the ontological nature of what we call a mathematical object. The distinction between abstract and mathematical on the one side, and not abstract and physical on the other side is unambiguous in the cases that are relevant for our purpose; my explications of data, patterns and phenomena do not depend on pathological cases in between these two strictly distinguishable sides of ontological classifications. Therefore, the vocabulary is very pragmatic and the philosophical argumentation is, in principle, open for a translation into nominalistic vocabulary. However, nominalistic vocabulary is much more impractical than the talk of *mathematical objects*. Note that, even if we deem the distinction between mathematical and non-mathematical objects as unproblematic for the cases relevant to us, a possible distinction between mathematical on the one side, and abstract, but non-mathematical on the other side is intricate; I elaborate on this distinction later in this section for further clarification of my use of vocabulary.

as abstract objects by the negative criterion of being non-mental and non-physical, as elaborated on later in this text (p. 79). In his much earlier text (1884, §47–§49) he suggests an interpretation of numbers as *concepts* (Ger., *Begriffe*), which are independent of any physical instantiation (Ger. *Gegenstand als Träger)*. However, discussions on the historical interpretation of Frege as a realist are held by, amongst others, Burge (1992) and Weiner (1995).

[1]Hilbert's position is often titled as *formalism*. However, I believe it is fair to mention him as the historical father of mathematical nominalism in the realism/nominalism dichotomy, in particular due to his role as an opponent to Frege. Blanchette (2014) provides an historical overview over the Frege-Hilbert debate.

[2]The realists' terminology seems more intuitive and pragmatically applicable to me.

In my terminology, mathematical objects are not restricted to any form of complexity, uncomplexity or primitivity whatsoever. One can argue—but this rarely happens—that more complex mathematical objects are defined by basic atoms of mathematics and are therefore no mathematical objects in a very strict sense of *foundationalism*. However, reference to and quantification over mathematical objects are applied in everyday mathematics in the same general form regardless of the specific philosophical interpretation of the mathematical terms that denote these objects; quantification over relatively complex mathematical objects, such as operators in functional analysis or measures in measure theory is in no fundamental manner different from quantification over very basic mathematical objects, such as natural numbers in number theory.[1] Thus, in my terminology, mathematical objects are simply everything that is or can be denoted by a mathematical term in principle. Note that mathematical objects that were explicated by computers can be vastly more complex (*e.g.* measured by the used numbers of characters to denote the object) than any human-made mathematics.[2]—Note also that *complexity* does not imply *creativity*.

Some controversies from the philosophy of mathematics and the arguments that are used to strengthen or weaken a specific position in it shed some light on the differences between mathematical and non-mathematical objects. Mathematical objects can be characterised by having the properties as listed below. Some of these properties depend on each other in one way or another, but this list is intended to highlight one specific particular aspect of mathematical objects per list item, that show an important defining aspect of mathematical objects. I briefly mention some historical discussions concerning these properties. These discussions make the properties, their philosophical implications and their (partial) absence at non-mathematical objects more comprehensible.

(a) Mathematical objects can be imagined in exactly the same way by different individual agents (including non-human agents). The number four I think of does have exactly the same mathematical properties than the number four you think of. Note how this is not true for non-mathematical objects (*e.g.* a pen; a tiger; your grandmother). Realists state, motivated by this property, that mathematical objects exist in an extra, agent-independent realm and

[1] It is interesting to note that mathematical problems that are considered as being hard to solve by the mathematician's community do not depend on any form of complexity of the mathematical objects that are necessary to formulate these problems. The list of proofs discussed in Aigner and Ziegler (2004) is an exemplary manifest of the idea that mathematics is a discipline of problems that are expressed very simple, but only solvable by employing new perspectives of mathematical thinking.

[2] The discussion on the proof for the four-colour theorem is an example for this. Another example are geometric 3D sceneries in modern computer games that are composed of millions of triangles.

independent of the space and time of the physical realm. Thus, they state, the number four is an object in this realm and we both perceive it. As also endorsed by the following properties of mathematical objects in this list, Platonism is so attractive due to its simplicity and explanatory power that Bernays (1935) alleged that mathematicians are usually—outspoken or silent—Platonists.

(b) Mathematical objects can be fully described in a finite and pragmatic way. Under *pragmatic* I understand, that we are able to communicate about mathematical objects as unambiguously as necessary (which may not be the case for many non-mathematical objects). This fact is independent from ones detailed position in philosophy of mathematics and the specific language that is used.[1]

(c) Mathematical objects are epistemically accessible to human beings without the implication of any lack or approximation of their *referential*[2] properties as necessarily involved in the epistemic treatment of empirical objects (*e.g.* a zoologist knows more about a tiger than I do). We can imagine them and we can even reason about them. The scientific discipline of mathematics is nothing else than defining (or from the perspective of a true Platonist: finding) mathematical objects and studying their mathematical properties. Kant (1787) in his epistemology was strongly motivated to find a solution to the riddle why human beings can reason about mathematical objects and how this reasoning is related to reasoning about empirical objects.[3] Benac-

[1]This point can be illustrated by the historical fact that mathematics (which is formulating theorems and finding proofs) was successfully accomplished by Greeks or Leibniz before or during the invention of a more advanced formal calculus. According to Gödel's realism, mathematical objects are accessed by human beings with mathematics similar to empirical access to physical objects with physics: via descriptions, which can be incomplete or wrong. However, there is no reason why well known mathematical objects cannot be described fully and correctly (which is not the case for non-mathematical objects).

[2]A number theorist may know more about 73 (*e.g.* that it is part of a prime twin) than you, but these are not referential properties, since we know exactly what 73 refers to without this knowledge.

[3]For Kant, in his *transcendental aesthetics* of his *Critique of Pure Reason* (Ger. *Kritik der reinen Vernunft*) mathematical objects can be epistemically accessed via *pure intuition* (Ger. *reine Anschauung*), whereas *Anschauung* is the ability to fully imagine an object in the mind without direct empirical contact to it. In Kant's epistemology, *pure intuition* plays a central role, because, for him, it gives us the ability for non-empiric and non-trivial reasoning, the *synthetic judgements* (Ger. *synthetische Urteile*) *a priori*. For him, mathematical reasoning is of this kind of reasoning. Note that Kant developed his views before Frege's invention of a predicate calculus (in the *Begriffsschrift* 1879) and in the *Grundgesetze der Arithmetik* (1893/1903), and before the general discussion on a mathematics-describing axiomatic formal system, which comprises a large part of the history of the philosophy of mathematics from the ending decades of the 19$^{\text{th}}$ century on. By the successfully formalised mathematical calculus the epistemology of mathematical reasoning received further insights and arguments in favour of a nominalistic account. However, Kant's ideas still shed some light on the specific metaphysical nature of mathematical objects

erraf (1973) formulates an argument against Platonism on the basis of the epistemic accessibility of mathematical objects: if we accept that knowledge can be gained only by causal interaction with the objects we gain knowledge about,[1] then Platonists are forced to explain how we can causally interact with objects from another ontic realm, which is, in this case, the abstract realm of mathematical objects.

(d) Mathematical objects are assembled from more simple mathematical objects, or can at least be interpreted and formulated as being so. Even if it is a relatively late historical discovery for many parts of mathematics, most, if not all, mathematical objects can be and are usually defined solely on the basis of more simple mathematical objects.[2] Many discussions in the history of the philosophy of mathematics (including the debates on the positions of *formalism*, *logicism* and *intuitionism*) make implicit or explicit claims about the appropriate fundament of mathematics. This fundament is characterised by the basic objects and the calculus with which mathematics is build up. Today, sets usually play the role of the fundamental building blocks for mathematical objects for the *working mathematician* (for an exemplary survey see Deiser's (2010) introduction into the "basic notions of scientific mathematics"), but types and categories are alternatives.[3] Structuralists in the philosophy of mathematics, such as Hellman (1989), Parsons (1990), Shapiro (1997) and Resnik (1997),[4] infer from the simplicity of the basic building blocks with their very simple own properties (in case of sets: some set is an element of another set or not) and the missing uniqueness

and reasoning, due to its intricate relation to the empirical world.

[1]The causal theory of knowledge, as initiated by Goldman (1967; 1976), states—in the most simplified version—that agent A's belief in fact f is caused by fact f. For a survey see Shope (1983, ch. 5). The crucial question for Benacerraf is how an object in the mathematical realm can have causal relation to a human being, which is an empirical object.

[2]Axiomatic rules for geometry and the construction of geometric objects out of its geometric parts may be a mathematical concept since Euclid's (1908) antique geometric manifest. However, an axiomatic foundation and the definition of basis objects for other mathematical fields, such as analysis, is a *modern* endeavour usually referred to as *Grundlagenstreit* or *Grundlagenkrise* and chronologically located in the first third of the twentieth century.

[3]For a survey on categories see Mac Lane (1971). Type theory was initiated by Russell (1903, 1908) to deliver an alternative approach to set theory, which he considered to be substantially problematic, due to the paradox he formulated by strictly using set theory. It underwent further elaboration and extension by (amongst others, but most influential) Church (1940) by the introduction of a useful calculus for type theory and Martin-Löf (1970; 1973) by the addition of further types to make the calculus more useful for a foundation of mathematics.

[4]Historically, Benacerraf (1965) is often mentioned as an initiator for mathematical structuralism, but Dedekind (1888) formulates similar arguments against the identification of natural numbers with specific sets and therefore in favour of mathematical structuralism. For a comprehensive list of literature about mathematical structuralism see the references in Reck and Price (2000), and in particular footnote 1 for a remark on the historical development of mathematical structuralism.

in the formulation of a mathematical object, as highlighted by Benacerraf (1965) that mathematical objects are *structures* or positions in a *structure*. Structures are nothing else than descriptions of arrangements for positions without any *own* properties apart from their position in the arrangement. To say it in other words: for structuralists mathematical objects are not really objects, but places or structures in a larger structure. Note that many, in different aspects varying forms of structuralism are worked out in the philosophy of mathematics. For a more comprehensive survey on structuralism in contemporary philosophy of mathematics see Reck and Price (2000).

(e) Mathematical objects can be, and are in fact often, infinite (in various forms)[1]—at least in everyday mathematics.

As Shapiro (1997) highlights, the common use of infinities in mathematics is an argument in favour of Platonism, since without the acceptance of abstract objects, we need an explanation on how we can epistemically deal with (or: how we are motivated to define a calculus including) infinities on the basis of our finite perception of the empirical world. Opponents of abstract objects, such as Hilbert (1926), Goodman and Quine (1947), or Field (1980), argue for a nominalism in mathematics instead of a Platonism and are forced to explain exactly this point, which turns out to be an intricate task.

On the other hand, empirical sciences make extensive use of mathematics, especially in physics, to formulate their theories. Accepting Quine's (1969) naturalistic thesis that our best scientific theories give guidance to the philosophical questions on what we know and what entities exist and applying this approach to mathematics, as Putnam (1971) argues, mathematical objects must exist, since without them it is hard to understand scientific theories at all. Furthermore, as Colyvan (2001) states, since empirical evidence and experimental confirmation is available for successful scientific theories, mathematical objects can even be seen as *indispensable* in a naturalistic ontology.

The *intuitionist* Brouwer (2004) states that mathematics is a mental activity of construction by human beings. Therefore, infinite mathematical objects need to be finitely constructed. This leads to a rejection of *actual* infinity in mathematics, as—seemingly—used by working mathematicians everyday;

[1]The existence and nature of *infinity* or *the infinite* is a classical topic of discussion in metaphysics as Moore (1990) examines by his survey. The most influential mathematical formulation of infinities is Cantor's (1895, 1897) concept of transfinite cardinal numbers. However, mathematicians outside the specialised field of set theory and scientists who apply mathematics for theory formulation, modelling or theoretical inferences do usually not further distinguish between different concepts of mathematical infinities; the most obvious indicator for this pragmatic treatment of mathematical infinities is the common and simple (and even for the pragmatic purposes of applied mathematics sloppy) use of the symbol '∞' to denote it.

in his view, mathematical objects can only be *potentially* infinite, such as a construction of the natural numbers by repeatedly applying the successor operator $+1$. From the intuitionist's point of view, many objects in today's mathematics are unacceptable or their intuitionistic justification needs still to be shown.

There is, in principle, no difference for all of these properties between their relevance for very simple mathematical objects and their relevance for very evolved ones. *Simple mathematical objects* include integers, simple geometric objects, such as triangles and parallelograms, or finite sets of simple mathematical objects. More *evolved mathematical objects* are those objects, which may not be epistemically accessible to individuals without a specific mathematical training, due to the complexity of these objects or the high order infinities that are involved. These mathematical objects include foliations from differential topology, Sobolev spaces H^S, and the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ induced by the Brownian motion (mainly due to the high cardinality of Ω). The distinction between simple and evolved mathematical objects is purely pragmatic without any systematic relevance, which is the point I want to make.

More on Mathematical Objects Being Structures

The use of the terms 'structures' and 'patterns' are manifold in the philosophical literature. Some authors use it interchangable and for some structures and patterns are not the same things, whereas the differences in the meaning are not identical as well. Therefore, we need a clarification of the terminology for the purpose of explaining what it means that mathematical objects are structures. Both terms ,'structures' and 'patterns', refer to something purely mathematical. Briefly summarised, structures are mathematical objects in general, and patterns are properties of data, which are mathematical objects (see chapters 2 and 4).

From the structuralists' description of mathematics follows that today's worked out mathematics is only a part of what we understand as mathematics. The mathematical objects that were subject to mathematical work in the history of it were defined by different motivations. The need for a mathematical framework for applications (*e.g.* functional analysis; probability theory) or a specific intellectual appeal (*e.g.* number theory; algebra) may be reasons why certain structures are defined and investigated by mathematicians, but others are not. When I refer to mathematics or mathematical objects I do not refer only to the mathematics that can be find in the worked out theory of mathematics; I refer to proposition and objects that are mathematical in principle, that is, that can be worked out by mathematics in principle.

Following Reck and Price (2000), I want to introduce structures as discussed in the philosophy of mathematics by citing and commenting some of the main proponents of structuralism. I outline my notion of structures on the basis of characterisations given by Resnik and Parsons (1997), which Reck and Price discuss under "pattern structuralism" (p. 363–71)—which is, of course, already a violation of my suggested terminology.

> The underlying philosophical idea here is that in mathematics the primary subject-matter is not the individual mathematical objects but rather the structures in which they are arranged. The objects of mathematics, that is, the entities which our mathematical constants and quantifiers denote, are themselves *atoms*, structureless points, or positions, in structures. As such they have *no identity or distinguishing features outside a structure*. (p. 201, my emphasis)

Are structures objects? Benacerraf (1965) emphasises that for him positions in structures, which is what mathematical denotations refer to, are not something that we should call 'objects', due to the lack of properties that go beyond their purely structural properties:

> Therefore, numbers are not objects at all, because in giving the properties (that is, necessary and sufficient) of numbers you merely characterize an *abstract structure*—and the distinction lies in the fact that the "elements" of the structure have no properties other than those relating them to other "elements" of the same structure.
>
> (...)
>
> That a system of objects exhibits the structure of the integers implies that the elements of that system have some properties not dependent on structure. It must be possible to individuate those objects independently of the role they play in that structure. But this is precisely what cannot be done with the numbers. (p. 70)

As the second part of this quote indicates, for Benacerraf something qualifies as an object, only if it has more properties to it than its position in (or its being an "element" of) a structure. Benacerrafs consideration highlight our differing use of the vocabulary. As noted, I use the notion of objects in mathematics for merely pragmatic purposes, even if I commit myself to a version of structuralism from the philosophy of mathematics.

A precise definition of a structure is available and well-established in mathematics. Structures are associated as being fundamental mathematical concepts in the branches of *model theory* and *universal algebra*. Model theory is the genuine field in which models that provide the *semantics* for logically expressed axiom system are investigated, universal algebra studies the most general forms of the algebras, which include *groups*, *rings* and *fields*. Apparently, structures in this sense (find the definition below) are very fundamental building blocks in every algebraic or set-theoretic approach to build up a corpus of mathematical objects and theorems. However, I argue that there is an even more general understanding of structures that is rather an ontological notion than a mathematical one.

From a historical perspective, Bourbaki (1939–) is commonly associated as having undertaken "the monumental task of reorganizing mathematics in terms of basic structural components" (Thom[1] 1971, p. 699), which includes a concept of a structure that is intended to play the mentioned fundamental role for mathematics in general. 'Bourbaki' is a pseudonym for a working group of mainly French mathematicians that was founded in the 1930s and became very influential among mathematicians with their program to define and discuss a basis for mathematics in general. The reason for this influence in the mathematicians' community may not only be the claim to offer a fundamental, comprehensive and structural account for mathematics, but also the presentation of their result in a pragmatic and well adaptable notation in a well-organised overall notational framework.

Sets are the most basic building blocks of mathematics in Bourbaki's account, which is not surprising given the historical state of the discussion on a foundation for mathematics at this time.[2] They introduced a definition of a structure (1968, ch. IV), which is based on sets and not the one that is widely used in today's mathematics; but one of Bourbaki's special forms of a structure, the *algebraic structure*, resembles the established definition that we mention below. As Corry (1992) claims, the formal notion of a mathematical structure in Bourbaki's sense as well as in the modern definition cannot do the job of explicating mathematical structures in the non-formal sense of the philosophical position of mathematical structuralism. Corry also claims that Bourbaki himself does not make much use of his notion of a structure from the *Theory of Sets* volume, which he uses as an argument to strengthen his point.

I introduce the among mathematicians widely shared modern definition of a mathematical structure from model theory, which is equivalent to the one given in Ebbinghaus' (1994, p. 26) textbook:

Definition (S-structure)

Be $A \neq \emptyset$ a set and S a set of *symbols*. $\mathfrak{A} = (A, \mathfrak{a})$ is a *S-structure*, if

(a) $S = \mathcal{R} \cup \mathcal{F} \cup \mathcal{C}$ with $\mathcal{R}, \mathcal{F}$ and $\mathcal{C}$ being pairwise disjoint, and

[1]René Thom was a topologist and awarded with the Fields Medal in 1958.

[2]The ground for ZFC (***Z**ermelo-**F**raenkel axioms plus axiom of **c**hoice*) as an axiomation for set theory according to which many mathematical objects and reasoning can be formulated was prepared by, mainly, Dedekind (1888), Cantor (1895; 1897), Zermelo (1904; 1908; 1908; 1930), Fraenkel (1922), Skolem (1923) and von Neumann (1923; 1925). See Kanamori (2011) for a historical survey on ZFC, including a detailed overview on how the mentioned references led to ZFC.

(b) $\mathfrak{a}$ is a map with $\mathrm{dom}(\mathfrak{a}) = S$ and

(1) $\forall R \in \mathcal{R} \,(\, \exists n \in \mathbb{N} \,(\, \mathfrak{a}(R) \subset A^n \,)),$

(2) $\forall f \in \mathcal{F} \,(\, \exists n \in \mathbb{N} \,(\, \mathfrak{a}(f) \equiv A^n \to A \,)),$

(3) $\forall c \in \mathcal{C} \,(\, \mathfrak{a}(c) \in A \,).$

The set of symbols S denotes relations $\mathcal{R}$, functions $\mathcal{F}$ and constants $\mathcal{C}$ of the structure, *e.g.* $S_{\mathrm{ar}} := \{<, +, \cdot, 0, 1\}$ for basic arithmetics. Since often a first-order logical language L is given by some contextual background and the symbols S are taken from L, the $S-$*structure* is also called an '$L-$*structure*'.

'A^n' with $n \in \mathbb{N}$ denotes the set of all n-tuples with elements from A. '$A^n \to A$' denotes a function. Bourbaki's structures are defined purely on sets, whereas this common sense definition mixes n-tuples and functions with sets. However, this notation is pragmatic rather than something substantially different from Bourbaki's intentions. Deiser's (2002) introductory mathematical textbook about set theory expresses this view that is widely shared among mathematicians:

> It has been shown that the new framework of axiomatic set theory was large enough to interpret all mathematical objects—numbers of all kind, functions, geometric objects[, relations] etc.—in it. This means there is a definition for all of these concepts that is based on sets and provides all wanted and for mathematics necessary properties of these objects. Set theory therefore is a basic discipline for mathematics and it is unrivalled universal regarding the interpretation of mathematical constructs. (p. 11, translation from German by me)

S-structures are generally successful in covering a lot of mathematical objects, in particular algebras and spaces of various kinds. Algebraic objects like groups, rings and fields are straightforwardly S-structures. With a proper amount of mathematical work objects like vector spaces, measurable spaces, geometric oder topological objects can be expressed by S-structures.

I now highlight why such a definition cannot do the job of explicating structures in the sense of mathematical structuralism. The problematic aspect with Bourbaki's account and also the pragmatic definition of S-structures is that they are defined on sets and therefore a mathematical structure in the structuralists sense or object cannot be *uniquely* identified with an S-structure. The classical example from Benacerraf (1965) suffices to fully explain this point: two alleged structuralists regarding mathematics, who use set constructions as structures, cannot agree on what natural numbers exactly are. Among the infinitely many possible solutions are the following two:

(E) $\{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots$

(J) $\{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, \ldots$

The denotations of the two versions are intended to honour the mathematicians, who proposed these structures, Ernst Zermelo and John von Neumann.[1] Both solutions follow clear construction rules to express succession, which is the fundamental characteristic of the natural numbers structure, by certain constructions of sets. In particular, (E) translates $n+1$ into $n \cup \{n\}$ and (J) translates n into $\{n\}$. But there is no further argument that can be applied to rule out one of the two options to be the structure in question.

The problem of mathematically explicating structures in the sense of mathematical structuralists leads Corry (1992) to distinguish between structures in two different meanings, the "formal one and the nonformal one" (p. 315). Parsons often cited characterisation of mathematical structuralism describes the targets of a worked-out structuralism:

> By the "structuralists view" of mathematical objects, I mean the view that reference to mathematical objects is always in the context of some background structure, and the objects have no more to them than can be expressed in terms of basic relations of the structure. (Parsons 1990, p. 303)

In fact, the philosophical discussions about mathematical structuralism are to a large extent non-formal.[2] And if we can find formal aspects that do not include sets, then it is not clear how the formalism to express a very specific structure should be generalised to express mathematical structures in general.

But, are alternatives available to sets in which we can express the "basic relations" in a unique way? This implies that they express mathematical objects in a way that, strictly speaking, if the expression differs, then the referred mathematical objects always have to be different as well. One important attempt into this direction has to be mentioned at this point and is widely discussed among mathematicians and philosophers of mathematics since the 1940s: *category theory*.

As Marquis (2014) describes in his survey, Eilenberg and Mac Lane (1945) introduced *categories* "in a purely auxiliary fashion" (sect. 1.1). In particular due to the often referred-to work of Lawvere (1963, PhD thesis), category theory is discussed to provide an alternative to set theory as the foundation of mathematics. It is "a general mathematical theory of structures and systems of structures" (Marquis, introduction) and therefore we should have a look whether it can provide a

[1]Benacerraf refers to them via the representative fictional child characters *Ernie* and *Johnny*, who learn basic algorithm as a set theoretic structure in the first place and struggle with the mentioned problem of identification.

[2]See Reck and Price (2000) for an introductory survey that is also exemplary for the use of formal methods in the discussion. Further examples are the often referred to books about the topic from Hellman (1989), Shapiro (1997) and Resnik (1997) and the article from Parsons (1990).

notion of structures that is general enough to capture structures in the sense of mathematical structuralism.

Most of the following discussion revolves around the question whether categories can provide a foundation for mathematics that is not afflicted with Benacerraf's identification problem from above and is at least as general as set theory is. This means that all the mathematical objects for which we have set constructions, constructions based on categories must be possible as well. However, I aim to explicate what structures are and not what the foundation of mathematics is. The discussion of this issue helps to clarify the ontological and epistemological implications of my explication of data, patterns and phenomena. Structures in the most general meaning of the term are something purely mathematical and every mathematical object is a structure.

I introduce a very common definition of categories that is used by, among others, Lambek (1968; 1969; 1972), Mac Lane (1971) and Awodey (1996). I follow their functor style notation.

Definition (category)

A *category* consists of *c-objects* $A, B, C, ...$ and *morphisms* (or *arrows*) $f, g, h, ...$ with

(i) Every morphism f has the form $f : A \to B$, *i.e.* it has a *domain* (here A) and a *codomain* (here B)

(ii) The composition of morphism is unique and associative, *i.e.*

$$h \circ (g \circ f) = (h \circ g) \circ f$$

(iii) For each object B there is a morphism $1_B : B \to B$ with

$$1_B \circ f = f \text{ and } g \circ 1_B = g \text{ for } f : A \to B \text{ and } g : B \to C$$

Morphisms and c-objects are *primitive notions* at this point, meaning that have have no further properties than these defined ones. Throughout the literature, c-objects are usually denoted as 'objects', but I want to pragmatically establish a distinction in wording here to be always sure whether we refer to c-objects from a category or to a mathematical object in the very general sense; however, a strict proponent of category theory as a way to express mathematical objects would reject that there can be a distinction. In the application of categories, c-objects are often restricted to well known mathematical objects, such as sets or topological spaces. But this aspects of "categories for the working mathematician" (this is the title of Mac Lane's 1971 text book) is not in the main scope of my survey.

The very simple, and often mentioned[1], example of a Cartesian product $X \times Y$ of two c-objects shows how categories help to avoid Benacerraf's identification problem.

$$\begin{tikzcd} & W \arrow[dl, "f"'] \arrow[d, "h"] \arrow[dr, "g"] & \\ X & X \times Y \arrow[l, "p"] \arrow[r, "q"'] & Y \end{tikzcd}$$

This plotted graph shows the categorical specification of a product, which is, in other words, an ordered pair of two c-objects in a category. In set language, a product of sets $X_S \times Y_S$ can be expressed in infinitely many ways. The Cartesian product $\{(x_s, y_s) : x_s \in X_S,\ y_S \in Y_S\}$ can, for instance, be $\{\{\{x_s\}, \{x_s, y_s\}\} : x_s \in X_S,\ y_s \in Y_S\}$ (as proposed by Kuratowski 1921) or $\{\{\{\{x_s\}, \emptyset\}, \{\{y_s\}\}\} : x_s \in X_S,\ y_S \in Y_S\}$ (as proposed by Wiener 1914) or one of infinitely many other solutions. In the illustrated category, to define the product $X \times Y$, for any given W and morphisms f and g the morphism h has to be unique such that $g = h \circ q$ and $f = h \circ p$. W and h assure that the product is an ordered pair. X and p, as well as Y and q assure that $X \times Y$ is in fact a product of two c-objects X and Y.

Along this simple example we can point at the advantages and disadvantages of category theory. This category of a product is a very pragmatic way to specify what in set language mathematics is usually described as *identity up to isomorphism*. It is not prima facie given what the class of isomorphism of (E) (p. 72) are, but the category of the product $\times$ provides a structural description of what all the isomorphic sets that represent ordered pairs have to express. This is the advantage that categories have over sets.

The disadvantage is that in most mathematical elaborations on categories the c-objects are assumed to come from a certain class of mathematical objects like sets or topologies. But we are interested in a specification of structures in general and not in one under the assumption that the basic objects of our epistemic conceptualisation are sets or topologies. Lawvere's *categories of categories* are a specific approach to overcome this problem. He introduces them as categories "whose maps [*i.e.* morphisms] are 'all' possible functors [which are morphisms between categories], and whose [c-]objects are 'all' possible (identity functors of) categories" (1963, p. 26). This is a way to naturally define categories as structural in the sense of category theory all the way down.

In the following, I discuss two further major concerns regarding the acceptance of categories from category theory as a general notion for structures. I claim that structures are mathematical, and I want to find out whether categories can do the

[1]It is mentioned by Lawvere (1963), Mac Lane (1971) and Awodey (1996).

job of describing the mathematical objects that are structures in this most general sense.

The following quote from Marquis (2014) introduces the first problem:

> There is no such thing, for instance, as the natural numbers. However, it can be argued that there is such a thing as the *concept* of natural numbers. Indeed, the concept of natural numbers can be given unambiguously, via the Dedekind-Peano-Lawvere axioms, but what this concept refers to in specific cases depends on the context in which it is interpreted, *e.g.*, the category of sets or a topos of sheaves over a topological space. It is hard to resist the temptation to think that category theory embodies a form of structuralism that it describes mathematical objects as structures since the latter, presumably, are always characterized up to isomorphism. Thus, the key here has to do with the kind of criterion of identity at work within a categorical framework and how it resembles any criterion given for objects which are thought of as forms in general. One of the standard objections presented against this view is that, if objects are thought of as structures and only as *abstract* structures, meaning here that they are separated from any specific or concrete representation, then it is impossible to locate them within the mathematical universe. (See Hellman (2003) for a standard formulation of the objection, McLarty (1993), Awodey (2004), Landry & Marquis (2005), Shapiro (2005), Landry (2011), Linnebo & Pettigrew (2011), McLarty (2011) for relevant material on the issue.)

It seems he thinks of mundane mathematical objects, like the natural numbers or triangles, which have to be "locate[d] (...) within the mathematical universe" and the "*concept* of natural numbers" cannot be thought of as an "*abstract* structure" alone without any "specific or concrete representation" *e.g.* "the category of sets". This position is one of an *in re* structural realist about mathematics, which I do not share.[1] It implies that structures, which mathematical objects are, have to, as Shapiro (1997) puts it, be "exemplified in a nonstructural realm" (p. 89). Marquis further describes this point:

> A slightly different way to make sense of the situation is to think of mathematical objects as *types* for which there are tokens given in different contexts. This is strikingly different from the situation one finds in set theory, in which mathematical objects are defined uniquely and their reference is given directly. Although one can make room for types within set theory via equivalence classes or isomorphism types in general, the *basic* criterion of identity within that framework is given by the axiom of extensionality and thus, ultimately, reference is made to specific sets. Furthermore, it can be argued that the relation between a type and its token is not represented adequately by the membership relation. A token does not belong to a type, it is not an element of a type, but rather it is an instance of it. In a categorical framework, one always refers to a *token* of a type, and what the theory

[1] See Reck and Price (2000, p. 366 ff.) with reference to Shapiro (1997, ch. 3) for a discussion on *in re* structural realism about mathematics and the distinction between *in re* - and *ante rem* structuralism. Shapiro is an *ante rem* structuralist.

> characterizes directly is the type, not the tokens. In this framework, one does not have to locate a type, but tokens of it are, at least in mathematics, epistemologically required. This is simply the reflection of the interaction between the abstract and the concrete in the epistemological sense (and not the ontological sense of these latter expressions.) (See Ellerman (1988), Marquis (2000), Marquis (2006), Marquis (2013).)

I believe that here he has no substantial point of criticism to offer. Cantor famously introduces a set as "a gathering together into a whole of definite, distinct objects of our perception [Ger. *Anschauung*] or of our thought—which are called elements of the set." (1895; 1897) This indicates that the elements of a set do epistemically refer to mathematical or non-mathematical objects. However, this is not the way mathematicians use sets in der everyday use and in particular not in the discussion about a mathematical foundation. The cases (J) and (E) (p. 72) show how a primitive element $\emptyset$ that refers to nothing is used with the basic relation $\in$ to construct the structure of the natural numbers without any further referencing elements at all.

In abstract algebra, elements $a, b \in A$ from the set A of a S-structure $(A, \mathfrak{a})$, for instance a group, do not refer to anything different than what the S-structure expresses. The elements of A are simply variables for something pairwise non-identical, but they are still variables, *i.e.* placeholders for something arbitrary; and this is how the mathematicians of the field epistemically treat these elements of A. Even in probability theory the set Ω of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ has elements $\omega_1, \omega_2, ...$ that do not have any specified reference. Probability theorists usually construct probability distribution functions and stochastic processes without even specifying to what scenario a certain element refers to. The elements, while *doing* mathematics, are seen as variables for arbitrary mathematical or, in contexts of modelling applications, non-mathematical reference objects. Seen from this perspective, mathematicians are true structuralists, even when they work in a set-theoretic framework. Another very common and simple examples of missing references is the imaginary constant $\mathtt{i}$, which is the solution of $\mathtt{i}^2 = -1$. It makes no sense to speak of any "concrete" $\mathtt{i}$. Your grandmother will not be able to offer you a $\sqrt{-1}$ share of birthday pie, even if she has a PhD in math. That is why the only remaining criterion of any *instantiation* of a mathematical object could be an agent, who thought about it or wrote it down. The ontological status of mathematical objects then boils down to the classical Heideggerian question whether mind processes should influence our ontology or not. I reject this view regarding mathematical objects[1] and advertise an *ante rem* structuralism about

[1]The arguments are, briefly, that (i) the vast amount of possible agents (*e.g.* artificial intelligences; aliens from Zeta Reticuli) make it likely that every mathematical object is already instantiated somewhere, and (ii) since mathematical objects are truely intersubjective in the

mathematical objects.

I expressed a view concerning a question that is deeply imbedded in the history of philosophy of mathematics. It was an aspect of the early 20$^{\text{th}}$ century *Grundlagenstreit* that is usually historically connected to Hilbert (mathematics is a game of signs and axioms) and Frege (mathematical terms refer to objects of thought).

However, even if we do not believe that Marquis has a substantial point of criticism to offer, the second concern is more threatening. Awodey (2004), one of the most convinced proponents of the view that category theory helps us to understand what mathematical structures are, is very reluctant when it comes to nominate category theory as the approach to define mathematical structures in some *global* sense. Under a global account of mathematical structures I understand an approach that is structural, *i.e.* it is purely based on relations without any further referenced objects, and *foundational* for mathematics. One approach of this sort would be to formulate the whole body of discovered (or *known* or *defined*) mathematics in Lawvere's categories of categories. Awodey states:

> No one doing category theory thinks we are someday going to find the one 'true topos', in which all mathematics happens. The translations of set theory into topos theory (and other categories) are intended to show that categories like toposes can be used to do a lot of mathematics *for those used to doing mathematics in set theory*; they are not supposed to show that topos theory is the new universal 'system of foundations', intended to replace set theory. (p. 55)

A topos is a sort of category axiomatically defined by Tierney (1972). Marquis (2014) summarises that "an elementary topos is a category possessing a logical structure sufficiently rich to develop most of 'ordinary mathematics', that is, most of what is taught to mathematics undergraduates" (sect. 2). What can category theory offer regarding our interest to explicate what structures are? Awodey goes on:

> As opposed to this one-universe, 'global foundational' view, the 'categorical-structural' one we advocate is based instead on the idea of specifying, for a given theorem or theory only the required or relevant degree of information or structure, the essential features of a given situation, for the purpose at hand, without assuming some ultimate knowledge, specification, or determination of the 'objects' involved. The laws, rules, and axioms involved in a particular piece of reasoning, or a field of mathematics, may vary from one to the next, or even from one mathematician or epoch to another. (p. 56)

sense that a four is the same for you than it is for me, it makes little sense to make the reality of a four for me dependent on your thoughts about it.

This quote fits well to the actual development in category theory. Our simple example of ordered pairs $(x, y) \in X \times Y$ (p. 75) provides the structural interpretation of a Cartesian product, but in this case we have to stipulate that X and Y are sets, which a true category structuralists according to a "'global foundational' view" would have to replace with a categorical account of sets. Such an account is worked out by Lawvere (1964), but this is not the case for the whole body of mathematics. Awodey:

> No one claims that category theory is the only way to talk about structures of structures of ... . Or even that it is the best way (although I know of no better one). The only claim being made in this connection is that it is a very good way. (...) Category theory was developed so extensively because the notion of a category, and the related notions of functoriality, naturality, and adjointness, proved to be so effective in modern, abstract, mathematics. And the reason for this broad applicability has a lot to do precisely with their effectiveness at specifying and manipulating structures. (p. 61)

I agree with Awodey's position. Categories are the best explications of structures that today's mathematics has to offer. They are based on very simple axioms (p. 74) and their graph-like style is epistemically easily accessible. As shown in our discussion above, they are a true step forward in our endeavour of an explication of structures in general in comparison to sets, which are still the common approach for the foundation of mathematics among most communities of mathematicians.

However, I use 'structure' not with a strict reference to categories, but in the usual non-formal sense from philosophy of mathematics. The hereby given insight into the discussion on how globally fundamental category theory can or cannot be for mathematics helps to grasp a better understanding about what a mathematical object, a structure, is (in particular to readers without a strong mathematical background). In my view, actual approaches to mathematically explicate the notion of structure are too often neglected in the philosophical discussion on mathematical structuralism. As Lawvere emphasises, category theory is a step forward regarding an explication of the notion of structures in comparison to set theory.

Are There Abstract and Non-Mathematical Objects?

The introduced notion of abstractness is relatively simple: I call an object abstract, if it has no spatiotemporal location. This criterion is the descendant of Frege's (1918–1919) introduction of the ontological "third realm" where he locates thoughts to be. Objects in Frege's "third realm" are neither materially located in the world, nor do they depend on being the content of someone's conscious imagination.[1] I interpret the criterion of non-spatiotemporality to fulfil Frege's

[1] Frege writes:

criteria of being in the "outside world" and the independence from being a "content of consciousness". There are opponents of such an interpretation putting forth legitimate doubts, but I do not aim to discuss these problems in this text.[1]

I make use of the concept of abstract objects, as introduced above, to formulate a sharp counterpart to empirical objects. This twofold universe is of specific interest to us, due to the clear ontological distinction between the scientific phenomena, which are features of the empirical world, and the pattern in data, which are abstract and also mathematical objects (or terms).

Physical objects are not abstract, but every mathematical object is. Furthermore, the only sort of abstract objects I use in the argumentation are mathematical objects. That is why I use the two adjectives 'abstract' and 'mathematical' synonymously in this text. The title of this section is somewhat provoking, since many examples of abstract objects come to mind that are *usually* not considered to be mathematical: propositions, concepts, the game of chess, Goethe's Faust. If the distinction between abstract and concrete objects is interpreted as being identical to the distinction between particulars and universals, such as Quine (1961, esp. ch. I, IV, V, VI) suggests, then the possibility of non-mathematical abstract objects is even more appealing.

However, here is a justification for my use of the terminology. The starting point of the argumentation is my choice of a pragmatic position concerning mathematical objects: they are positions in structures or structures themselves. Furthermore, I stated that a set of scientific data is a mathematical object that

---

> Thus, the result seems to be: Thoughts are neither objects of the outside world, nor of imagination.
>
> A third realm needs to be acclaimed. Objects in this realm have something in common with imaginations: they cannot be perceived with the senses. But they have in common with the objects from the outside world that they do not require a mental bearer of whoms consciousness they are content of. For instance, the thought that is expressed by Pythagoras' theorem is timelessly true and its truth is independent of any individual mind's opinion about its truth. No mental bearer is necessary. The theorem is not only true, because someone discovered it. It is true like it is true that a planet has already interacted with other planets before someone discovered it.

translated (by me) from German:

> *So scheint das Ergebnis zu sein: Die Gedanken sind weder Dinge der Außenwelt noch Vorstellungen.*
>
> *Ein drittes Reich muß anerkannt werden. Was zu diesem gehört, stimmt mit den Vorstellungen darin überein, daß es nicht mit den Sinnen wahrgenommen werden kann, mit den Dingen aber darin, daß es keines Trägers bedarf, zu dessen Bewußtseinsinhalte es gehört. So ist z. B. der Gedanke, den wir im pythagoreischen Lehrsatz aussprachen, zeitlos wahr, unabhängig davon wahr, ob irgendjemand ihn für wahr hält. Er bedarf keines Trägers. Er ist wahr nicht erst, seitdem er entdeckt worden ist, wie ein Planet, schon bevor jemand ihn gesehen hat, mit andern Planeten in Wechselwirkung gewesen ist.*

As Rosen (2012) mentions, similar ideas were already proposed by Brentano (1874).

[1]Rosen (2012) provides an encyclopaedic overview.

can be represented physically or mathematically plus the additional empirical information of its experimental origin. (see chapter 2) The main reason for this assumption is that patterns can be explained only mathematically and scientific data analysis is explained to be the detection of patterns in scientific data. Data are abstract objects not only in the interpretation of abstractness I chose, but also in other common interpretations of it. A set of data fits well into the list of exemplary abstract objects above. If mathematical objects are structures and do have the expressive power to distinguish all possible sets of scientific data, why do they not have the power to express all abstract objects? I want to disentangle this thread of argumentation more thoroughly in the following passages.

First, structures, which are a composition of relations between propertyless primitive objects, have a high expressive power. But are structures possible, which are not mathematical? I believe the description of mathematical objects as structures is an equivalence: every structure is a mathematical object and every mathematical object is a structure. To say it in other words, close to those of Resnik (1997)[1], mathematics is the science of structures.

Second, assuming you agree with my conclusions from chapter 2 about data being mathematical, other abstract objects can easily be regarded as forms of *possible* data. A written down proposition (in a logical or any other language) or a thorough description of a concept differs only—if at all—from data by the irrelevance of the information of its origin, which is highly important for some sets of data. The claim that there is no difference between the class of all abstract objects in general and the class of all mathematical objects is strongly related to Russell's (1927) idea (that I already mentioned at page 57): "wherever we infer from perceptions, it is only structure that we can validly infer; and structure is what can be expressed by mathematical logic" (p. 254). This claim is *epistemic*, but as elaborated on later (p. 82), I believe that epistemic criteria that are not restricted to human agents serve perfectly as indications of what counts as an *ontology* (whatever this is in me detail). If this is the case, then a Kantian-Russellian argument serves to strengthen the idea that the ontological class of abstract objects concurs with the one of mathematical objects. Kant suggests that objects of thought can be accessed only by inner illustration (Ger. "*innere Anschauung*"), Russell claims that all perceptions (*inner* and *outer*) result in mathematical objects or terms. To summarise the argument: abstract objects must be epistemically accessible, what is epistemically accessible needs to be a result of possible perception (Kant), and perceptions result in structures (Russell), which are mathematical objects.

The biggest step in my argumentation is the following. Abstract objects, such as propositions, concepts, the game of chess or a piece of literature may in many

[1]In his structuralism, Resnik uses 'patterns' interchangeable with 'structures'. (cf. p. 202)

cases be describable as structures. And if every structure is a mathematical object, then how can an abstract object be non-mathematical? I want to examine this statement more thoroughly on the basis of the mentioned examples.

Propositions may be reasonably decomposable into a logical structure. If so, then their logical form can easily be identified as a structure in a more general sense. Usually, parts of a proposition may refer to non-logical objects or properties, such as concepts or predicates of empirical objects or empirical objects themselves. The important question at this point is: is the proposition only the logical form or are the referenced objects and propositions parts of the proposition itself? I do not think that this latter option is a good one. The physical sun is not part of ⌜the sun raises every morning⌝.[1] One argument in favour of this position is the reference to *fictional* objects, which do not physically exist. If unicorns have to be a part of ⌜unicors neigh⌝, then this proposition could not exist, which would be very odd.

If We Accept Computers as Agents of Mathematics, then the Ontology regarding Mathematics already Implies all Epistemological Aspects for Mathematics

Epistemology is, very roughly, the study of human perception and cognition, or, in other words, how human beings perceive information, think and judge. *Ontology* is, very roughly, the classification of things in this world on a level of abstraction that is not covered by any (other) scientific field. Given these descriptions, there should—one may think prima facie—not be much debate about whether a specific philosophical claim is an epistemological one or an ontological one: *Being* seems to be a rather different philosophical concept than *perceiving* or *reasoning* are.

However, Heidegger (1927) influentially incorporates *being a conscious human* as a crucial and non-reducible aspects into his ontological study. I discuss how Dennett infers from human perception an extra ontological class of "real patterns" in section 4.3. The idea behind these approaches is that something that can be perceived or thought about by a human agent *is* something distinct and cannot be ontologically *reduced* by any means.

In my approach, data and pattern, and even background assumptions in science are something structural and therefore mathematical. Otherwise, intersubjectivity of scientific evidence and inference would not be possible. How epistemological and how ontological is this claim? If the notion of epistemology is restricted to human epistemology, which it is in most epistemological accounts, then my claim does not have anything to do with epistemology due to the extensive use of tech-

[1] '⌜' and '⌝' denote quasi-quotation.

nical aids in science (*e.g.* statistical pattern analysis performed by computers)[1]. One may now oppose this claim by stressing that, even if a set of data may be a vastly complex mathematical object, propositional background assumptions are still human-made and must be conceivable for human agents. However, as stated earlier in this chapter, for pragmatic and historical (meaning here: future) reasons background assumptions, sets of data and patterns are not restricted to human agents. According to my notion of mathematics, Deep Blue's considerations about its next move are surely something mathematical and qualify as (scientific) reasoning, too, but no human agent may be capable of *understanding* it. Applications of a calculator are common in all fields of quantitative science and are part of scientific reasoning.

If the notion of epistemology is not restricted to human agents and also aims to describe non-human agents, such as calculators, artificial intelligences, extraterrestrials and the like, then, regarding mathematical objects, the notion of epistemology is identical to the notion of ontology. The reason is, as elaborated on in this section above, mathematical objects are explicated as being structures. Without any extra epistemic restrictions (*e.g.* complexity) that force us to exclude specific structures from the class of all mathematical objects, everything that counts as mathematical from an ontological view point does so from the epistemological point of view, too. Therefore, a notion of structure is sufficient to *extensionally* define all mathematics. This extensional *foundation* for mathematics sufficiently covers the *epistemology* of mathematics. Why is that? If we restrict the epistemology of mathematics to human agents,[2] then we had to exclude all the humanly inconceivable mathematical objects that AIs *use* in their reasoning. And this would not be an epistemology of mathematics, it would be an epistemology of restricted mathematics. That is why an epistemology of mathematics has to be an epistemology that applies to possible AIs and aliens, too. But since this class of reasoning machines and beings is arbitrary large and diverse, the only reasonable claim we can make about mathematical epistemology is what we already know from its ontology: mathematical objects are structures. And the foundational approaches for mathematical structuralism, such as category theory, provide all there is epistemically to know about mathematics: what the primitives are, how objects are composed of its components etc.

---

[1] The statistician or programmer may understand aspects of the computer aided analysis, but the scientist himself does often not have to.

[2] One exemplary episode for such a restriction of mathematics to human agents is the discussion about the proof of the four-colour theorem. The proof is no proof according to an epistemology of mathematics that is restricted to human agents. However, it is a proof if we eliminate this restriction and allow non-human mathematical agents. More on the discussion can be found at Tymoczko (1979) and Swart (1980)—both are more concerned with forms of *truth*, whereas I am mainly concerned with the epistemological implications.

3.2 Everything Computable is Mathematical

Section Abstract

Mathematics is explicatively more powerful than a purely logical language under pragmatic considerations and provides therefore a more suitable language to explicate data and patterns in science. The Church-Turing thesis shows that computer tasks are equivalent to a certain class of mathematical functions. It follows that the notion of a mathematical object or term incorporates all possible computer data and processes. However, non-constructive mathematics is not computable, which makes the class of mathematical objects a superset to computable data and processes.

Why do I put so much emphasis on explaining that data and patterns in them are *mathematical*, as opposed to, for instance, explicable in *predicate* logical terms, which arguably has more support in the recent explicative philosophical literature, beginning with Frege's, Carnap's and Hempel's works? As stated throughout this thesis, the reason is that today's scientists use more and more aid from computers to process data, find patterns and even explicate them.—Just try to explicate a 64×64 pixel image in predicate logic. Data sets or patterns are often stored and processed as statistical *signals*, *time series*, encoded files or the like. That is why for actual examples of scientific reasoning, purely logical explications fall pragmatically short, since it does not seem pragmatically feasible to explicate actual data sets (see the examples discussed in chapter 2) and computer aided pattern detection processes in common predicate logic. In this section I defend the claim that the notion of mathematics, as discussed in the former section, is powerful enough to play the role of the language of the demanded explications of data and patterns in science, including all possible involved computer processes.

Briefly outlined, the argumentation in this section is as follows. *Logic* is motivated to explicate the language of (human) reasoning and is therefore the preferred candidate for formal explications by the majority of today's philosophers. *Mathematics* is the language (or provides the objects) for the study of structures, which has more similarities to visual imagination (Kant) and relational thinking than what logics aim for. *Computers* are deterministic machines that are designed to execute computational tasks. These tasks are often, but not always, formulated in a *programming language*, which may have some substantial similarities to a logical language and may make use of a mathematical calculus. Tasks that are results of machine learning applications, are examples for tasks that are not directly formulated in a programming language. However, every task for any computer can be described in the compiled *computer language*, which describes the task for the computer, but is in many cases not easily epistemically accessible to a human agent.

```c
#include <stdio.h>
int main()
{
 int a = 2 + 2;
 printf("%d", a);
 return 0;
}
```

```
¨ELF###          # > #    #@     @
p#             @ 8   @ # # #   #   @        @
@      @ @     ø#      ø#      #       #
#   8#     8#@     8#@     #       #
#       #  #             @       @
¬#      ¬#              #   #   ##
##‘     ##‘     (#      O#              #
#   (#      (#‘     (#‘     Ð#      Ð#
#       # #   T#      T#@     T#@     D
D       #       Påtd#   _#          (...)
```

Figure 3.1: Left: a C script that calculates two plus two and shows the answer on the screen. Right: the first of 8520 bytes of the compiled version for my computer and visualised as characters with UTF-8 interpretation.

Figure 3.1 provides a small example of a program in the programming language and in a computer language. Roughly stated, the computer language determines which sequence of electric currents has to flow to which part of the computer.

For our purposes of discussing computers on a basic level there is no useful distinction between *software* and *hardware*.[1] A *Turing machine* is a specific architecture for a computer. For mainly historical reasons[2], Turing machines are theoretically well understood.

The crucial argument for my claim that it is very reasonable to assume that every computer task can be explicated in mathematical terms is the Church-Turing thesis. Many versions of this thesis were discussed and the two original versions from Alonzo Church and and Alan Turing are sometimes misinterpreted and misrepresented. (cf. Copeland 2002) I present a version that is, in my view, sufficiently historically accurate, understandable without specific theoretical background from the field of computational logic, and strong enough to provide sufficient support for my claim.

Church–Turing thesis

Be $\mathbb{N}$ the set of positive integers. For a function $f : \mathbb{N} \to \mathbb{N}$

$$f \text{ is effectively calculable} \Longrightarrow$$

[1]This sentence should avoid confusion for readers who are very unfamiliar with computer architecture. Software is always physically stored and therefore not distinguishable from an architectural perspective.

[2]Turing machines were first discussed by Turing (1937). I speculate that Turing's theoretical architecture for a computation machine was appealing to logicians and philosophers due to its conceptual simplicity, which helps to theorise about computability without any constraints regarding engineering feasibility. Note that Leibniz (1646–1716) and others already correctly discussed mechanical computing machines in detail (cf. Davis 2000, ch. 1) and Konrad Zuse's patented the mechanical switch for the *Z1* in 1936 (2017).

$$f \text{ is general recursive} \underset{(a)}{\iff} f \text{ is } \lambda\text{-computable} \underset{(b)}{\iff} f \text{ is Turing computable.}$$

Being *effectively calculable* is an informal concept expressing that a function can be calculated by (i) using a finite number of finitely formulated exact instructions, (ii) produce the result in a finite number of steps, (iii) can be carried out by a human being with only paper and pencil, and (iv) demands no further insight, intuition, or ingenuity. (cf. Copeland 2002, 1) General recursive functions are introduced by Herbrand (1932) and Gödel (1934). Without giving a detailed introduction, we can roughly describe general recursive functions to be functions $f : \mathbb{N} \to \mathbb{N}$ that are identical to a finite repetition of equations from a finite list (*e.g.* $3 \cdot a = c \Leftrightarrow a + a = b \wedge b + a = c$).[1] They are introduced to *explicate* the arithmetic of positive integers. The equivalence (a) is discussed by Church (1936) and Kleene (1936a; 1936b). λ-definable functions are introduced by Church (1932; 1936) and Kleene (1935). As opposed to general recursive functions, functions in the λ-calculus can be expressions with variables for integers and without any reference to recursions (*e.g.* a polynomial). The λ-calculus is very sparse and the allowance for *abstract terms* make it more similar to large parts of actual mathematics (*e.g.* analysis; algebra) than general recursion.[2] Equivalence (b) is discussed by Turing (1937). The function f is Turing computable, if f can be executed by a Turing machine in finitely many steps (cf. Turing 1937); more explanation follows below.

The thesis[3] is, to my knowledge, not proven, but also not disproven. For a disproof it is sufficient to show that there is at least one task that can be performed via a general recursive function in finite time, but not by a Turing machine or vice versa. Given the attention that computational logic received over the last decades it is at least fair to assume that the validity of the thesis is widely *accepted*.

How much does the Church-Turing thesis tell us about my claim, which gives the title to this section? We have to look at the two interesting notions from the Church-Turing thesis and relate them to computability in a more pragmatically applicable sense for the computers that we actually use on the one side, and to mathematical objects in general on the other side. Namely, we should have a look at how much Turing computability tells us about our actual computers and how

[1] A condensed English definition for general recursive functions can by found at Kleene (1936a, §1).

[2] Alama (2017) provides an introduction into the λ-calculus.

[3] My given version of the Church-Turing thesis does not refer to the pre-formal notion of *effectiveness*, which is the most appropriate historical starting point in any discussion about the thesis. However, the referred to notions of general recursivity and Turing computability are suitable for my aim to explain the relation between the capabilities to perform tasks by a computer and by an agent who is able to use only mathematical reasoning. *Effectiveness* is not thoroughly formalised and is intended to specifically explicate the capabilities of a human agent, who uses *paper and pencil* to perform (mathematical) tasks.—That is why Turing machines write and erase on tapes, which seems to be a rather unpractical approach in comparison to the architecture of actual computers (*e.g.* von Neumann architecture).

much general recursive functions tell us about mathematical objects.

I start with the computability aspects and want to keep it brief. Turing computability is very powerful and every digital computer that was built[1] until now can be emulated by a Turing machine. A Turing machine has a program and some storage space on its *tape*. As Hopcroft, Motwani and Ullman (2001, sect. 8.6), to which I refer regarding any computational theoretic claim in this paragraph, show, the number of steps to simulate any of today's known computer program by a Turing machine is at most polynomial to the number of steps of the original computer program. The relevant information for our theoretical investigation is that this number of steps is always finite, since our today's actually built digital computers can execute only a finite number of computations.

The computation of the value of a certain mathematical function $f(\vec{n})$ for some $\vec{n} \in \mathbb{N}^m$ with some $m \in \mathbb{N}$ is, intuitively speaking, a *recursive* process. That is why, intuitively, general recursion theory provides a suitable approach for the explication of actual mathematical functions. But what do theoretical insights about general recursive functions tell us about mathematical objects or terms in general? Note that set theory, which is today's most common approach to formulate a foundational vocabulary for mathematics, does not have to imply any recursion. However, *constructive* approaches for the real numbers (*e.g.* Cauchy sequences; Dedekind cuts) hint to the fact that recursive methods are seen as epistemically more suitable for a foundation of mathematics by a significant proportion of working mathematicians. An illustrative example for such a constructive and also recursive approach is Heron's method for the description of the irrational number $\sqrt{2}$ via an iteration:

$$x_0 \in \mathbb{R}_{>0}\,,\ x_{n+1} = \frac{1}{2} \cdot \left(x_n + \frac{2}{x_2} \right) \text{ for which } x_n \overset{n \to \infty}{\longrightarrow} \sqrt{2}.$$

Due to the build-up approach that the fundamental mathematical vocabulary implies which resulted from the *Grundlagenstreit* (*e.g.* set theory; predicate logic), and due to the many approaches to constructively define mathematical objects, it seems reasonable to stipulate that many mathematical objects are recursively constructable objects, even if our everyday mathematical vocabulary does not indicate this in every case clearly. That is why I believe that the theory of general recursive functions tells us a lot about mathematical objects in general and I can adopt the conclusion of the Church-Turing thesis to some extent for mathematical

[1] In the field of *hypercomputation* various sorts of computers are theoretically discussed that cannot be emulated by a Turing machine. These include computers that can solve the *halting problem*, that can store irrational numbers, or that can compute infinitely many steps in finite time. However, no actually hypercomputing machine exists today, including *(qubit) quantum computers*, which can solve only Turing computable tasks (cf. Bernstein and Vazirani, 1997).

objects in general.

If the reader is not convinced by these arguments (which are not proofs), I remind him that the Church-Turing thesis is convincing since no known theoretical problem can disproof it (*e.g.* halting problem). However, it should be noted that, as is widely discussed, the *axiom of choice* stipulates the existence of functions, which cannot be *constructively* described in set theoretic terms or by any known logical language. Since such a function is not constructive it can therefore also not be *recursively* defined. This counterexample is sufficiently highlighted throughout the literature[1] and the epistemological implications for mathematics were discussed to a sufficient degree.

However, my title giving claim in this section is not threatened by any non-constructive parts of mathematics. The claim is not build upon the equivalence that is stated by Church-Turing thesis; it gives a strong interpretation to the implication that Turing computable functions are always mathematical, but it does not imply any further interpretation for the other direction of the equivalence from general recursiveness to Turing computability. This leads to the conclusion that computers cannot execute tasks, which cannot also be described in mathematical terms (even if there are non-constructive mathematical objects that can never be instantiated by a computer).

As elaborated on in section 3.1, I suggest that category theory is the best widely discussed approach to explicate mathematical objects and I chose Grenander's general pattern theory as the epistemically preferable approach to explicate patterns due to its crucial criterion of providing a way of construction for a pattern. An epistemically easily accessible way of construction does not have to be given for sets of data, but data sets are always finite and can therefore always be trivially constructed by a finite description. That is why, for the purposes of this thesis, I can simply restrict the notion of mathematics to the one that mathematical constructivists suggest[2], who reject the axiom of choice. Mathematical objects in accordance with this notion are the mathematical objects or terms that can be constructed in mathematical terms, and these objects are equivalent to objects or terms that can be constructed by a computer, such like a von Neumann computer, if we accept this very strong version of the Church-Turing thesis.

[1] For a survey on the history and literature about the axiom of choice see Bell's (2015) encyclopaedia entry.

[2] Bridges and Palmgren's (2013) encyclopaedia entry provides an overview and a history of constructive mathematics.

3.3 Algorithms, Machine Learning and Epistemology

Section Abstract

Recent and not so recent developments in *artificial intelligence*, including *machine learning*, help to strengthen the view that scientific inferences, data and patterns can be mathematically explicated. Every computer program can be translated into a mathematical algorithm via decompiling. Actual recent technical developments help to undermine important implications of positions from Dennett and Searle from the philosophy of mind. In principle, only computable tasks can be rational scientific inferences by human agents. I take the following as a premise: an open question is *when* artificial intelligences have the computational power to perform all tasks that humans can perform regarding science, but not *that* this will happen.

In this thesis, mathematics plays a central role for epistemological and ontological explications. One of the reasons for this is that computer-made inferences play an important role in scientific practice (*e.g.* statistical analysis; data base searching). As mentioned earlier in this chapter, these inferences can, in some cases, only hardly be explicated in purely logical terms under pragmatic considerations, but they can always be explicated in mathematical terms. In this section I want to further strengthen the view that this epistemology of mathematics via computability is very powerful. With this I mean that the following doubt is not very threatening: computers and mathematics may play an auxiliary role for science, but any purely mathematical explication of data and pattern recognition falls short in regard to grasp the full epistemic role of data and patterns in the epistemic process of the human scientist's mind.

The argument implied in this threat has some similarities to the old philosophical question whether human decision making ist completely determined or in other words: is there *freedom of will*? In this section I, of course, cannot provide an answer to this long debate that would be convincing enough to persuade every reader and this is not my aim. What I aim for is to refer to very recent philosophical speculations about future computer technology that are usually referred to by the umbrella term 'artificial intelligence' (abbreviated 'AI'), and from these I follow that, firstly, the human organism does not, regarding the making of scientific inferences, provide any special capabilities that cannot be replicated by a machine,[1] and, secondly, since the arguments from the former section are

[1] I feel convinced that I am aware of the radical metaphysical, epistemological and religious implication of such a claim.

still valid for these future computers, reasoning and objects of reasoning can be mathematically explicated.

In section 2.2, commenting figure 2.4, I already highlighted my view that the *qualia* impressions that are induced by a set of data (*e.g.* the impression of red from an image) are not a part of the data, which can be explicated in mathematical terms and computer image storing is an example for this. Analogously, if an artificially intelligent agent makes inferences from a colourful image, this qualia seems to be absent from this process anyways, if we believe that qualia impressions are something only biological organisms (*e.g.* humans; bats[1]) can have. If we accept that qualia are not restricted to those organisms, and should be seen as aspects of artificial intelligences (made from silicon and copper instead of proteins, water and fat), too, then the elaborations from section 2.2 still apply.[2]

It is not a simple task to explicate what *artificial intelligences* are. They are most often described as computers or computer programs that are in some sense *intelligent* (whatever this may mean in detail)[3]. *Intelligence* may be defined by the classes of problems that can be solved—a program needs to be more intelligent to beat a grand master in chess than to calculate $\sqrt{17}$. But, obviously, the classification of tasks that can be solved only by an intelligent agent is as intricate as defining intelligence is. Turing (1950) suggests an *imitation game*, according to which a machine has a human level of intelligence, if it can convince any human after a merely textual interrogation that it could be a human agent. However, for the purposes of this thesis it is not necessary to explicate (artificial) intelligence. It is sufficient to say that there are artificial intelligences (find examples in the list below). The reason for me to still use this notion is that it is widely used in certain parts of the literature and it is used to solve problems that do not seem to be solvable by more *classical* computation methods like the implementation of a static algorithm that was preconceived by a human agent. *Static* means that the algorithms does not change itself with iterations.

I believe that there is no empirical ground for believing that humans can accomplish tasks that a machine could principally not and artificial intelligences are very suitable to support my conviction. These tasks include inferences in all scientific fields, creating *art* (whatever our notion of art is), using any language and driving a car. I also believe that a rejection of such a claim is less grounded

[1] This example is, of course, a little tribute to Thomas Nagel's (1974) paper.

[2] The idea of extending the notion of qualia from humans and animals to machines is not so far off as it may seem to some. For example, for a physical reductionist's or supervenience notion of qualia a perfect computer replication of a biological organism can instantiate pain or feelings by instantiating the corresponding physical state in the replication.

[3] An introductory discussion on the notion of (artificial) intelligences can be found at Russell and Norvig (2010, ch. I). A collection of recent papers on notions and tests for intelligence of computers can be found in an *AI Magazine* issue (2016).

on convincing arguments from the philosophy of mind, but rather on a dull belief in human uniqueness or *grandeur*, which has an influential philosophical and, more prominently, religious history.[1] This false traditional belief is nicely exemplified by Gary Kasparov's claim after being defeated in the first round of the chess match against IBM's computer, as Hsu (2002, blurb and p. 265) reports, "that only human intervention could have allowed Deep Blue to make its decisive, 'uncomputerlike' moves" that lead to its victory over the human world champion.

My argumentative technique—which is also the one of Bostrom, Kurzweil, Musk and others—could be called 'optimistic induction'[2]: the success of artificial intelligences that we observed so far makes me (and many others) believe that a human-level or even superior artificial intelligence is possible. Furthermore, we do not need more than to upscale our known computer architecture to built computers that match or exceed the *cognitive performance* (which I identify with the *computational power*) of a human brain. Figure 3.2 shows a numerical prediction of future computational power with some further explanation in the caption.

For a more detailed history and explanation on AI I refer to Bostrom (2014). For the purposes of this thesis it is sufficient to know that the upscaling in computational power of supercomputers does not have a general upper threshold[3] and that many intelligent solutions from computers (*e.g.* beating master Go players) are achieved by *machine learning* (and in particular *deep learning*[4]). Machine learning is the ability of a program to adapt its own program as a reaction to some feedback it received regarding an earlier application of one of its algorithms.—Figure 3.3 illustrates the case for an AI that is aimed to identify objects that are shown in an image.

One simple strategy to address an opponent's claim that human epistemic capabilities can never be matched by a silicon-copper based computer is to present a list of historical incidents at which machines accomplished tasks that only human

[1]The best study that I am aware of on the interplay between human consciousness, Christianity and specific human cultural capabilities is Jordan Petersons lecture on the psychological relevance of the Biblical stories (2017).

[2]This is, of course, a little reference to the *pessimistic (meta-)induction*, which is an idea from the debate about realism about unobservable objects and scientific theories. It roughly states that we should be pessimists about the reality of an unobservable entity or about the truth of an empirical theory, because so many earlier theories and stipulated unobservable entities turned out to be false or non-existent.

[3]Scales of computer chips are usually restricted by upper thresholds (due to cooling issues) and lower thresholds (due to the avoidance of quantum effects). However, arbitrarily large computers can be built by the use of cluster architectures and are constrained only by the availability of resources. Note that all computers that are connected to the internet can be seen as one large computer cluster.

[4]Deep learning is the application of artificial neural networks with more than one hidden layer. For our purposes it is enough to say that this implies that it is, for a human investigator, not easy or even epistemically impossible (depending on the sizes and number of hidden layers) to understand the causal processes of the artificial neural network *after* its training.

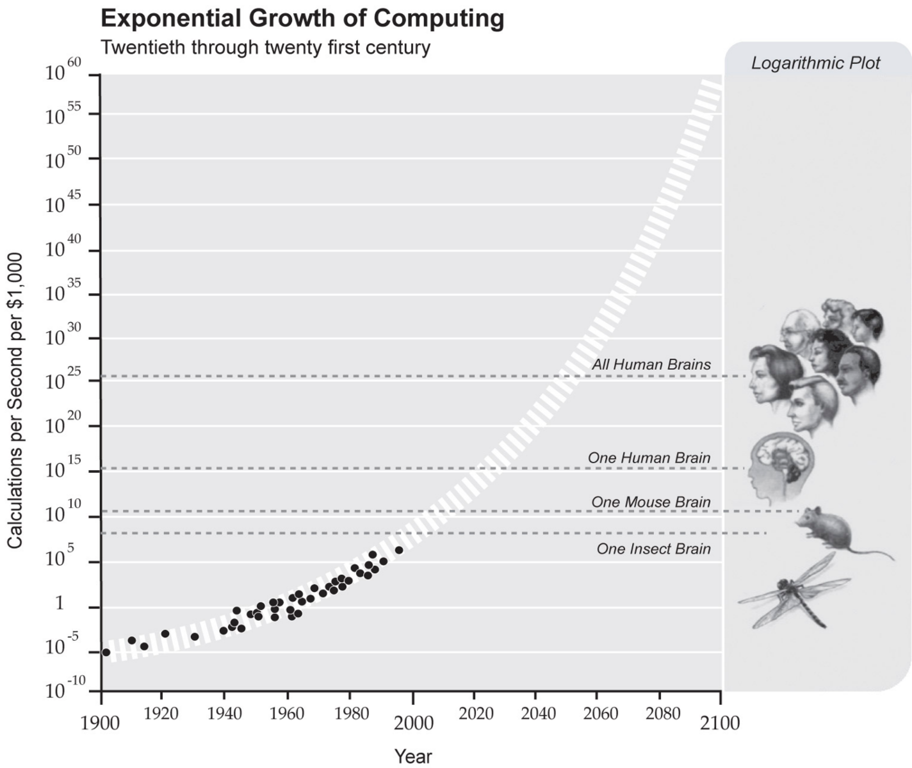


Figure 3.2: Prediction of future computational power in instructions per second for a fixed financial budget by Kurzweil (2005, p. 70). The growth rate is justified by an adaption of Moore's law about efficiency gains in processor production. The computational power of human (and animal) brains are estimated by calculating the necessary number of calculations that the retina has to do for visual image processing empirically observed at humans and extrapolating this performance to the mass of the entire brain.—Kurzweil adopted this method from Hans Moravec (cf. ch. 3, see there for further references).

beings were able to accomplish before the development of the machine. These are the starting points of the *optimistic induction*. However, the optimistic induction is neither a fully convincing empirical study (since we have no super AI yet), nor a proof of anything. But it serves as *falsifications* for the claims that the mentioned accomplishments demand human consciousness. To provide convincing arguments to confirm my belief in human-level future AIs to a sufficient degree much more work would have to be done. A lot more insights into the workings of the human brain and the mind are necessary. Therefore, for the aim of this thesis, I take this belief as a premise and the following list, at least, corroborates this belief. The examples include AI and machine learning:

- IBM's Deep Blues beat reigning world champion Gary Kasparov in a game of chess in 1996. Deep Blue uses an *evaluation function* to assign numerical values to the possible next moves from which the highest ranking one gets chosen. It demands, according to 1996 standards, massive parallel processing power. The evaluation function is statically programmed, meaning that no

`canine`	**`backseat`**	**`flaming poppy`**
`cavy`	`wheeled vehicle`	`corn poppy`
`batnam`	`self-propelled vehicle`	`iceland poppy`
`domestic dog`	`motor vehicle`	`opium poppy`

Figure 3.3: Three examples of images that were used as illustrations or characters in this thesis and the suggestions from Wolfram's (2015a) Image Identification Project. The first name in bold print is the AI's preferred solution. The next three names are the most likely alternatives according to the AI. These are results from August 9, 2017.

machine learning is involved. (cf. IBM 2001)[1]

- IBM's Watson achieved a victory over a group of human *Jeopardy!* champions in 2011. Watson is based on IBM's DeepQA software, which aims to *operationally understand* content that is phrased in the English language. (cf. IBM 2009) This operational understanding is the ability to provide adequate answers and questions in a conversation. For the match of Jeopardy, Watson was given an amount in the ballpark of 200 million pages of content from various sources as machine learning input. (cf. Best 2013) As Deep Blue, Watson demands massive parallel processing power. Watson received the answers, to which he has to find the correct questions, in text form and did not use speech recognition.

- DeepMind's AlphaGo beat the European champion Fan Hui in the board game Go, and after that it beat world champion Lee Sedol in 2015. DeepMind (2016) summarises their conclusions regarding AlphaGo's style of play and the game Go:

  > During the games, AlphaGo played a handful of highly inventive winning moves, several of which - including move 37 in game two - were so surprising they overturned hundreds of years of received wisdom, and have since been examined extensively by players of all levels. In the course of winning, AlphaGo somehow taught the world completely new

[1] Campbell, Hoane and Hsu (2002) provide a more detailed computer scientific description of the Deep Blue project. The leading computer scientist of the IBM project, Feng-Hsiung Hsu (2002) also provides a historical survey, including social aspects of the working group and the personal interaction with Kasparov.

knowledge about perhaps the most studied and contemplated game in history.

As Silver *et al.* (2016) calculate, Go is combinatorially vastly more complex than chess. AlphaGo uses machine learning and a neural network, which was trained by human supervising experts and via reinforcement learning from games of self-play with introduced random events. It uses GPU processing architecture.[1]

- The Wolfram Language Image Identification Project (2015a) uses machine learning of an artificial neural network to identify objects in photographs (*e.g.* a cat; a bridge). It is trained with a few tens of millions of curated photographs and uses GPU processing architecture. (cf. Wolfram Research 2015b, Stephen Wolfram 2015). Figure 3.3 shows some results of the software.

- In 2015, a machine learning algorithm was able to reach human-level results in a narrowed version of the ImageNet Large Scale Visual Recognition Challenge. (cf. ImageNet 2016) For all of the five tasks a single visual example is given, and a classification or reproduction of the shown object has to be made by the AI. Figure 3.4 shows an exemplary classification task. The program runs on a GPU architecture and uses probabilistic reasoning. The relevant information for this example is that it seems particularly challenging to build an AI that is able to provide human-level results for a tasks with only one single input image, which significantly differs from mastering chess or Go via millions of examples and training games. (cf. Lake *et al.* 2015, and NVIDIA 2017, p. 16)

Figure 3.4: The task is to find a similar object in the list with the framed object as single input (Lake *et al.*).

- Libratus, an AI programmed by Tuomas Sandholm and Noam Brown from Carnegie Mellon University, beat a team of four professional human poker Texas hold'em players in January 2017. Libratus uses machine learning and, as opposed to the other mentioned AIs, it uses a regret minimization function. The case of poker is of particular interest in comparison to chess or

[1] **G**raphical **p**rocessing **u**nits can execute a vast amount of similar calculations much vaster than the classical **c**entral **p**rocessing **u**nit, which is the common processing unit in the von Neumann architecture. GPUs were first applied for the rendering of large images on computer screens, and more powerful GPUs were developed for video gaming applications. The architectonical difference between CPUs and GPUs is that for CPUs all information that enters the processing on the input side has to be buffered (*i.e.* physically stored), whereas GPUs receive a their input as a direct stack and they are optimised for automated parallel processing.

Go, since it is a game with *imperfect information* (*i.e.* you do not know your opponent's cards, but he does) and possible deception (*e.g.* bluffing). (cf. Hsu 2017) Libratus' success hints to the possibility that AIs can be superior to humans in other imperfect information games, such as price negotiations or military battle strategies.

- Almost all major car makers develop hard- and software to achieve full autonomous driving. The chip manufacturer NVIDIA claims to provide artificial intelligence chips with sufficiently high performance, low scale and economic affordability. (cf. NVIDIA 2017, p. 19-20)

Recent discussions about (super-human) AI focus on civilizational and ethical implications of future AI development. Relevant questions are how future AI may act towards humans and what regulatory constraints should be enforced rather than what they are able to do. The answer to the last question is that they can do everything computable. Note that no human can solve an incomputable problem either. However, human cognition does usually not require *explications* of neither the tasks, nor the solutions that a human can provide. This thesis is not the place to discuss these mentioned problems. However, we can very reasonably assume that AI will be more and more capable of achieving things only human beings can do so far (and even beyond that). The only relevant remaining question is how fast does the available computability power grow (see figure 3.2).

Some disagree with my convictions at this point. In particular, Dennett (see section 4.3) introduces an ontological distinction between patterns in the sense of information technology or mathematics on the one hand, and “real patterns” that can be recognised by humans on the other hand. Obviously, AIs are not restricted to “real patterns” in Dennett's meaning. Such a *anthropocentric* epistemological or even ontological perspective can also be directly aimed to the comparison between humans and AI. Another fitting example of anthropocentric is Searle's (2014) recent comment concerning Bostrom's (2014) predictions (and warnings) regarding super-human AI. His ideas have a strong relation to Dennett's thoughts by claiming that the human mind adds something extra to the process of reasoning and this extra is in principle unachievable with any silicon-copper architecture.

I strongly reject Dennett's and Searle's claims and I speculate that they have their roots in a profound misconception of AI and machine learning. It is true that computers can execute only algorithms like a common pocket calculator. But the examples above (*e.g.* chess; Go) show that this leads to remarkable accomplishments, if we consider the mere vastness and calculation speed that can be realised with a computer.—Recall the confidence in their abilities of the challenged human master players *before* the matches. I show Searle's reasoning in a longer passage to further discuss this opposing position and to highlight the misconception:

> When I, a human computer, add 2+2 to get 4, that computation is observer independent, intrinsic, original, and real. When my pocket calculator, a mechanical computer, does the same computation, the computation is observer relative, derivative, and dependent on human interpretation. There is no psychological reality at all to what is happening in the pocket calculator.
>
> (...)
>
> If we ask, "How much real, observer-independent intelligence do computers have, whether 'intelligent' or 'superintelligent'?" the answer is zero, absolutely nothing. The intelligence is entirely observer relative. And what goes for intelligence goes for thinking, remembering, deciding, desiring, reasoning, motivation, learning, and information processing, not to mention playing chess and answering the factual questions posed on Jeopardy! In the observer-independent sense, the amount that the computer possesses of each of these is zero. Commercial computers are complicated electronic circuits that we have designed for certain jobs. And while some of them do their jobs superbly, do not for a moment think that there is any psychological reality to them.
>
> Why is it so important that the system be capable of consciousness? Why isn't appropriate behavior enough? Of course for many purposes it is enough. If the computer can fly airplanes, drive cars, and win at chess, who cares if it is totally nonconscious? But if we are worried about a maliciously motivated superintelligence destroying us, then it is important that the malicious motivation should be real[sic!]. Without consciousness, there is no possibility of its being real.
>
> What is the argument that without consciousness there is no psychological reality to the facts attributed to the computer by the observer-relative sense of the psychological words? After all, most of our mental states are unconscious most of the time, and why should it be any different in the computer? For example, I believe that Washington was the first president even when I am sound asleep and not thinking about it. We have to distinguish between the unconscious and the nonconscious. There are all sorts of neuron firings going on in my brain that are not unconscious, they are nonconscious. For example, whenever I see anything there are neuronal feedbacks between V1 (Visual Area 1) and the LGN (lateral geniculate nucleus). But the transactions between V1 and the LGN are not unconscious mental phenomena, they are nonconscious neurobiological phenomena.
>
> The problem with the commercial computer is it is totally nonconscious. In earlier writings, I have developed an argument to show that we understand mental predicates—i.e., what is affirmed or denied about the subject of a proposition—conscious or unconscious, only so far as they are accessible to consciousness. But for present purposes, there is a simpler way to see the point. Ask yourself what fact corresponds to the claims about the psychology in both the computer and the conscious agent. Contrast my conscious thought processes in, for example, correcting my spelling and the computer's spell-check. I have a "desire" to spell correctly, and I "believe" I can find the correct spelling of a word by looking it up in a dictionary, and so I do

> "look up" the correct spelling. That describes the psychological reality of practical reasoning. There are three levels of description in my rational behavior: a neurobiological level, a mental or conscious level that is caused by and realized in the neurobiological level, and a level of intentional behavior caused by the psychological level.
>
> (...)
>
> Bostrom tells us that AI motivation need not be like human motivation. But all the same, there has to be some motivation if we are to think of it as engaging in motivated behavior. And so far, no sense has been given to attributing any observer-independent motivation at all to the computer.
>
> (Searle, 2014)

What went wrong here? He infers from *consciousness* the possibility of having a *motivation*. Obviously, his notion of motivation requires consciousness. But Bostrom's notion of motivation is a different one. For him it is reasonable to say that a chess computer has the motivation to beat his opponent, because it was programmed to aim for a victory. First, in this thesis I do not want to argue about the notion of consciousness and also not about the notion of motivation. Second, I do see how we can rule out that a computer, which we may build in a distant future may have consciousness; one way to try this is to perfectly replicate the physical processes in a human brain with a computer. Searle discusses this idea more thoroughly in an earlier paper:

> Any attempt literally to create intentionality artificially (strong AI) could not succeed just by designing programs but would have to duplicate the causal powers of the human brain. (1980, p. 417)

A replication can be a mere program and therefore the above mentioned author's and I disagree with Searle. Third, even if we totally agree with Searle's claim that AIs are and will ever be principally non-conscious, what *operational* difference does it make for our society and science? A non-conscious future AI may be better than any human can be at proving mathematical theorems or formulating physical theories. For the practical reality of science it does not make any difference, if the super-human AI provides better scientific solutions consciously or not. Operationally, consciousness is nothing more than a terminology of specific human pride, which has religious roots, and in particular religious roots. More specifically, Searle's idea is, at its core, a rebranding of the classical Christian idea of a hierarchy of souls, as discussed by Thomas Aquinas' *Summa Theologiæ* (1265–1274, Ia.75) and others.[1] According to this tradition, a human inherits another class of soul than a thing or a plant or an animal or an angel. This classification is normative without any physical arguments like a reference to neurological complexity.

[1] These include Aristotle's *De anima* (see Bekker 1831–1870) and Leibniz' *Monadology* (1714).

That is why according to this Christian account, a mentally disabled person or a living human person with extreme physical brain damage is as much a human person as any other human is, because every human inherits a human soul. One could bring forward the argument that a certain brain function must be intact at a human organism to classify this human as a person. And this brain function is exactly what has to be replicated by a AI to count as a person as well.

Searle adapts Dennett's meaning of the term 'real', which is very central in their account and profoundly misleading, as I elaborate on in a later part of this thesis (section 4.3). However, to give a vivid example, if an AI sends a drone to kill you, this problem will become very *real* (by every possible meaning of the word) to you very soon. In such a discussion it is *fruitless* and irrelevant for Bostrom to attribute *conscious motivation* to it or not.

Bostrom warns of the danger that a future AI can be programmed by not well-meaning people (*e.g.* terrorists), or that it can ethically misinterpret a programmed maxim (*e.g.* autonomous car kills his driver to avoid accidents that are caused by the driver's flawed interventions). My aim is to strengthen the view that scientific inferences *can* always be mathematically explicated. A scientist may have consciousness, but so does the chess player. Nevertheless, an AI can be the superior chess player. The fact that they can be outperformed by an AI shows that consciousness is not a relevant factor for these cognitive achievements. As I elaborate on later (4.3), for my aim it is sufficient to restrict the analysis of scientific reasoning to the operationally explicable aspects and those are the aspects that matter, whereas states of minds or consciousness are not relevant for my topic. With this I mean that these things or processes do not add anything *operationally relevant* to a scientific inference.

In order to make things more clear, as one substantial claim of this thesis I defend the position that scientific data, patterns, inferences and background assumptions are mathematical, ontologically speaking. An important argument for this position is the introduction of AI agents into science for auxiliary purposes, but also for more profound tasks like hypotheses formulation. This implies that a lot of epistemological questions of science can be answered with relative ease via the reference to mathematics. However, as will become more apparent in the next chapter, this position does by no means imply any ontological commitment towards the empirical world *out there*, the Kantian *Ding an sich*. My position is agnostic towards the option of us living in a Cartesian-Bostromian dystopia and only because we think and we reason in mathematically explicable terms, which are *structures*, this does not imply any arguments in favour of an ontic structural realism as Ladymann and Ross (2007) propose it. In my ontology, science pictures the world mathematically, but the world itself is something we have insufficient metaphysical access to, to get to know what it is.

3.4 Conclusion

I presented a structuralist's notion of mathematical objects that may best, but not perfectly, be explicated by category theory. Furthermore, according to this notion, we do not have to distinguish between epistemological and ontological aspects regarding mathematics and regarding scientific inferences. The reason is that I reject any fundamental distinction between mathematical and scientific knowledge by a human agent on the one side and by an artificially copper- and silicon-based intelligent agent on the other side. Mathematics is a superior approach to predicate logic for the pragmatic explication of scientific inferences since it is already widely used in explications for everyday science and computer programs can much more easily be explicated in mathematical terms than in terms of predicate logic.

The Church-Turing thesis shows that computer tasks are equivalent to a certain class of mathematical functions. That is why computer programs, including all forms of today's and future artificial intelligences, can be explicated in mathematical terms via decompiling. This is true even for cases in which human agents were not epistemically able to understand this mathematical explication due to its complexity or magnitude. Current developments in the field of artificial intelligence can be extrapolated to the assumption that any epistemically possible human scientific achievement can be a result of an artificially intelligent problem solver. I take this as a premise. It follows that we have very good arguments to suppert the view that scientific inferences can always be mathematically explicated and this general explicative classification of scientific reasoning also holds for non-human scientific agents, which are already in use in everyday science in form of, for instance, statistical pattern recognition programs.

Chapter 4

From Data to Patterns

Chapter Abstract

The relation between patterns and phenomena is not as simple as (Bogen and) Woodward's examples indicate. We need to distinguish between *concrete* and *general* patterns. Patterns cannot depend on human sensory and cognitive capabilities, but non-anthropocentric epistemological criteria play a substantial role for an explication of the notion of a pattern. Patterns are purely mathematical, but they have to be *constructive* for epistemic reasons. That is why we cannot explicate patterns with classical information theory and I apply the more adequate *general pattern theory* for this.

Due to the great variety in which data occur in science and that are used to detect phenomena, the concept of patterns in data needs to be very flexible. If data are abstract objects in the sense described in chapter 2 (*i.e.* mathematical objects), then patterns in data are abstract objects in this sense as well. Mathematical objects can have only mathematical properties. An important, but rather epistemological question is whether all pattern recognition techniques that are used in science can be described as mathematical algorithms. This implies that pattern recognitions by human beings would be nothing else than a very efficient execution of mathematical algorithms. As we will see, patterns are mathematical objects, even if we are not always able to explicate them precisely as such in practice.

In this chapter with 'patterns' I refer to something that plays the role in scientific inferences that (Bogen and) Woodward imply by their use of term 'pattern'. This notion is more appropriately named 'patterns in data', since patterns are always an actual or possible feature of an actual or possible set of data. As will become clear in this chapter, this notion excludes other meanings of 'pattern', such as what in this thesis is named 'structure' (*e.g.* Resnik 1997). However, the notion of patterns that is applied here is neither trivially explicated, nor restricted to occurrences in strictly scientific contexts in a narrow sense.

It is rather the case that data in science are something mathematical *because* inferences from data to phenomena (or to theory) can in principle be explicated as mathematical pattern detection routines than the other way round.

4.1 General Patterns, Concrete Patterns and Pattern Similarity

Section Abstract

I introduce the concepts of a *general* and of a *concrete* pattern. Both of them are mathematical objects. General patterns are the patterns that are most similar to what (Bogen and) Woodward refer to with 'pattern'. Concrete patterns are the specific mathematical objects that are identified in data. I also introduce the notion of pattern similarity. Two concrete patterns are similar to each other, if there is a not very complex mathematical algorithm definable to transform the one concrete pattern into the other one.

A pattern is a property that a set of data can have and, if so, the pattern is called to be *detected* or *recognised* in the data. I already mentioned that data is epistemically accessible only via physical or mathematical representations. Given the explanation of data and of data representations in chapter 2 and in particular 2.3, a pattern can be explained to be detected in data, if at least one representation of the data exists in which the pattern can be detected. The distinct advantage of this explication of pattern detection in data is that we can apply it to the different representations of data, but still have an appropriate explication for the data as a concept that fits more to the everyday use of the term 'data' in scientific contexts.

Different sets of data from the same part of the empirical world under investigation that are scientifically used to detect the same phenomenon can be very difficult. A simple illustrative example is to detect albinism in an animal or in a human being. One set of data can be measurement results from the colour of skin, hair and eyes. The other set of data can be results from a genetic analysis. Figure 4.1 illustrates this example. Scientists would likely not say that they can recognise *the* according pattern of albinism in both sets of data. That is why it is not easy to define this specific exemplary pattern as *a* feature of one set of data. This problem does also occur in a simplified scenario, in which both sets of data are translated into purely numerical representations and the pattern(s) are recognised solely by the application of mathematical methods.

In 2.3 I introduced the distinction between the data and its representations that human beings can epistemically access. If a phenomenon can be detected

```
  1 | MLLAVLYCLL WSFQTSAGHF PRACVSSKNL MEKECCPPWS GDRSPCGQLS
 51 | GRGSCQNILL SNAPLGPQFP FTGVDDRESW PSVFYNRTCQ CSGNFMGFNC
101 | GNCKFGFWGP NCTERRLLVR RNIFDLSAPE KDKFFAYLTL AKHTISSDYV
151 | IPIGTYGQMK NGSTPMFNDI NIYDLFVWMH YYVSMDALLG GSEIWRDIDF
201 | AHEAPAFLPW HRLFLLRWEQ EIQKLTGDEN FTIPYWDWRD AEKCDICTDE
251 | YMGGQHPTNP NLLSPASFFS SWQIVCSRLE EYNSHQSLCN GTPEGPLRRN
301 | PGNHDKSRTP KAPLFSDVEF CLSLTQYESG SMDKAANFSF RNTLEGFASP
351 | LTGIADASQS SMHNALHIYM NGTMSQVQGS ANDPIFLLHH AFVDSIFEQW
401 | LRRHRPLQEV YPEANAPIGH NRESYMVPFI PLYRNGDFFI SSKDLGYDYS
451 | YLQDSDPDSF QDYIKSYLEQ ASRIWSWLLG AAMVGAVLTA LLAGLVSLLC
501 | RHKRKQLPEE KQPLLMEKED YHSLYQSHL
```

Figure 4.1: Left: The amino acid sequence of the 11$^{\text{th}}$ chromosome with the permutations (underlined) that indicate albinism. (Tomita *et al.* 1989). For an explanation of the used abbreviations see footnote 3 on page 114. Right: The appearance of a child with albinism (Wikimedia).

in two different representations of a set of data by the according pattern(s), then it is vague whether we want to call this *a* pattern or *two different* patterns. On the one hand, in the case of an entirely mathematical describable scenario, the two representations of the data and the according pattern(s) can be very similar. Consider the example of stock prices: we can analyse the volatilities on the basis of the data given as prices in Euro and two digits for the cents (*e.g.* EUR 45.37), but also on the basis of the same data in a log-return format of full percentages (*e.g.* -17%).—An example is plotted by figure 2.5 on page 49. Rather simple calculations without additional empirical knowledge are sufficient to translate one representation of the data into the other. But in the example of albinism it is very unlikely that a biologist will be able to find this sort of similarity between sets of data from a genetic analysis and results from the measurement of skin colour.

This intuitive notion of *similarity* between two sets of data or patterns in them is hard to mathematically specify fully. Price times series and return time series seem to be very similar sets of data, because very simple rules for conversion between them are available. The same holds for many other scientifically interesting patterns in these sets of data. For the two mentioned sets of data regarding albinism a simple conversion rule is exactly what is missing. An epistemically unbounded agent could explain how exactly, biologically speaking, one of these patterns always coincides with the other (*e.g.* via melanin production) but a vast amount of empirical scientific *knowledge* (not only assumptions) is necessary to do this. We would have to *perfectly understand* DNA (and RNA) to say which amino acid lead to which phenotype.

Even if all pattern recognition procedures in science can be described by mathematical algorithms that are employed on mathematical data, a possible similarity cannot be explained by some form of sophisticated *isomorphism*. Other specifications of simple relations between mathematical structures, fail to explain this similarity to its full extent as well. Therefore, albinism can be detected by patterns that are structurally completely dissimilar to each other.

Again, consider a scenario in which not only the representations, but also the two sets of data are structurally very different, and have been gathered with the intention to detect the same phenomenon by pattern recognition. In the jargon adopted by (Bogen and) Woodward, scientists would talk about only one pattern in this scenario (*e.g.* "the pattern of albinism"). But it might be hard to justify this ambiguous designation even in a very simple scenario setup with only numerical data and (a) mathematical pattern(s). Depending on the type of data, the mathematical techniques that are used to recognise the patterns can be completely different; this applies even if the relevant theory indicates that these two patterns in the data can be recognised due to the same phenomenon, as the example of albinism shows.

The solution to the described problems is to conceptually distinguish between the notions of a *concrete* and of a *general* pattern that I introduce here. This idea is in some aspect analogue to my explanation of data, but in some not. What is usually called a pattern in the terminology of (Bogen and) Woodward is in my terminology a general pattern, which is a class of concrete patterns. Two of these concrete patterns from the example of albinism above are the following. Firstly, the measurement results show a white-pink colour of skin, white hair and light blue eyes with red parts. Secondly, in the records of a genetic analysis the tyrosinase (TYR) gene in the 11^{th} chromosome is mutated in a certain way.[1]

We can outline a formalisation for the resulting general pattern. The actual general pattern of albinism may consist of much more concrete patterns, but this is a simplified example based on the discussed two concrete patterns.

$$\begin{bmatrix}\text{general pattern}\\ \text{of albinism}\end{bmatrix} = \left\{ \underbrace{\begin{bmatrix}\text{white hair and}\\ \text{light skin and}\\ \text{reddish eyes}\end{bmatrix}}_{\text{concrete pattern}}, \underbrace{\begin{bmatrix}\text{mutation of positions}\\ 311,\ 312,\ 314,\ 315,\ 316,\ 317\\ \text{at the } 11^{\text{th}} \text{ chromosome}\end{bmatrix}}_{\text{concrete pattern}} \right\}$$

How then are general and concrete patterns related to data and its representations? Let us, at first, discuss an example of one set of data with two different representations of it. Analogously, we then speak of one (general) pattern and two different concrete patterns of it. But we need to be careful with the following point. Let us assume two sets of data are available in which a phenomenon can be

[1]For simplicity, I focus only on a certain strong form of albinism biologically classified as *OCA1A*.

detected in different representations for each set of data. Since we want to assign only one pattern to a phenomenon, we, again, speak of one general pattern with two concrete patterns that are detected in the different representations for each set of data.

To further prepare the ground for the explanation of concrete patterns I want to provide a little excursus concerning data representation and mathematics to recall the results from chapter 2. Representations of data are not all mathematical. Photographs or texts are examples. The more important question for us is whether the pattern recognition procedures that are used in science are reducible to the execution of mathematical algorithms in principle. If this is the case, then the interesting aspects of representation of data are always mathematical. To illustrate this point more simple: if pattern recognition techniques executed by human beings can be translated into the executions of (possibly extremely complex) mathematical algorithms, then these algorithms must be applied to a purely mathematical representation of the data. In this case, every non-mathematical aspect of a representation of data is only pragmatic to support the accessibility of data to human beings with their sensory and cognitive capabilities.

Should we reduce our philosophical analysis to mathematical representations of data? Pattern recognition techniques that are conducted by computers are a relative novelty in the history of science. Furthermore, for many needs concerning pattern recognition in science computational routines are not developed enough. To illustrate this point with an example from outside of the sciences, already very simple tasks for pattern recognitions in images are designed for computer security reasons to distinguish real human IT users from programmed automatisations.[1] These tasks can easily be performed by most human beings, but they cannot easily be performed with today's computers. Thus, the picture of today's actual pattern recognition in science as a discipline of applied mathematics seems wrong. But can it be reduced to this *in principle*?

The applications of statistical methods in various scientific fields are examples for the use of mathematical algorithms to detect patterns. But the question whether all pattern recognition procedures that are conducted solely by human beings are nothing else than a pragmatic way to conduct a possibly unknown and very complex mathematical algorithm is unanswered. Due to the historical success of computational data analysis (find a thorough discussion in 3.3) and the lack of arguments for the opposite position we should rather affirm than refute this position. This point can be illuminated and strengthened by simply asking for the

[1] These *CAPTCHA*s (*Completely Automated Public Turing test to tell Computers and Humans Apart*) are a simple illustration of the boundaries of today's computerised pattern recognition in images.

falsifying fact: Is there a pattern recognition procedure in science that *in principle* cannot be described by a mathematical algorithm for some reason?

Under the assumption that all aspects of the data itself that are relevant to science can be expressed by a mathematical representation of this data, what then are the concrete patterns? A concrete pattern in a mathematical data representation is a mathematical *model* (in the sense of: exemplifying certain aspects) to which the data representation has to *fit* (in the statistical meaning of ‘fit’). Fitness in this sense is a concept that is very difficult to mathematically explicate. Analyses of photographs, texts or sounds in science must be applicable to this theoretical concept of mathematical models and its fitness to data. Therefore, this explication is an ambitious task. In this text I can only assert that this concept seems intuitively adequate. A photograph of an elephant fits to the visual concrete pattern of an elephant. This example shows how hard it is to explicate fitness, but since we and also modern AI software (cf. 3.3) are able to recognise the elephant, there must be a concept of fitness in the described sense.

Why is the relation between a phenomenon and a corresponding pattern not a *representation*? Two cases need to be considered: first, does a concrete pattern represent a phenomenon, and, second, does a general pattern represent a phenomenon? However, the answer to the full question can be outlined briefly. Whatever our notion of representation may be in detail (see 5.3 for further discussion), a precondition for a relation of representation is that one object or term represents another object or term. But as the example of albinism exemplifies, the notion of the phenomenon is empty, if we strip away its general pattern (and therefore all concrete patterns). This implies that the relation between a phenomenon and the corresponding pattern is much more defining for the phenomenon than the relation of representation between objects or terms is. In other words, this is a case of *relation without relata*-problem on the phenomenon side. The general pattern *specifies* the phenomenon on the background of further empirical assumptions and nothing but the general pattern can be a specification of the phenomenon. I elaborate more on the relation between phenomena and representation in section 5.3.

To summarise, patterns exist in two ways: in a concrete form and in a general form. The concrete patterns can in principle be formalised by a mathematical model that can be applied to certain mathematical representations of data. Not mathematically explicated pattern recognition procedures in science occur due to merely pragmatic reasons. General patterns are classes of concrete patterns and specify the phenomenon.

4.2 Problems and Literature Regarding Patterns in Science

Section Abstract

Patterns are related to repetition and the human sense for aesthetics. However, it is not an easy task to explicate the class of all patterns. Pattern recognition, unlike rationality, depends on specific agents or types of agents (*e.g.* humans; aliens). Mathematical approaches for the explication of the class of all patterns are available.

The term 'pattern' is widely used by scientists and philosophers of science. It is also used in various artistic and everyday contexts. Most scientists and also philosophers would agree on the claim that patterns have a lot to do with repetition and, in more particular, visually perceivable repetition. According to the common understanding, mosaics of tiles or certain repetitions in an one-dimensional time series are examples of patterns; see figures 4.2[1] (p. 108) and 4.3 (p. 109) for illustrations of everyday examples according to these descriptions.

Furthermore, human beings seem to apprehend certain visual patterns as easy-to-detect with their natural sensory and cognitive capabilities. To other patterns, which are, for instance, very *complex* (in the sense of: we would need a lot of place on a sheet of paper to describe them) or spread out over a very vast amount of data, we are more or less blind. Visual patterns are also even considered *aesthetically harmonic* (whatever that means in detail), as their various uses for decorational purposes indicate. These intellectual and aesthetic compatibilities of patterns to the human perception and mind indicate that patterns are a very cornerstone of human perception and the cognitive processing of their surrounding world in general. Aiming for a description of basic human epistemic capabilities and the *scientific method*, pattern detection seems to be an ability that is pragmatically at least as fundamental as rationality (in the sense of: applying logical reasoning) is. However, scientists and philosophers often refer to a concept of patterns without a further explanation, what a pattern is, and without any references to other clarifying texts about the topic. If patterns are at least as fundamental to the scientific method as rationality is, then we need a convincing description of what a pattern is. The disregard of this explication indicates that most philosophers do not find it to be an intricate task (or deem it not worth doing); however, as

[1] The Islamic dome is often considered as an iconic landmark in Jerusalem and Islamic architecture in general. It was built in the seventh century and is heavily influenced by the architecture of churches in the city. The shown part of the facade is the result of a redesign under the Ottoman Suleiman the Magnificent in the 16th century. For the art-historical survey that is the source of this information, see Avner (2010).

we will see, some philosophical and mathematical discussions are available and I claim that this explication of patterns is neither trivial, nor fruitless concerning our main task, which is a better understanding of scientific phenomena and their relations to patterns in data.

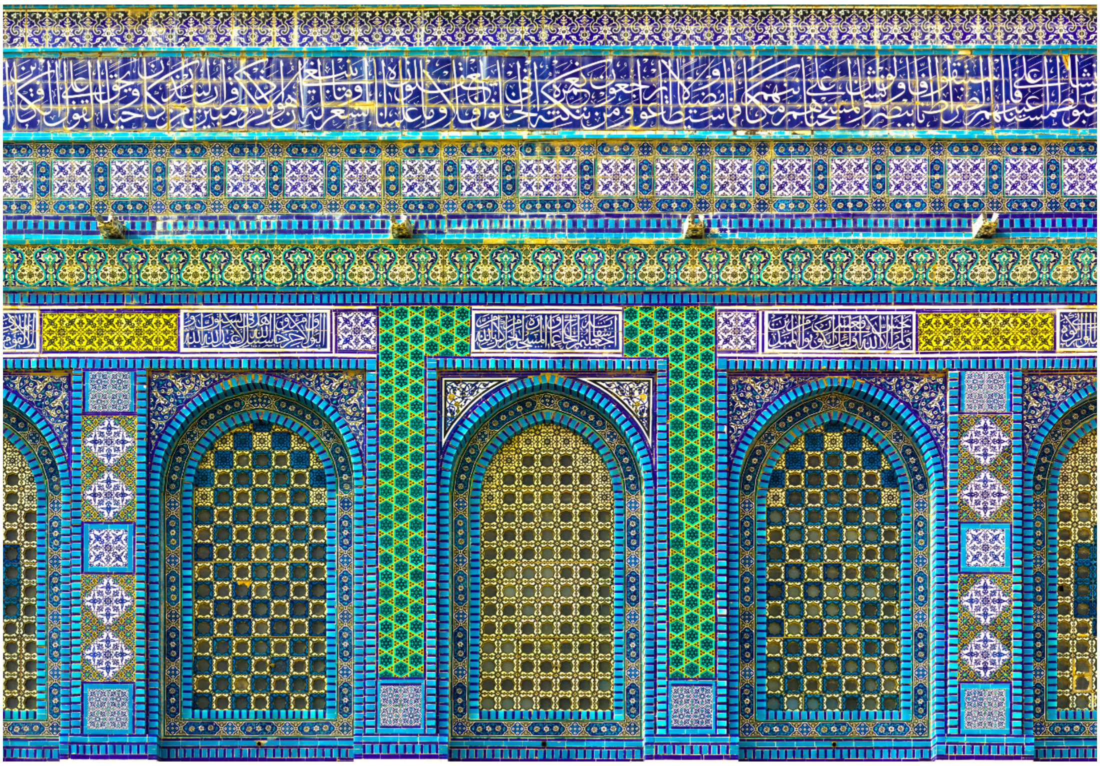


Figure 4.2: Part of the facade of the Dome of the Rock in Jerusalem. The tiles are arranged in (a) pattern(s). (Photo by A. Shiva via Wikimedia Commons)

The illustrated examples show some important aspects concerning the explication of a concept of patterns. I selected them with care to exemplify various important aspects of patterns. They stimulate some specific questions one may have about patterns prima facie.

Do we see one or many patterns in figure 4.2? Are there patterns of patterns, such as known from *self-similar*[1] geometric objects as more apparent at, for example, the Sierpinski triangle (figure 4.4)? The aesthetic effect that the mosaic has for human beings is obviously influenced by the colours, too; do we have to include not only shapes, but also colours into our concept of patterns? Is the text in the top part a pattern or not or a part of the pattern? It surely contributes to the aesthetic effect, but it behaves differently concerning the repetition of shapes.

The plot in figure 4.3 is not directly *motivated* by aesthetic purposes, but rather by (socio)scientific or business analytical needs. Apparently, the number of monthly passengers follows seasonal trends (i.e vacation travel) and overall growth

[1] I (2012) offer a brief introduction into self-similar geometry and stochastics in my diploma thesis about multifractal stochastic processes.

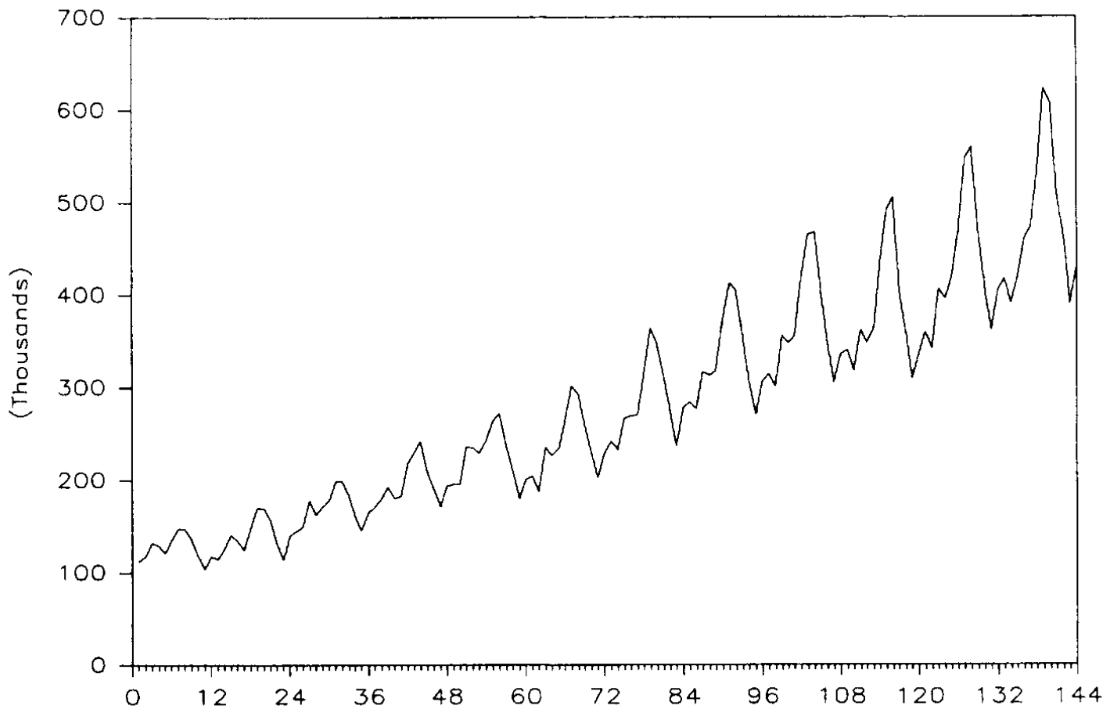


Figure 4.3: Plot of monthly international airplane passengers from January 1949 to December 1960. The exponential drift and seasonal spikes are a pattern. (image from Brockwell and Davis 1991, data from Box and Jenkins 1976)

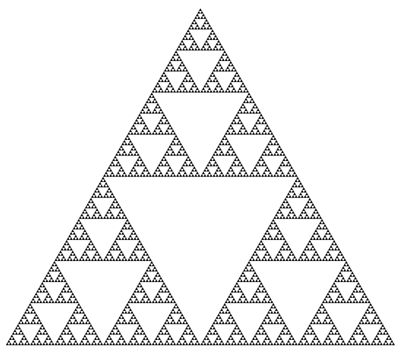


Figure 4.4: A Sierpinski triangle, a *self-similar* geometric object. (Wikimedia)

(*i.e.* decreasing ticket costs in average working hours). A mathematically more detailed discussion on statistical pattern analysis and on this example in particular is given in the appendix, section A.1. Do the exponential drift and the seasonal influences make one pattern or many patterns or one pattern with sub-patterns? This *ideal* pattern is comparatively easy to explicate in common mathematical terms and seems to be deformed by noise, but it is very easy to detect with the naked eye. What separates a pattern from noise and what exactly is noise? Can a pattern be noisy by itself or are there *ideal* patterns that are distorted by noise? What does it tell us epistemologically that this pattern is very easy to grasp for a human being?

Do, ontologically speaking, patterns depend on human sensory or cognitive capabilities in general, or do they even depend on single investigating agents? Recall that—let us compare—for *rationality* most philosophers would deny similar claims, but for patterns many, such as Dennett (1991), would not. Rationality is, if anything, normative and therefore a postulated ideal for reasoning in scientific contexts and in general. In the case of the dependence of a pattern on a single agent, would this agent be idealised or does it have to be an actual agent, with all of his/her individual peculiarities and limitations like specialised education, professional experiences, blindness, colour-blindness, autism, dyslexia, racism and so on? If patterns depend on human investigating agents, how are they related to the solely visual capabilities—note that for Kant (1787) it is a cornerstone of his conception that synthetic *Urteile* (engl. *judgements*) a priori are possible by only our *innere Anschauung* (engl. *inner illustration* (sic!)[1]), which has very strong relations to our visual capabilities.

In the following course of this chapter I defend the view that there is a normative and descriptive account for patterns in science, in a similar way as there is one for rationality (even if there is a lively philosophical discussion about the *correct* account). Our specific human sensory and cognitive capabilities do play only a pragmatic role regarding patterns without any ontological implications. In my explication and the examples of patterns I focus on patterns that play a role in a process of scientific discovery. But I believe that there is not much more to add to the concept of patterns, if we loosen this restriction and refer to also non-scientific purposes, *e.g.* patterns that occur in the context of engineering, of art, of decoration or in children's games. My account also provides a descriptive answer to the question what exactly a pattern extensionally is, but not so much what sorts of patterns are easy-to-detect with basic human sensory or cognitive capabilites, whatever these may include. This account is epistemologically normative by providing a precise structural and therefore mathematical and logical[2] account of patterns; the account is descriptive by covering all of our empirical examples—but also many more—of patterns from several scientific fields.

Apart from the discussion about scientific phenomena from (Bogen and) Woodward and in the 2011 Synthese special issue, patterns are a subject of philosophical examination in some further philosophical writings. Dennett (1987; 1991), Haugeland (1993) and Ladyman and Ross (2013) discuss the ontology of patterns; I discuss their results and I present my view in section 4.3.

[1]'Anschauung' at Kant is often translated as 'intuition' or 'contemplation' (see Carus 1892), and the issue of translating the word appropriately is a subject of discussion among Kantians. The reason vor my suggested translation is that *Anschauungen* for are in space and time, which form a sort of inner and intersubjective epistemic framework. They are as similar as possible to what we can *see* from the empirical world, but they are our inner illustration of something in space and time, like triangles in the ideal mathematical sense, but they are not ideas or thoughts. The extension of a body is something that can be an *Anschauung*, but its weight, its smell and its reputation cannot.

[2]I am a structuralist about mathematics (see section 3.1).

Grenander (1970) presents a detailed mathematical explication of the concept of a pattern in scientific contexts with a maximum of generality. His *general pattern theory* is an algebraic approach that is general enough to cover not only examples from very quantitative fields, such as physics, but also patterns from linguistic, historical or social studies, for which no mathematical explications are known. In later works, Grenander and Miller (2007) show the applicability of the general theory by various well known examples of pattern recognition in the well established fields of statistics, communication technology and image recognition.

I present the basic algebraic ideas from the general pattern theory in section 4.4. This theory is of importance for this thesis for two reasons. Firstly, the algebraic approach provides a purely mathematical, and therefore purely structural and precise account of patterns and pattern recognition. This structural explication helps us to grasp a better understanding of these topics. Secondly, I believe that there is not much more to say about the specific philosophical matter of patterns and pattern recognition, after accepting this algebraic explication as an adequate mathematical and philosophical explication; general pattern theory, by it being purely mathematical, provides epistemological and metaphysical implications concerning patterns and pattern recognition. If we deem this mathematical explication of patterns as suitable for our philosophical purposes, then a pattern is ontologically nothing else than a mathematical object, such as the number 4 or the vector space of all polynomials.

However, the algebraic approach of general pattern theory does not provide detailed information of the class of patterns from actual scientific data. To say it more precisely, general pattern theory provides a general structural that is mathematical, explication of scientific patterns (*i.e.* a description of the data, certain operations on the data, a very general definition of the pattern based on these operations), but it remains an open question how to fill some of the proposed algebraic mappings with more concrete criteria that tell us what separates a pattern from mere noise in widely occurring cases of actual scientific data. Grenander and Miller's (2007) later pattern theory provides a series of answers to this question for many widely used pattern detection techniques from different scientific fields and engineering. I chose a route for explicating patterns that is on the one hand not to a large extent adapted to the specific techniques that are used in actual science, but that is, on the other hand, much more informative than the very general notion of patterns from general pattern theory. Based on Shannon's (1948a; 1948b) information theory, Kolmogorov (1965) introduces an algorithmic notion of complexity of data; I use this *Kolmogorov complexity* to define patterns in data and discuss its advantages and disadvantages compared to the solution suggested by general pattern theory.

4.3 Patterns are Mathematical: Discussion and Objections

Section Abstract

Patterns are mathematical objects without any criterion of complexity or repetition necessary to define them. The examples of albinism and long-range dependence illuminate this. Dennett's notion of "real patterns" restricts patterns to human agents, which is inadequate for science with all its technical auxiliaries. Wallace and Dennett imply unjustified ontological roles of patterns, which are in fact epistemic and structural rather than metaphysical. Haugeland confuses patterns with different degrees of awareness for patterns by certain agents.

As already introduced, I understand patterns as mathematical objects such as the number $\sqrt{2}$ or the set $\mathbb{N}$. We face two challenges to strengthen this claim. First, I need to defend this general ontological claim against critics. Second, I need to describe what kind of mathematical objects patterns are; to say it in other words, is there a certain class of mathematical objects, which are the patterns? In this section I aim to address possible concerns against my claim; these concerns are, in particular, brought forward by Dennett (1987; 1991) and Haugeland (1993), who propose a richer ontology concerning patterns. The detailed mathematical explication of patterns follows in the section after this one.

Furthermore, I focus on examples from actual science, but imply that my descriptions and arguments in favour of patterns being mathematical do apply to everyday reasoning in the same way. As already state earlier in the introduction, section 1.2, of this thesis, I claim, on a general level, that everyday reasoning is well described as the *science of laymen*. That is why our explication is not only a matter of interest concerning philosophy of science, but also for epistemology in general. In the following paragraphs, I briefly introduce the critiques of Dennett and Haugeland. I do not agree with their philosophical position, but their role as well-established opponents helps to sharpen the view on the subject.

Due to the simplicity of many patterns in data that correspond to empiric phenomena, we cannot reasonably propose any criterion of minimal *complexity* for patterns in data whatsoever. Recall the example of the melting point of lead: we have a very simple time series of measurement results showing temperatures (*e.g.* 600.59 K, 600.61 K, 600.61 K, 600.60 K). The pattern simply is the melting point of 600.61 K, which is nothing else than a constant rational number, or it is, more general, a certain form of clustering around this number. Due to the lack of knowledge we may not be sure about the *actually* correct pattern in many cases.

There are not many mathematical objects that are less complex, disregarding how exactly one understands complexity of mathematical objects in detail.

What are patterns mathematical more concretely? Measurement results in science can look very different. In the simple case of measuring the melting point of lead there can be a time series of measured temperatures were the lead was detected to melt as given above. In a little bit more artificial approach the results could also be given as a time series of tuples showing arbitrary temperatures next to an measured aggregate state from repeated runs, for example

$$\begin{pmatrix}600.50\\ l\end{pmatrix}, \begin{pmatrix}600.55\\ l\end{pmatrix}, \begin{pmatrix}600.60\\ l\end{pmatrix}, \begin{pmatrix}600.65\\ s\end{pmatrix},$$
$$\begin{pmatrix}600.50\\ l\end{pmatrix}, \begin{pmatrix}600.55\\ l\end{pmatrix}, \begin{pmatrix}600.60\\ s\end{pmatrix}, \begin{pmatrix}600.65\\ s\end{pmatrix}, \ldots$$

whereas the upper scalars denote the measured temperature in degree Kelvin and the lower scalars denote a measured aggregate state (s for solid, l for liquid). For this two simple examples one concrete pattern of the melting point of lead is detected by two mathematically different concrete patterns.

How can we explicate the concrete patterns of albinism? The measurement results for the phenotype could be given by a triple of numbers (x, y, z) with, for instance, $x \in \mathbb{Q}$ indicating a skin tone with 0 being the palest white and 9 the darkest black skin tone, $y \in \mathbb{Q}$ indicating the colour of the hair from purely white to black and z describing the colour of the eyes. Since eye colours are not easily scalable by brightness, the z indicator has to be more thought out. A solution would be

$$z = z_1 + 0.1 \cdot z_2 \in \mathbb{Q}$$

with $z_1 \in \mathbb{N}$ indicating a general colour tone and $z_2 \in \mathbb{Q} \cap [0{,}10[$ the general brightness. In this toy scenario the pattern that corresponds to albinism could be a simple criterion of the data (x_H, y_H, z_H) from an individual denoted with H: $x_H \in [0, 0.5], y_H \in [0, 0.5], z_H \in [11.0, 11.5]$. The pattern in this example is the set

$$\{(r_1, r_2, r_3) \in \mathbb{R}^3 : r_1 \in [0, 0.5], r_2 \in [0, 0.5], r_3 \in [11.0, 11.5]\},$$

but only for this specific measurement procedure. With an alternative measurement procedure the patter that corresponds to albinism has to be adapted, but this newly adapted pattern is still an element of the same concrete pattern.

The other concrete pattern that corresponds to albinism is a certain result of a gene analysis. The form of albinism, in which we are interested in, *OCA1A*, is represented by a defect in the *TYR* gene in the 11th chromosome. Specific genes of organisms can be measured[1] as DNA, RNA or protein sequences. The amino acid sequence of the *TYR* gene can be coded in the form shown by figure 4.1, left side.[2] In this case of an individual with albinism, the pattern can be described mathematically as indicated by the following table (cf. Tomita *et al.* 1989, p. 992–3): [3]

Position	Amino acid in non-mutated gene	Amino acid in mutated gene
311	R	K
312	L	A
314	S	L
315	S	F
316	A	S

Patterns that correspond to phenomena may be not fully discovered by the scientists. Amongst others, Oetting *et al.* (1998), Spritz *et al.* (1997), or Wang *et al.* (2009) specify further mutations that seem to lead to *OCA1A*. In the *Human Gene Mutation Database*[4] a comprehensive list of analysed mutations of the *TYR* for *OCA1A* patients can be found. Thus, in fact the scientific results are not sufficient to fully explicate the pattern that corresponds to the phenomenon of albinism.

The sought-for pattern of *OCA1A* albinism in gene analysis records may be identified by one of many sufficient mutations: many different positions of amino acids in the *TYR* gene may cause albinism, if properly mutated. There is no clarity

[1] In fact, reading out DNA is a biochemical process, which may not properly be described as *measurement* in the sense physicists use this term.

[2] In protein isoform *Iso 1* according to http://www.nextprot.org/db/entry/NX_P14679/sequence

[3] A list of the abbreviated amino acids for completeness:

Amino acid name	abbreviation	Amino acid name	abbreviation
Alanine	A	Leucine	L
Arginine	R	Lysine	K
Asparagine	N	Methionine	M
Aspartic acid	D	Phenylalanine	F
Cysteine	C	Proline	P
Glutamic acid	E	Serine	S
Glutamine	Q	Threonine	T
Glycine	G	Tryptophan	W
Histidine	H	Tyrosine	Y
Isoleucine	I	Valine	V

[4] See http://www.hgmd.cf.ac.uk/ac/gene.php?gene=TYR, the entry for the tyrosinase gene *TYR* in a database for published gene lesions responsible for human inherited disease by the *Institute of Medical Genetics* in Cardiff. The database is openly available to read after registration.

about whether these mutations can be disjoint in the sense that for two individuals with albinism no mutated position in the *TYR* gene of the one individual is mutated in the *TYR* gene of the other individual. The literature presents a list of positive examples of mutations of the *TYR* gene from patients with *OCA1A*, and this known list of mutations is very likely not complete to describe all possible mutations that lead to the phenotype of *OCA1A*.

The selection of the *TYR* gene from the whole DNA already counts as a part of the pattern corresponding to *OCA1A* in the strict sense. We do not know the exact pattern that corresponds to *OCA1A*, but there is scientific evidence that the amino acid sequence of the *TYR* gene is a sufficient set of data to detect it. My explication of phenomena, of patterns and their relation to each other does not depend on the validity of this statement, but only on the reality of a pattern in a set of data that corresponds to albinism. I use this example and the discussion about its biological details to illustrate the aspects of patterns in science I am interested in. These are its explicability in mathematical terms in principle, as well as the lack of scientific knowledge to do this exactly in realistic cases from everyday scientific work.

Various examples can be found to further illustrate that patterns in science that correspond to phenomena are often very uncomplex. This is the case for most patterns that are already explicated mathematically due to the common use of the mathematical language by the science in question. Patterns of this kind include the so called *long-range dependence*, which is a statistical property of a time series of real numbers, but also a pattern in data that corresponds to a phenomenon or, respectively, to a group of phenomena. I outline the definition of long-range dependence for stochastic processes to discuss the general idea of it. The statistical test for it in time series and the statistical estimation of the calibrated parameter $1-\alpha \in]0,1[$, the *intensity*, is a topic of intense mathematical research on its own.[1]

Definition (long-range dependence)

A covariance stationary, real-valued stochastic process $(X_t)_{t\in\mathbb{Z}}$ is called **long-range dependent** (or **persistent** or **strongly dependent** or having **long memory**), if for all $t \in \mathbb{Z}$

$$\lim_{n\to\infty} \frac{\operatorname{cov}(X_t, X_{t+n})}{c \cdot n^{-\alpha}} = 1 \text{ for a } c \in \mathbb{R}_{>0} \text{ and a } \alpha \in \{r \in \mathbb{R} : 0 < r < 1\}.$$

[1] For a discussion from various perspectives see Doukhan, Oppenheim and Taqqu (2003). Several estimation procedures were developed to improve the statistical stability and convergence speed of the estimation. For an overview of estimation methods see also appendix C of Biagini, Øksendal and Zhang (2008).

In other words, a process or a time series is long-range dependent, if the autocovariance decreases hyperbolically with increasing lag above a certain threshold of the lags. As the term already indicates, a prima facie assumption in many time series from repeated measurements along a time-line is the complete vanishing of any autodependence at a sufficient increase of the lag. This would amount to a convergence of the autocovariance to 0 with increasing lag, which is opposed to the criterion given above. The divergence of a long-range dependent stochastic process $(X_t)_{t\in\mathbb{Z}}$ can be seen by flooring the covariance series by the harmonic series, which is already diverging:

$$\sum_{n=1}^{\infty} \operatorname{cov}(X_t, X_{t+n}) \overset{\text{large } \hat{n}}{\approx} \hat{C} + \sum_{n=\hat{n}}^{\infty} cn^{-\alpha} = \hat{C} + c\sum_{n=\hat{n}}^{\infty} \frac{1}{n^{\alpha}} \overset{\alpha\in]0,1[}{>} \hat{C} + c\sum_{n=\hat{n}}^{\infty} \frac{1}{n} = \infty$$

for sufficiently large $\hat{n} \in \mathbb{N}$ and a constant $\hat{C} \in \mathbb{R}$. The diverging autocovariances of long-range dependent stochastic processes or time series is illustrated by figure 4.5.

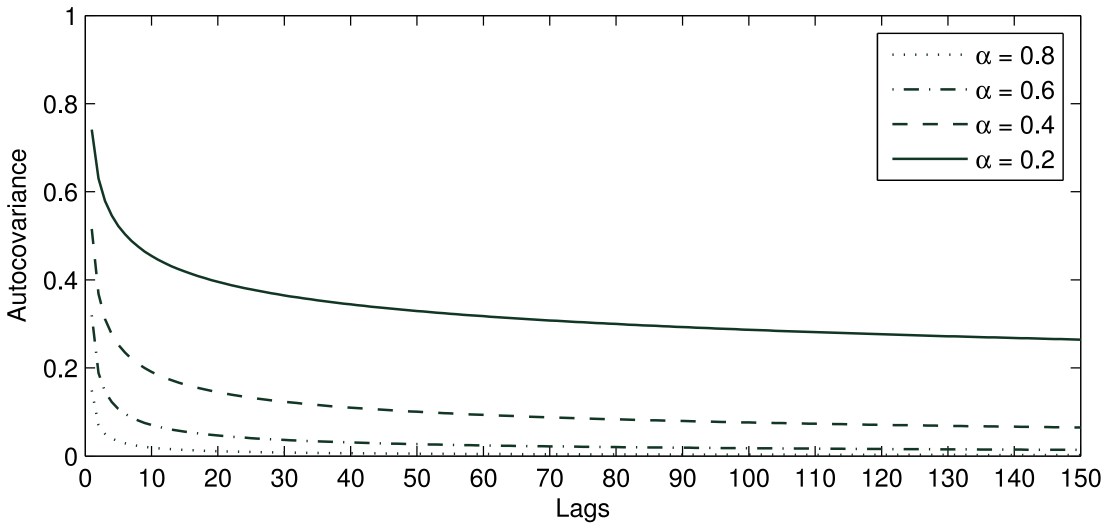


Figure 4.5: The autocovariance of a process $(X_t)_{t\in\mathbb{Z}}$ with long-range dependence for different given intensities $1-\alpha$ and for lags from 1 to 150. The volatility of the process is constant $\sigma_t = 1$ for all $t \in \mathbb{Z}$. The process is perfectly long-range dependent in the sense that the covariance does not converge to the defining property for only large lags $n \in \mathbb{N}$, but fulfils the defining criterion for all lags $n \in \mathbb{N}$.[1]

The pattern of longe-range dependence does not play the role of only a concrete pattern, but also the one of a full description of the phenomenon itself, which is its general pattern. Contrary to many empirical phenomena in science, there are no other pragmatic ways to identify the empirical behaviour of long-range dependence in an empirical system than by the statistical analysis itself. There is nothing

[1]To produce this plot I adapted a script that I wrote for my diploma thesis (2012).

like a phenotype or any other concrete pattern of long-range dependence that is accessible to human beings with the naked eye. Long-range dependence is a case in which the general pattern consists of only one concrete pattern. However, long-range dependence should be explained by scientific theories and is a phenomenon that occurs in very different empirical data.

Long-range dependence is identified in historical measurement data of the Agyptian Nile flood from A.D. 662 to 1469 (cf. Hurst 1951) or hydrological time series in general (cf. Mandelbrot 1968 and Klemeš 1974). It is discussed as a phenomenon of financial asset price time series or its in absolute or squared financial asset returns (cf. Mandelbrot 1969 and 1971, Lobato and Savin 1997, Lux 1996, and Grau-Carles 2000). Furthermore, long-range dependence is a phenomenon of traffic in data networks, such as the World Wide Web (cf. Crovella and Bestavros 1997) or in Ethernets in general (cf. Leland *et al.* 1994).[1]

I want to discuss two noticeable properties of the phenomenon of long-range dependence. First, it is already expressed in its form of a statistical pattern. Second, this pattern can be found in scientific data from various very different and not related parts of the empirical world. Due to this first property, it is really easy to define the concrete and the general pattern of long-range dependence. Both of them are fully given by the mathematical definition of long-range dependence. But do we have to distinguish between the occurrences of long-range dependence in this very different parts of the empirical world? Is long-range dependence in water levels a different phenomenon than long-range dependence in the volatility of daily asset returns in the financial market?

My answer to this question is that, even if the pattern that corresponds to these phenomena is identical, the phenomena themselves are not. The empirical fact (or “feature”) that the Nile floods are long-range dependent is not the same empirical fact that the data traffic in the internet is long-range dependent. It may be possible to formulate theories on a micro level that describe the internet and the meteorological circumstances influencing the Nile floods by the same theoretical propositions. In this case, not only the statistical pattern of long-range dependence could be identically detected in the data of these two different empirical systems, but also the reasons in the systems that lead to long-range dependence could be theoretically identical. However, the pattern of long-range dependence is identical in all these cases, but it corresponds to different phenomena, which are selected

[1] As Willinger *et al.* (2003) point out, the phenomenon of long-range dependence in network traffic is caused by user and application characteristics rather than by the architecture of the network. The scientific discussion of it does not revolve around only the popular—how they call it—“black box” approach of statistical time series analysis, which is the mere statistical test for the occurrence of a statistical phenomenon. Instead the physical mechanisms of the broader networking context is scrutinised to not only identify the occurrence of long-range dependence, but also explain it on a single agent level.

and defined with the use of a whole body of further background assumptions, as I elaborate on in chapter 5.

Let us consider—purely hypothetically—a theory about the two systems can explain the long-range dependence to be caused by a certain seasonal clustering of larger amounts of flow in the system. The record of the Nile flood may be strongly influenced by some meteorological circumstances in the rainy months of the summer; in a theoretically identical way the data flow in the internet may be strongly influenced by the users' surfing behaviour directly after the office ours in the United States and in Europe. In both cases the theory may explain the long-range dependence in the data by these theoretical influences. Even in this case, were not only the phenomenon, but also its theoretical explanation in a smaller scale is identical, the phenomena are still different. The reason for this is that the theoretical explanation is not identical to the empirical part of the world under investigation by the science that needs to be explained. The application of the same theory does not involve the explanandum to be the same scientific phenomenon, as I elaborate on in chapter 5.

Objection: Dennett's "Real Patterns" for Humans' Pragmatic Concerns

Dennett[1] distinguishes between patterns in general and—how he famously coined the terminology—"real patterns" (cf. Dennett 1991), which are, in his view, discernible by intentional states, such as beliefs or desires. *Real patterns* have an extra, real, metaphysical status. The following passage indicates his view:

> I claim that the intentional stance provides a vantage point for discerning similarly useful patterns. These patterns are objective – they are *there* to be detected – but from our point of view they are not *out there* entirely independent of us, since they are patterns composed partly of our own "subjective" reactions to what is out there; they are the patterns made to order our narcissistic concerns (Akins 1986). It is easy for us, constituted as we are, to perceive the patterns that are visible from the intentional stance – and only from that stance.(...) Martians might find it extremely difficult, but they can aspire to know the regularities that are second nature to us just as we can aspire to know the world of the spider or the fish. So I am a sort of realist. (Dennett 1987, p. 39–40)

I comment this quote. The *intentional stance* is a central idea in Dennett's contributions. It refers to an—as he claims—metaphysical stance about intentional

[1] My thanks go to Max Kistler, who organised the conference *New Trends Metaphysics of Science* in Paris in December 2015. He brought to my attention that the claim of patterns being purely mathematical is attacked by Dennett and this attack received a significant amount of attention among philosophers.

states, such as beliefs or desires. It roughly says that such a state is a "perfectly objective phenomenon" (Dennett 1987, p. 15) like "the question of whether a person is infected with a particular virus" (p. 14) and unlike "the question of whether a person is immoral, or has style, or talent, or would make a good wife" (p. 15). This objectivity is what makes Dennett "a sort of realist", whereas this *realism* incorporates human intentional states into the ontology.

However, not all possible intentional states are allowed according to the intentional stance; agents have to be rational and their "behavior is reliably and voluminously predictable" (p. 15) and they are therefore "true believers" (p. 15). Such as we use medical procedures to detect a virus in an organism, we can detect an agent's set of beliefs by investigating his behaviour. Since the agents are nevertheless bounded by the cognitive and sensorial equipment of a human being, Dennett's "similarly useful patterns", which can be detected in the world, are only a certain selection of all the possible patterns, as I define them later on in this text. These patterns may be detectable by algorithms, artificial intelligences or—a little bit outdated—"Martians".

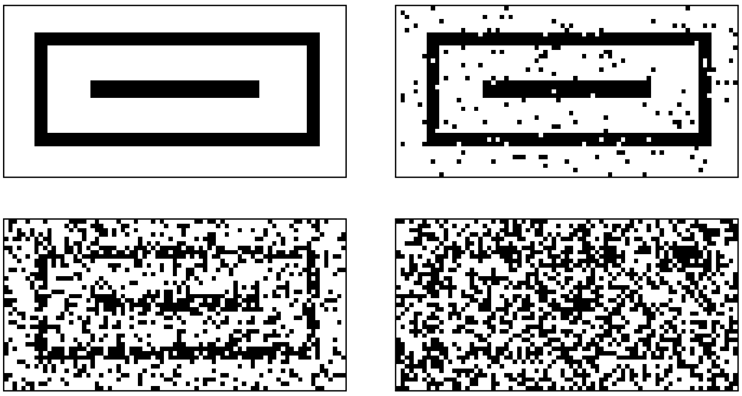


Figure 4.6: Four simple visualisations of sets of data: $80 \times 40 = 3200$ binary information per image in a specific order. We can stipulate the upper left one as an ideal pattern; this choice is *epistemically natural* to a human, since this pattern is easy to remember and describe. Based on this ideal pattern, the other three images show realisations of the pattern plus a noise with a level of 0.1 (upper right), 0.5 (lower left) and 0.8 (lower right). Here, a level of noise of $\rho \in [0,1]$ means that $(\rho \cdot 100)\%$ of the points in the image are white or black by random and the rest are coloured according to the ideal pattern.

The metaphysical difference between a *real pattern* according to the intentional stance and any other pattern in a set of data is given by the fact that they are "composed partly of our own 'subjective' reactions to what is out there". Akins' (1986) referenced PhD thesis *On Piranhas, Narcissism, and Mental Representation* is a fruitful neuroscientific perspective for our endeavour. This reference highlights Dennett's guiding principle to put actual human cognition much more into the centre of the ontological investigations than I am willing to do. We read:

> This dissertation is motivated by the following question: Is the portrayal of mind/brain processes as representations–as entities that in some sense reflect, correspond with, or symbolize the world–particularly apt? Through detailed examples from the neuroscientific literature, with an emphasis on sensory processing, I argue that this way of viewing brain functioning is typically misleading. It depicts neural functioning as a bipartite process: first the production of a set of neural "calibrational" states with properties in the world, and then their interpretation by "higher" functions. On the contrary, even at the transducer level, sensory organs cannot be characterized as relay mechanisms for the brute facts. The form and content of all information gleaned about the external world conforms to the particular needs, hence neural functions, of the organism. Evolution, it seems, is not concerned with "the truth", but only with that which proves necessary or expedient. Relaxing the grip of the representational metaphor, I argue, affords us the means to reconstrue or even dissolve some standard philosophical questions about content and intentionality. (...) (Akins 1986, abstract)

Her conclusion of mind processes being guided by rather pragmatic "narcissistic concerns" is not surprising whatsoever *prima facie*—at least to me. A little example similar to Dennett's[1] (1991) about visual patterns illuminates this point even more. Figure 4.6 shows four visualised sets of data. If we ask laymen in regard to statistics and data analysis where he can find a pattern, then most would answer that the upper left one shows the pattern and the other three show it, but distorted by different levels of noise. However, this choice of a pattern is rather arbitrary and guided by the sensory-cognitive capabilities of human beings in general and laymen (regarding *e.g.* computer imagery) in particular. On the contrary, the bottom left image may show the exact pattern of interest for a certain analysis of the data, and there is no need apart from the sensory-cognitive capabilities of human beings to prohibit this choice.

Akins' depiction, which leads Dennett to his conclusion about "real patterns", is a neurologically derived version of neo-Kantianism. Kant (under the label of *transcendental aesthetics*) already provided a description how all our perceptions are framed by our distinct human cognitive framework.

Disregarding any lack of specific focus on ontological questions, it is helpful to elaborate on our arguments why any ontological or other philosophical distinction between patterns and "real patterns" in the described sense is in principle futile to describe patterns in science—Dennett's texts also have ambiguous parts concerning what his final position on the subject is.[2]

---

[1] Dennett's (1987; 1991) favourite example is the *Game of Life*, a simple zero-player game that produces patterns of two binary states on a surface; it starts with a certain state and in deterministic subsequent steps the binary states of the single pixels change according to a very simple transition rule. Some very distinctive patterns, such as the *gliders* occur that are stable along the next steps of the Game. The Wikipedia entry *Conway's Game of Life* (English) is very informative about this *cellular automaton*.

[2] Simply compare the following two passages with the one already cited above (p. 118):

In the following paragraphs I explain why the distinction is futile for any characterisation of patterns in science. According to Dennett, the criteria that distinguish a "real pattern" from any other pattern are determined by the pragmatic needs of the human organism; obviously those criteria make different "real patterns" for different individuals, such as chess players, experimental physicists, decorators or any other group of trained persons. These differences do not occur only among different contemporary groups, but also in historical contexts. Newton deemed other patterns in experimental data interesting than modern statistical physicists. The same holds for painters like Rubens and Cy Twombly.

Thus, it is hard to define an agent-independent class of "real patterns", but is it impossible *in principle*? If someone has a perfect scientific knowledge about how the human brain works, can she then define this class that describes all the patterns that a human being can detect in principle? Dennett seemingly wants to affirm this question with the aim to secure his point by stipulating for his intentional stance only rational agents who are "true believers". His strategy is then to sort out all patterns that do not contribute to any rationally and somehow personally beneficial strategy in life. But this strategy does not work, since scientific and aesthetic criteria of interest for patterns change drastically from context to context. Even with a perfect knowledge about how the human brain works, there is no convincing criterion about a pattern that excludes it from being interesting for a rational strategy of a human being. The only acceptable claim, if we include scientific applications of pattern detection by human agents into the notion of patterns, is to refer to epistemic restrictions (*e.g.* patterns that are too complex in some sense to be detected by human agents), but this criterion is rather arbitrary bounded by physical brain power, which is exactly what Dennett's intentional stance opposes.

> My view is, I insist, a *sort* of realism, since I maintain that the patterns the Martians miss [but we humans or at least some of us recognise] are really, objectively there to be noticed or overlooked. How could the Martians, who "know everything" about the physical events in our world, miss these patterns? What could it mean to say that some patterns, while objectively there, are visible only from one point of view? (Dennett 1987, p. 37)

but also

> A pattern exists in some data—is real—if *there is* a description of the data that is more efficient than the bit map, whether or not anyone can concoct it. (Dennett 1991, p. 34)

Ladyman, Ross and Collier (2007) put specific emphasis on Dennett's presentation regarding metaphysical implications, too:

> However, in a now-classic paper 'Real Patterns' (RP; 1991), he [Dennett] emerged from this neutrality [about metaphysics] to frame his view of mind in the context of what Haugeland (1993) rightly regards as a distinctive metaphysical thesis. According to RP, the utility of the intentional stance is a special case of the utility of scale-relative perspectives in general in science, and expresses a fact about the way in which reality is organized—that is to say, a metaphysical fact. (p. 199)

Scientific development repeatedly produces new patterns of interest in various sets of data (*e.g.* long-range dependence).

Furthermore, it seems futile to draw a strict demarcation line between genuinely human pattern detection and procedures that make use of technical or other auxiliary equipment or procedures. The pattern of a chess board is easy-to-detect by the naked eye, but this does not apply to patterns in thousands of pictures from the Hubble telescope, which are analysed by either powerful computers or a group of scientists or laymen.

Another strategy to defend Dennett's metaphysical implications is to stipulate that there must be actual human agents, who actually recognised this pattern to make it a "real pattern". But this defence contradicts with the idealisation of dealing with rational agents and true believers, as stated above. All in all "real pattern" turns out to be a vague notion. Therefore, the notion is not very useful for an ontological or epistemological classification. This is one reason why I want to drop Dennett's notion.

A more compelling reason to reject the use of Dennett's notion for my purpose is the following. Even if the convictions of a human scientist and the human scientific communities play a crucial role in science, scientific pattern detection undergoes a constant process of technical enhancement. Scientists do not use mathematical algorithms to only detect certain patterns, they can use artificial intelligences to define relevant patterns in the first place. The detection of cognitively accessible patterns is replaced with patterns that are defined by machines in which the scientists *trust*. This trust is the same trust that the economist may have in a human expert in statistics, who provides him certain results. Under this view, the patterns that serve our human "narcissistic concerns" cannot play a prominent role for an explication of patterns in science.

Another important remark regarding my view on patterns and the philosophy of mind is that my view is by no means hostile to an ontological acceptance of *consciousness*. I claim that patterns, pattern detection and scientific inference have nothing to do with consciousness. We do not have any indication that complex *cognitive* tasks require consciousness and the optimistic induction (see 3.3) rest on empirical evidence of the contrary.

Objection: Wallace's Patterns in Science as Universals

Wallace (2003) develops a view about indefiniteness in quantum mechanics according to which macroscopic physical objects should be understood in terms of Dennett's patterns rather than by the common physicists formalism of quantum mechanics. I do not elaborate on this specific view, which is not in the direct scope of this survey, but I want to focus on his usage of the concept of a pattern; he refers directly to Dennett's texts but adds some further ontological aspects to it. They help us to get further insights into Dennett's view and to sharpen my suggested view about patterns.

We start with a quote in which Wallace introduces his notion of patterns that he named "in recognition of a very similar view proposed by Dennett (1991)":

> *Dennett's criterion:* A macro-object is a pattern, and the existence of a pattern as a real thing depends on the usefulness—in particular, the explanatory power and predictive reliability—of theories which admit that pattern in their ontology. (Wallace 2003, p. 93)

For him, a pattern is a "macro-object" and therefore also some part of the empirical world (assuming he is not a strict nominalist, which he is not). In his text he exemplifies this notion of a pattern by a tiger (sic!), which is *prima facie* "to be understood as a pattern or structure in the physical state." (p. 92) But do physical states "admit that pattern [of a tiger] in their ontology"? No, this ontology is rather based on particles, but the science and language of "zoology and evolutionary adaptationism" (p. 93) does admit tigers and therefore "a tiger is any pattern which behaves as a tiger" (p. 93) according to this language.

He is a *structural realist*, whereas he uses 'structure' and 'pattern' interchangable. 'Structural realism' denotes the ontological orientation in philosophy of science according to which we should, in the context of scientific theories, enter into ontological commitments on the level of structures and not on the level of individuals. Worrall (1989) revitalised this general idea for contemporary philosophy and highlights it with the help of the example of Fresnel's[1] elastic solid ether theory and the transition to the theory of electromagnetic fields from Maxwell[2]; according to Worrall there "was continuity or accumulation in the shift [from Fresnel's theory to Maxwell's], but the continuity is one of form or structure, not of content"

[1] The French physicist Augustin-Jean Fresnel (1788–1827) explained the propagation of light, which can be observed in transparent matters, but also in empty spaces, by the existence of the ether. In his theory, the ether carries light waves and has to fill the whole physical space. The propagation speed of light indicates for the ether the properties of an elastic solid material.

[2] The British physicist James Clerk Maxwell (1831–1879) interpreted light as an electromagnetic radiation, because the electromagnetic radiation he investigated propagates through empty space with the speed of light, too. However, Fresnel's empirical laws of light propagation remain intact under this new interpretation of light.

(p. 117). The reason for this interpretation is that the existence of the ether as the matter in which light travels through space, could be fasified by experiments, but the relations that Maxwell inferred under the assumption of the ether hold empirically in electromagnetic fields.[1]

For Wallace a pattern is something that is usually called a *universal* in ancient and medieval metaphysical debates; examples of universals include tigers, the colour red and triangles. Universals are manifested by, at least, a name for them in everyday or a specialised language, which is, in essence, the criterion that Wallace expresses in the given quote by "the usefulness (...) of theories which admit that pattern in their ontology". We have a word for tigers because it is useful for grasping the animal kingdom. For Dennett real patterns are manifested by the *intentional stance*, which does not refer to terms in a language, but to something very similar that has its place in human minds.

Wallace's motivation for his realists' position is a version of the *no-miracle* argument, which in its most common version roughly states that our best scientific theories about the empirical world must at least be approximately true and the terms used by them must refer to actual objects by virtue of their predictive success and the functioning applications in engineering.[2] But in Wallace's version it is not the predictive success that has to be explained by something else than a miracle, it is the *explanatory* success of our scientific theories:

> Why is it reasonable to claim, in examples like these [of tigers], that higher-level descriptions [in the language of zoology] are *explanatorily* more powerful than lower-level ones [in the language of microphysical states]? In other words, granted that a prediction from microphysics is in practice impossible, if we had such a prediction why would not it count as a good explanation? To some extent I am inclined to say that this is just obvious–anyone who really believes that a description of the trajectories followed by the molecular constituents of a tiger explains why that tiger eats a deer means something very different by 'explanation'. But possibly a more satisfying reason is that the higher-level theory to some extent 'floats free' of the lower-level one, in the sense that it does not care how its patterns are instantiated provided that they are instantiated. (Hence a zoological account of tigers requires us to assume that they are carnivorous, have certain strengths and weaknesses, and so on, but does not care what their internal makeup is.) So an explanation in terms of the lower-level theory contains an enormous amount of

[1] Structural realism is not an entirely new view. Among others, Poincaré (1902; 1905), Russell (1927) and Cassirer (1936) expressed in crucial regards similar ideas. Structural realism is often subclassified into the *ontic* and the *epistemic* version; ontic structural realism expresses the view that the world is in fact composed of structures and not of individuals, epistemic structural realism expresses the view that our conceptualisation of the world is structural for epistemic reasons. Ladyman (2014) provides a survey on structural realism in which he focuses on the contemporary discussion.

[2] I refer to, mainly, Putnam's (1975, p. 73) often referred to introduction of the no-miracle argument. Find an encyclopaedic discussion at Chakravartty (2011, sect. 1.2).

> extraneous noise which is irrelevant to a description in terms of higher-level patterns. (Wallace 2003, p. 94)

To sum up his view: since our best explanations of our surrounding world introduce tigers, this universal must be real (no-miracle) and since our theories express patterns—or in terms of Ladyman and Ross (2007): *structures*—these real things must be patterns (ontic structuralism).

Structural realism is, like other ontological positions, to a certain degree a metaphysical commitment one can believe in or not rather than a debatable position. However, why should we reject Wallace's realists notion of a pattern? The reason is that this notion simply does not describe what we refer to, when we talk about patterns in the everyday language, as well as in almost every language of a scientific field. With the term 'patterns' I refer to examples of the sort illustrated by figures 4.2 (p. 108) and 4.3 (p. 109), and not to tigers or the colour red.

What Wallace calls 'patterns' are in fact universals, and some of them stand in a strong relation to patterns in data in the same way as (Bogen and) Woodward assume patterns to stand in a strong relation to phenomena. Referring to universals with 'patterns' is motivated by the stipulated identity of those with patterns in data according to a scientific theory in the same way (Bogen and) Woodward stipulate an identity between pattern and phenomena. However, both use a descriptively inadequate notion of patterns regarding the actual use of the notion by scientists. In Wallace's case, this does not have to count as a mistake, but may be misleading for a reader. In (Bogen and) Woodward's case, it is a mistake, as I discuss in section 4.1 with help of the example of albinism.

How do Wallace's ideas help to sharpen our view on patterns in data? Patterns in data have to be strictly distinguished from structures in the sense of ontic structuralism. But how exactly? Patterns in data are certain mathematical (*i.e.* structural) properties of some data, which are mathematised reports of an observation. Wallace's patterns are sub-structures or positions in the complex structure, which is a certain subset of our body of knowledge (of *e.g.* zoology). It is correct to say that a picture of a tiger shows a tiger by means of the visual pattern that corresponds to a tiger and can be detected in the picture. However, this pattern *is not* the tiger and we may not know what a tiger really is (*i.e.* scientific anti-realism) despite all sorts of descriptions we come up about it.

How does this distinction relate to the notions of *concrete* and *general* patterns from section 4.1? As a reminder, the phenomenon of albinism corresponds to the general pattern that includes the concrete patterns of its genotype and its phenotype. General patterns are classes of concrete patterns that are defined by the actual occurrences of the phenomenon and we may not know every concrete pattern that corresponds to albinism. However, according to Wallace's account we know every structural property of the tiger, since a tiger is an object that is completely described by our language with which we refer to a tiger. This leads to an important difference between phenomena/(general) patterns on the one side and Wallace's suggested tiger structure as "patterns" on the other side: we do not know a general pattern of a phenomenon due to newly developed forms of data and undiscovered properties of the phenomenon, but we know everything about a Wallace pattern since it is *defined* by our reference to it.

Haugeland's Objection: Patterns are more than a Compositions of Bits or Elements

For Haugeland[1] (1993) patterns, and any ontological classification of them, do not depend on the "special case" (p. 267) of intentional states, as Dennett suggests. In Haugeland's view, all patterns that are in any manner recognisable by a human agent, count as patterns from the same ontological and epistemological class.

What all counts as a pattern for Haugeland? The following quote shows how broad his notion is:

> For instance, when I recognize the faces of my friends, or the expressions on their faces, or the genre of a book, there are no particular bits or other elements that these are patterns of. A delighted smile is not a pattern of epidermal cells, still less of pixels or light waves; if anything, it's a concurrence of cheek lift and brow movement, of lip shape and eye sparkle. But these are no more antecedently determinate than smiles themselves, perhaps less so. Smiles, as the definition suggests, are what they are because we recognize them to be, and not the other way round. (p. 274)

If a smile or face or book genre counts as a pattern, then this notion seems to be a similar one to Wallace's notion. Again, in classical ontological terms, a smile or a tiger would be a *universal*. For Haugeland, a pattern is a pattern by it being recognisable for the human agent. It is central to his view that "there are no particular bits or other elements that these are patterns of", but this is the crucial mistake of his view. Yes, we are not *aware* what structural recognition of a human face makes us categorise it as smiling. However, one could explicate a smile in purely mathematical (*i.e.* structural) terms. In fact, modern facial recognition software does exactly that.[2]

He goes on regarding the claim that patterns are something merely structural (*i.e.* mathematical):

[1] My thanks go to Philipp Haueis, who brought to my attention that Haugeland criticises Dennetts depiction of the relation between intentional states and patterns, as well as he attacks the claim of patterns being purely mathematical.

[2] An example of software for the recognition of facial expressions is *Sightcorp Crowdsight SDK*.

> In the meantime, requiring determinate prespecification of the bits or elements, as the mathematical definition does, can be a philosophical embarrassment, in more than one way. First, many relevant patters—conspicuously including the behavioral patterns that support intentional interpretation—do not seem to be made up of well-defined bits or elements. Just which causal commerce with the environment amounts to perception and action is by no means specifiable in advance, nor can it be precisely delineated in any case. Second, the account of patterns as orderly arrangements of predetermined elements is an invitation to metaphysical reductionism: the thesis that these patterns are 'nothing but' their elements arranged. Clearly, however, (whatever else one thinks about it) this runs counter to Dennett's motivating insight that 'real patterns' might be of distinctive ontological status and interest. Third, if (in spite of all the foregoing) an attempt were made to merge the two notions of pattern, such that recognizable patterns *must* at the same time be arrangements of prior elements, then, arguably, their recognizability would have to be *via* prior recognition of those elements; and that would be a version of epistemic foundationalism. (p. 275, original emphasis)

In the following, I want to carefully reply to the three listed concerns.

Regarding the first point, *all* patterns are structural and can be explicated as such. All of the sensorial input that humans are capable of perceiving, can be fully described by images, waves spectrums or other structural forms. However, the more relevant aspect of the first point seems to be that it is not easy to mathematically explicate "in advance" what qualifies as a pattern by being relevant for human's "perception and action". Obviously, our well-developed skills in facial recognition serve our daily social routines, but I see no reason, why all of this could not be mathematically explicated. To illuminate my point, I, again, suggest some analogies. A human Go player recognises specific patterns on a Go board during a game. A competitive Go computer recognises patterns of no less complexity. When IBM's Watson wins a match of *Jeopardy!*, then this is a strong indication that Watson is able to recognise the relevant linguistic and semantic patterns to a similar degree than the competing humans can. And we can decompile the exact mathematical explication of the pattern detection routine and the rules why these patterns are relevant from the computer. I elaborate more on the argument that a computer performed task exemplifies its mathematical explicability in chapter 3, sections 3.2 and 3.3.

Regarding the second point, I obviously endorse a metaphysical and epistemical reductionism in the restricted area of explaining patterns, pattern recognition and inferences in science. This reductionism, on the one hand, is consistent with scientific practice that makes vast use of automation and computers, and on the other hand, promotes an ontology that is more sparse than Dennett's or Haugeland's accounts imply. And if this would be a foremostly ontological discussion (which it is not, since for me it is a discussion about the explication of patterns in science) than we would need to take the merits of a sparse ontology into consideration.[1]

The third point of concern is very misleading from more than one perspective. As described earlier in his text and cited above the last quote, Haugeland refers to patterns not only in the way how humans would explicate them, but also how they actually occur to us. To explicate a smile on the basis of the movements of individual facial muscles is "epistemic foundationalism" for him, since we recognise the smile directly, without us analysing the muscle movements before the recognition. However, as mentioned above, it is irrelevant to an explication of a pattern, whether we are *aware* of how we infer it as a structure. The crucial difference between his and my position is that I want to talk about what a pattern *is* and not how it is for a human being's lay awareness. Even a human autistic person, who may have trouble with reading facial impressions, can give us a good description of a smile by the individual muscle movements of the smiling person. Furthermore, there is no problem in mathematically explicating a smile in terms of image recognition software.

Overall, Haugeland and also Dennett, provide an important insight into the difference of an explication of patterns in science on the one hand and pattern awareness of an untrained person on the other hand. Only the first notion is of relevance for this thesis and should not be confused. Even if one has a very friendly position towards the philosophical fruitfulness of philosophy of mind, a person's individual gift, neurological disorders, extreme training, technical aids or artificial cognition enhancement make any narrow class of all "real patterns" notoriously incomplete. That is why, even on the background of philosophy of mind, the mathematical approach of general pattern theory to explicate the notion of a pattern is more convincing than any other approach. This will be discussed in section 4.4 below.

Objection: the Optimistic Induction does not Provide a Suitable Comparison with Scientific Tasks

An opponent of my views may claim that, even if AIs show levels of performances above any human agent in games like chess, Go and poker, this still does not justify a view according to which scientific inferences can be made by AIs without actual or even only possible human epistemic *supervision*. These opponents may hold the view that science is something substantially different than these games.

[1] Quine (1948) recommends sparseness as a relevant criterion for ontologies with analogies to William of Ockham (*Ockham's Razor*). However, he uses this criterion to make the case to exclude abstract entities, whereas my ante rem structuralism about mathematical objects or propositions reduces empirical objects from everyday experiences (*e.g.* patterns) and thoughts to mathematics.

My answer to this objection is that at least one crucial criterion for the acceptance of a theory as a scientific and also scientifically valid theory is its success as empirical *predictor*. Even if our notion of *scientific explanation* is restricted to explanations for human agents, the reference to predictability still allows for theories that might not be accessible to a human mind.

We use technical aids like computers to solve mathematical calculations in science for decades. *Monte Carlo* methods are widely used in the social sciences to test agent level properties on a macro level. The mental games of chess, Go and poker provide a foreshadowing empirical study on what AIs can provide for the solution of a very specific scientific question or broader scientific fields in future decades and centuries. Overall, there is no need to exclude AIs in the epistemically and socially vastly complex *game of science* if, for instance, neural networks provide successful solutions with regard to empirical predictability.

More descriptively, the following example helps to support my view. The rocket company SpaceX was able to significantly increase the thrust of their rocket engines in comparison to other rockets.[1] It is important to note that rocket engineering is a field that very likely attracted many highly motivated and gifted engineers who put intense effort into optimisation of rocket engines over several decades. Therefore the field that can be used to showcase roughly the results of the best human engineering efforts. Due to the importance of propellant flows and chamber pressures, fluid mechanics plays a crucial role in the theoretical considerations of this topic. However, the actual flow of the propellant is extremely complex and rocket engineers in earlier decades depended on rough models and estimations. Modern GPU based simulations helped to significantly increase SpaceX's engine performance. (cf. Lichtl 2015)

How can such an example from engineering be compared to scientific models or goals? The reason is that GPU based simulations helped to increase the engine's performance in a *testable* and *predictable* manner. Different than with examples of simulations from social sciences (*e.g.* economics), the example from SpaceX exemplifies superiority of model calculations that can be conducted only with computational aid. But are these GPU based simulations not mere applications of well known differential equations from fluid mechanics? Yes, but if we, as a toy example, define the scientific goal ⌜describe the ideal rocket engine⌝ than the computer aided model, which manifests how the chamber and tubes should be formed, provides an empirically significantly more adequate model than anything what could be expressed based on mere human cognitive capabilities.

[1] Rounded thrust-to-weight ratio of selected rocket engines: Space Shuttle's RS-25 (54), Saturn V's Rocketdyne F-1 (82), Ariane 5's Vulcain (84), Soyuz-2-1v's NK-33 (136), SpaceX's Merlin 1D FT (200), SpaceX's Raptor (estimated ≫200, still in development). Find a detailed list of references at the English Wikipedia entry *Comparison of orbital rocket engines* (2018).

Here is a list of well-known scientific problems that are easy to describe and in a similar manner too complex to be solved without computer aided help, whereas microscopic properties of parts of a larger empirical system are well understood by more simple models that are cognitively comprehensible by a human agent:

- Determining what setup of DNA, RNA, proteins, cytoplasma, nourishment and possible other influences causes an arbitrarily chosen phenotype (*e.g.* being born with exactly 100 upper eyelashes) in a human being without the possibility of analysing a sufficient number of affected individuals (as is the case for *e.g.* albinism). (cf. Weiß 2009) This task requires a super-human level of model complexity.
- The exact mapping of mental states and thoughts to brain processes down to the sub-cell level seems to be a task of super-human complexity.
- Financial market models with the aim to provide detailed predictions based on available information seem to exemplify super-human complexity, due to the vast amount of psychological agents, information, politics and regulation that influence financial market activity.
- The common and extensive use of idealisation and approximation in models of physics in general hint to the fact that physical reality is too complex to be grasped by human's scientific models—however, human's physics with this methodology of phenomenon isolation is one of the most successful fields regarding the criterion of predictability.

Given the empirical examples from the optimistic induction (3.3) and how theoretically non-trivial games like Go and poker are, it is reasonable to assume that the above listed problems can in parts be solved by future AIs with empirically predictive success but without cognitive access for human agents.

Can a Pattern be Arbitrarily Complex?

If a pattern is an arbitrary mathematical property of a set of data and a human agent must somehow grasp it, according to some accounts from the philosophy of mind, then is there an upper bound of the complexity (according to any reasonable notion of it) that a pattern can have? The reason for this suggestion is that a human brain has limited capabilities.

A closer look to our human cognitive capabilities regarding scientific pattern recognition reveals that our limits are surprisingly low. When it comes to numerical tasks, a very rudimentary calculator can vastly outperform a human in a task like writing down or recognise the multiplicative series of three (*i.e.* 3, 9, 27, 81, 243, ...). However, a pattern like this, and many other vastly more complex ones, occur in various actual scientific applications. On the other hand, humans are of course specialised to recognise certain patterns very good (*e.g.* facial impressions; distinguishing objects seen from various different angles and lighting conditions).

The conviction of the scientists in the believability of a certain hypothesis is a major goal of scientific endeavours. How could a human scientist be convinced in a hypothesis if the relevant reasoning makes use of pattern recognition procedures that are not accessible to a human agent? As stated earlier, he trusts in experts and machines that do the work of pattern recognition for him. At least an expert has trust in the machine and the scientists *lends* his trust from the machine expert (transitively, so to speak).

That is why, at least until now, there is no upper complexity bound for patterns necessary. Future scenarios, in which artificially intelligent non-human agents use vastly more cognitive resources to formulate theories and patterns that humans may not able to understand, demand further analysis regarding the mentioned lending of trust. But such an analysis depends on the actual scientific involvement of these non-human agents that we cannot assess today.

4.4 Patterns are Mathematical: Explication

Section Abstract

Modern information theory provides a theoretical framework for data and its applicability on computer hardware demonstrates its adequateness for this explicative role. The notion of information entropy illuminates how patterns are always discussed on the background of an information space. The notion of Kolmogorov complexity illuminates that an important property of a pattern is that it should be constructable for epistemic reasons. General pattern theory is a sufficiently general algebraic approach to actually define the class of all patterns.

In the former sections of this chapter I discussed how patterns should roughly be classified with regard to their relevance for phenomena into concrete and general patterns (4.1), why there is historically a philosophical debate about the notion of patterns (4.2) and why mathematics is all what we need to sufficiently explicate the notion for pragmatic, epistemological and metaphysical needs (4.3). In this final section I present an actual mathematical solution to the problem of the general explication of the notion of a pattern.

The route to the mathematical solution can be outlined as follows. First, we take a look into mathematical information theory and the notion of *complexity* to define a pattern in the broadest and conceptually simplest way. The problem with this view is that we gain a general notion of patterns without any further knowledge about a specific pattern. This notion is *not constructive* in a somewhat stricter sense of constructivism than common in the philosophy of mathematics. To say it in more concrete terms, according to this first explication of the notion of a pattern, a set of data may show a pattern to a certain statistical degree, but there is no way to find this out, because the pattern, according to this first notion, does not come with a description to *test* it in data or to *construct* it mathematically. That is why, secondly, I introduce general pattern theory as an approach to constructively define patterns without any unnecessary restrictions.

I prefer the constructive approach over the adaption of the notion of complexity due to its epistemological merits. Even if we do not restrict the agents of our epistemology to human agents, every agent (*e.g.* artificial intelligences; aliens) has to reason by constructive inferences.—It is another question whether a human agent is able to reenact these inferences or not.

A Simple but Unconstructive Approach: Complexity

It seems natural to appeal to established mathematical theory for the aim to further specify what a pattern in scientific data is. Since data can occur in various mathematical forms (from images; sound waves; via measurements *etc.*) the main obstacle to a mathematical specification of patterns in data is the necessary generality of such a specification.

A common philosophical idea for a most general structural definition is to make use of the theory of a logic. But for the sake of our endeavour, I do not want to specify relations between non-mathematical objects, such as propositions or *natural kinds*; my aim is to specify what property of a set of data makes it to show a pattern and what a pattern mathematically is. Fortunately, a fleshed out theory is available with mathematical theory that seems suitable for my endeavour: the so-called *information theory.* Ladyman and Ross (2013) refer to this solution as “mere patterns”:

> Mere patterns—stable but nonredundant relationships in data—are distinguished from ‘real’ patterns [in Dennett’s sense] by appeal to mathematical information theory. A pattern is redundant, and not an ultimately sound object of scientific generalization or naturalized ontology, if it is generated by a pattern of greater computational power (lower logical depth). (p. 108)

To me this description of patterns as “stable but nonredundant relationships in data” and the reference to “greater computational power” is not sufficiently precise. In my view, a thorough introduction into information theory and its notion of complexity is necessary.

In the following, I briefly introduce the field and explain what part of it is of specific use to us. I also mention other cornerstones of the discussion of data processing in the field to highlight why the approach that I chose is the most useful one for our purpose. My general idea here is to identify a set of data, as well as a pattern in data with *information*, as explored by this field; further specification follows.

Information theory is a discipline of mathematics and engineering that is motivated by the applications for electrical engineering and computer science. Claude Shannon, in his now classical paper (1948a; 1948b), articulated a theoretical framework for the transmission of discrete data over channels without and with noise. A noisy channel implies that a set of data was sent, and must be interpreted or reconstructed by a receiver, which receives a distorted version of the data set. The transmitted data is usually referred to as *information*. Figure 4.7 provides a schematic illustration. As we will see, due to its close relation to pattern recognition in the most general sense, we are mostly interested in the aspects of information theory concerning compressing data for transmission.

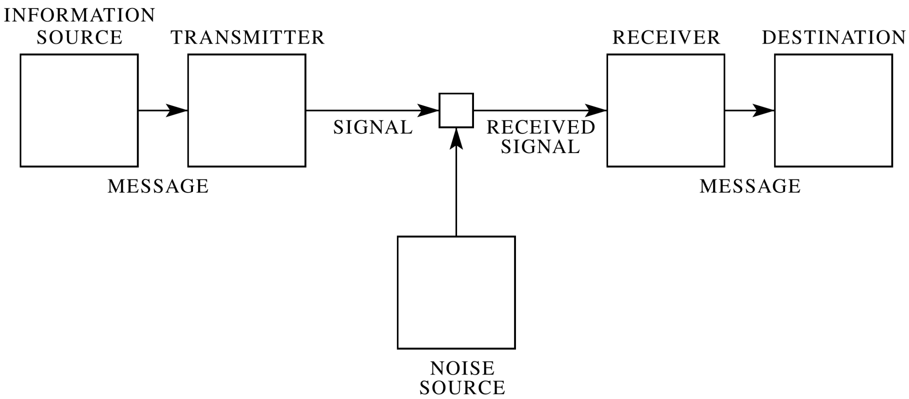


Figure 4.7: Schematic diagram of a general communication system (Shannon 1948a)

As Kolmogorov (1965) stresses, several very different approaches were introduced to mathematically quantify the amount of information in an information-theoretical framework. I want to explain the preconditions of the discussion by a very simple example. Assume we have a set of data d, which is a binary string[1] of the length of 3, *i.e.*

$$d \in \{(b_1, b_2, b_3) : b_1, b_2, b_3 \in \{\texttt{0},\texttt{1}\}\} =: \{\texttt{0},\texttt{1}\}^3,$$

and this set of data, for example $d = \texttt{010}$, has to be transmitted via a noisy channel. For the sake of applicability to our problem I want to emphasise that, in

[1] In information theory information is often referred to as 'strings' or 'words' (instead of *e.g.* 'texts'; 'numbers') to highlight that the treatment of the data is purely syntactical without taking any non-syntactical meaning (*e.g.* natural kinds; propositions; objects that are stipulated by a physical theory) of it into consideration.

principle, strings can also be texts, images, descriptions of waves or the like.

An important insight for the designer of a technical communication system concerning the space of possible data $D := \{\texttt{0},\texttt{1}\}^3$ is what the probability distribution of the $2^3 = 8$ possible different receivable outcomes

$$\texttt{000}, \texttt{001}, \texttt{010}, \texttt{011}, \texttt{100}, \texttt{101}, \texttt{110}, \texttt{111}$$

is; if the actually sent strings can be only one or two entries from this list, the transmitter design has to be different from a scenario in which every outcome is equally probable. That is why the *entropy* of the space of possible data D is a widely discussed measure in the context of quantifying the amount of information of a string. Be

$$\mathbb{P} : 2^D \to [0,1]$$

a probability measure (with 2^D being the power set of D) that induces a probability distribution over D. The entropy for D with regard to $\mathbb{P}$ is defined as

$$H(D,\mathbb{P}) := -\sum_{d \in D} \mathbb{P}(\{d\}) \log_2 \mathbb{P}(\{d\})$$

with the convention[1] $0 \log_2 0 := 0$. Figure 4.8 provides an intuitive illustration. A base of 2 for the logarithm is natural in a computer theoretic setup since the addition of one bit doubles the number off expressible strings. The entropy $H(D,\mathbb{P})$ is a measure of how much information is given by a string, *e.g.* `010`, with regard to the space of possibilities D and the probabilities $\mathbb{P}$. In fact, as Shannon shows (sect. 6, theorem 2), entropy is the unique measure for this under some rather weak and natural constraints. That is why entropy plays such a prominent role in thermodynamics (in particular due to its use in formulations of the second law).

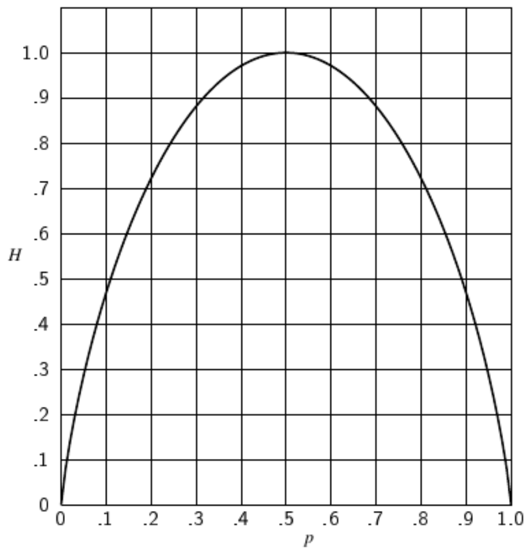


Figure 4.8: Entropy for two possibilities with probabilities p and $1-p$ (Shannon 1948a, p. 11)

[1] Due to $log_2(0) = -\infty$ we need to introduce this convention; null sets naturally occure in many setups: strings that will certainly never be sent. $0 \cdot -\infty$ is undefined in the common calculus.

How are entropy and complexity of information related? And how do these help to explicate patterns in data? As figure 4.6 (p. 119) illustrates and Ladyman and Ross' adaption of Dennett's notion of 'real patterns' hints at, the possibility of a distinction between the formerly defined pattern and the distorting noise in the data is *the* crucial aspect of a pattern. An information space $(D, \mathbb{P})$ with maximum entropy implies that every string has the same probability (or *propensity*) to occur, which means that no pattern is *expected* to play a relevant part at all. This is not a trivial point. Imagine the images from figure 4.6; maximum entropy means that all four images are realisations of the noise with the same probability to occur. And this is not want we want, when we talk about the pattern in this figure. Another example that Shannon and Kolmogorov refer to is the use of everyday language; the string '`house`' is much more likely to occur than '`KKHGU`', because our language shows patterns like words, grammar and sentence structure.

Complexity of information can be seen as a measure of how *epistemically simple* a string is regarding a defined information space $(D, \mathbb{P})$. Given the information space of the English language, '`KKHGU`' is more complex than '`house`'—just imagine how an English speaking agent could remember these strings. In other scientific discussions, like Bennett's (1990), complexity is used to describe the fundamental differences between living organisms (high complexity) and other matter (low complexity) on the background of the information space of physics and chemistry. In a low entropy information space (*e.g.* English language) many strings with low complexity (*e.g.* '`house`') occur. In a high entropy information space (*e.g.* the last four digits of the phone numbers in your personal phone book) most strings have a very high complexity.

What is a pattern? For epistemic reasons, a pattern should be relatively uncomplex—this seems to be Dennett's reason to introduce "real patterns" (on the background of the information space of human cognition and sensory capabilities) with the problems of them depending on human agents and having ontological implications. But the complexity depends on the information space—'`house`' is uncomplex in English, but complex in Latin. Therefore, it seems to make sense to loosely explicate a pattern as a string with relatively low complexity regarding the information space. Note that the information space in a scientific context is given by the full body of scientific background assumptions and the language that is used for it, which undergoes changes with time.

One could start an approach to explicate patterns by, firstly, explicating complexity and then by, secondly, stipulating that a pattern is a string under a certain complexity threshold, or with relatively low complexity in comparison to most other strings with regard to the information space, or the like. However, to make the notion of complexity more accessible and to make its weaknesses for our application more apparent, I briefly introduce an influential and useful approach, the *Kolmogorov complexity*. The Kolmogorov complexity, which is also applied to quantify the *degree of randomness*, also helps to further illuminate the important role of the information space.

Despite many definitions that are well-embedded in modern information theory, for instance by Vitányi and Li (2000), I use a definition and syntax close to Kolmogorov's (1965) original introduction with some simplifications, due to its brevity and clarity for our purpose. The intuition is that the Kolmogorov complexity of a pattern is given by the length of its shortest description in a given language. For simplicity, we assume to have

$$y \in \{0{,}1\}^n, \text{ which we call } \textit{pattern}_K$$

with some realistically large $n \in \mathbb{N}$, *e.g.* the number of pixels in the images at figure 4.6 (p. 119). We define

$$\{0{,}1\}^\infty := \bigcup_{i=0}^{\infty} \{0{,}1\}^i \times (0, 0, 0, ...)$$

to denote the set of all infinitely long binary series and

$$[\,\cdot\,] : \bigcup_{i=0}^{\infty} \{0{,}1\}^i \to \{0{,}1\}^\infty,\ [\,\cdot\,] : y \mapsto y \times (0, 0, 0, ...)$$

the translation of a finite binary series into $\{0{,}1\}^\infty$. Furthermore,

$$\begin{aligned} l : \{0{,}1\}^\infty &\to \mathbb{N} \cup \{\infty\}, \\ l : \quad s \quad &\mapsto \max_{i \in \mathbb{N}} (s_i \neq 0 \text{ for } s = s_1 s_2 s_3 ... \text{ with } s_n \in \{0{,}1\} \text{ for all } n \in \mathbb{N}) \end{aligned}$$

defines the *length* of a string $s \in \{0{,}1\}^\infty$. Be $p \in \{0{,}1\}^\infty$ with $l(p) < \infty$ a *program* and a φ with

$$\varphi : \{0{,}1\}^\infty \to \{0{,}1\}^\infty \text{ and}$$

$$\varphi \text{ is } \textit{partial recursive}$$

is the *programming method*. Partial recursiveness means that it can be computed by Turing machines and, intuitively speaking for our purpose, foremostly avoids that φ is chosen in a way that it cannot be explicated and computed in a finite way.[1] φ is fixed for an information space and can be thought of as, for instance,

[1]Partial recursive functions are most often defined on the domain of $\mathbb{N}^n$ for some $n \in \mathbb{N}$. But the notion can easily be redefined for $\{0{,}1\}^\infty$ by the use of the standard transformation of binary numbers to decimal numbers and vice versa. In other words, there are trivial bijective mappings $\{0{,}1\}^\infty \to \mathbb{N}$. For the definition of partial recursive function I refer to Minsky (1967, sect. 10.5).

the parser of a programming language or the interpreter of the English language that provides the physical reference to '`house`' (in a Fregean sense of reference).

Finally, we can define the *complexity*$_K$ K of a *pattern*$_K$ and a given programming method φ as

$$K_\varphi : \{0,1\}^\infty \to \mathbb{N} \cup \{\infty\},\ K_\varphi : [y] \mapsto \min_{\varphi(p)=[y]} l(p)$$

$K_\varphi([y]) = l([y])$ describes the case of a maximum *complexity*$_K$. Importantly, the more powerful the stipulated φ is (*e.g.* a C++ parser with a lot of libraries; a scientifically highly specialised terminology), the lower is the complexity of many patterns.

I give an example. Assume, $\hat{\varphi}$ maps the binary codified version of our standard mathematical set-theoretical language to the binary number that is expressed with this language. Be $\hat{y}$ a series of one million 1's and after that only zeros. A program could be outlined as

$$\hat{p} = \{1\}^{10^6} \times (0, 0, 0, ...)$$

and $K_{\hat{\varphi}}(\hat{y})$ would be very small, $K_{\hat{\varphi}}(\hat{y}) \ll l(\hat{y})$.

This is the general idea of Kolmogorov complexity. The approach is not restricted to series of binary numbers and can be adapted to every string and therefore every set of data and patterns with according definitions of the pattern$_K$, the length, the programming method and the program. It should be mentioned that Kolmogorov defines the complexity of a pattern originally based on some *data*$_K$ $d \in \{0,1\}^m$ with some sufficiently large $m \in \mathbb{N}$. His programming method is then $\varphi(p, x) = y$, but I ignored this further aspect for simplicity.

Some further theoretical insights are of interest here. Regarding the programming method φ, the *invariance theorem* roughly states that for a given pattern or class of patterns a complexity optimal programming method is only as good as any other descriptively sufficiently powerful programming method plus some constant that is necessary to describe the optimal programming method with the other programming method.[1] Chaitin (1992) showed that, roughly said and also very intuitive, the choice of the programming method φ and a maximum program length $l(p)$ (which is necessary to actually run the routines on computers) always determines a maximum threshold of Kolmogorov complexity $L \in \mathbb{N}$ that can be determined. He denoted this result 'incompleteness theorem', due to the unprovability of a statement like: $K_\varphi(s) < L + 2$, if we now that $K_\varphi(s) > L$ for some string $s \in \{0,1\}^\infty$.

[1]For a discussion on the invariance theorem, see Li and Vitányi (2008, sect. 2.1)

Kolmogorov complexity seems very promising after our discussion so far. What are the problems? For the definition we use $\min_{\varphi(p)=[y]}$, which refers to the set of all programs p. Even if the set of all p is countably infinity it is practically very hard to feasibly determine the optimal p under φ. If a certain pattern from a realistic example is given and we have a lot of computational power at hand, it may still be take millions of years until even our best computers made a decision about the optimal program to *compress* the pattern x in question. But this is not the way we (including non-human agents) epistemically talk about patterns. Usually, when we refer to a statistical or visual pattern, we are able to provide a (maybe vague) description of it in the first hand. We know that '`house`' is a pattern to us English speakers, since we already have a list of vocabulary at hand. It is not the case that we see '`house`' and then think about every possible combination of five letters, find possible references for all of these mostly made-up words and finally find out that houses are objects that can be referenced very easily. The pattern that is shown in the top left image from figure 4.6 (p. 119) is a pattern for us since we can *construct* the depicted geometric object very easily from everyday geometry by referring to rectangles and lines and not by going through every possible arrangement of black and white pixles, and then find out that it might be a pragmatically good idea to talk about lines and rectangles specifically. This route of explicating complexity is therefore not a good approach to provide a descriptive epistemological account of what patterns in science are; this inadequateness holds for all relevant agents (*e.g.* humans; AIs; aliens). Again, Dennett is right regarding his neo-Kantian implications, but he is wrong by his restriction to some kind of epistemically fixed human agent and the ontological implications.

These are the reasons why, in the following, I want to focus on general pattern theory which provides an answer to the problem of pattern construction and keeping the merits of the complexity approach, which is the distinction from compressibly describable patterns from noise.

The Constructive Approach: General Pattern Theory

In accordance with (Bogen and) Woodward's intentional use of the term, we discussed "patterns" ("in science") in the broadest possible meaning. Obviously, it is a very extensive endeavour to actually *show* that one can mathematically explicate all cases of patterns. Ulf Grenander's[1] œuvre revolves to a significant amount around exactly this goal. I want to point out that every judgement regarding how well he achieved his goal can be only unfair without a sufficiently comprehensive investigation of his work and this is not the aim of this thesis.

[1] Mukhopadhyay (2006) provides a helpful overview over Grenander's œuvre and academic career.

General pattern theory is *constructive* in the most general sense, meaning that every pattern comes with a finite and recursive construction rule, but these construction rules use one of the most general epistemic and mathematically fleshed out sets of vocabulary, which is mathematical algebra.—I justify this epistemological view in chapter 3 about mathematics and my ante rem structuralist's position. The epistemological aims of general pattern theory are also stated by Grenander in an interview with Mukhopadhyay (2006):

> [T]he emphasis in pattern theory is on the actual *act of* knowledge and *act of* understanding. The key phrase is "act of," that is, we want to learn the process of understanding and the emphasis is not necessarily on specifics of what it is that we understand. Pattern theory is more like mathematics of knowledge representation[.] (p. 14, original emphasis)

Regarding the patterns in science that are to a significant amount mathematically explicated (*e.g.* everything that is statistically tested) general pattern theory can be seen as a project of a diligent foundation for these patterns quite similar to the attempts of defining axiomatic set theory or category theory as a general foundation for the entire body of mathematics.

This is how I approach general pattern theory for our purpose. I introduce only the most basic mathematical ideas of Grenander's approach from one of his first papers (1970) on the topic. These provide sufficient insight into the program to roughly understand the motivations and solutions. However, patterns from Grenander's approach that are explicated in his later works, are used in various applied fields, such as biomedical engineering, computer science, electrical engineering, speech recognition, and computer linguistics. The most accessible comprehensive sources are his late book with Miller (2007) and (1993). To make the ideas more accessible, I provide examples for every definition.

Be S a set of *signs* with $\phi \in S$, which are assigned to the most primitive epistemically accessible objects. Examples include a pixel in the images from figure 4.6 (p. 119) ($S := \{\texttt{black}, \texttt{white}, \phi\}$) or the letters and punctuation marks for the English language ($S = \{a, b, c, ..., A, B, C, ..., ., ,, ?, !, ..., \phi\}$). ϕ denotes the *empty* sign. Note that the elements of S *denote* the most primitive epistemically accessible objects in a modelling sense and it is not the case that they *are* these objects. Otherwise the element `black` from the first example would have been replaced with ■. A crucial aspect of this approach to explicate also non-visual patterns, like patterns in sound data that we can hear. However, we can still maintain the view that the patterns are mathematical, even if the signs denote something phenomenologically distinct, like a most primitive piece of a hearable sound, like one note in Bach's Goldberg variation[1].

[1]Piano music is a very nice example for our purpose. The reason is that we might have the first impression that the notes should be the range of the signs, since the notes seem to already

Any sign $s \in S$ with $s \neq \phi$ belongs to one and only one *paradigmatic class* S_σ, which are disjoint except for the empty sign ϕ. The idea behind paradigmatic classes is to establish a very general separation of signs that can be exchanged only individually among each other. (cf. 1993, p. 547) For the example of the English language the signs for the end of a sentence $S_{\sigma_1} = \{., !, \phi\}$ or the capital letters for the start of the sentence $S_{\sigma_2} = \{A, B, C, ...\}$ may serve as an example here, but those have some obvious shortcomings. These are namely that for S_{σ_1} we need very restricted rules for punctuation and for S_{σ_2} we need to forbid capitalised proper names.

A vector $\vec{c} = (s_1, s_2, ..., s_n)$ with $n \in \mathbb{N}$ and $s_\nu \in S$ is a *configuration*. An example is an English sentence or word (*e.g.* '`house`'). Furthermore, we have a set of syntactical rules (*e.g.* ⌜a ' ' occurs only between two other signs⌝) $\mathscr{R}$, which is the *grammatic* for the configuration. '`Knnn k !   ..`' is not an English sentence and can therefore not be a configuration in the pattern explication of the English language.

Given the disjoint partition of S into paradigmatic classes, modulo the empty sign ϕ, for any paradigmatic class S_σ there is a semigroup $(T_\sigma, \circ)$ of transformations $T_\sigma \ni t : S_\sigma \to S_\sigma$ with $t(\phi) = \phi$. Furthermore, the identity mapping is in these semigroups of transformations, *i.e.* $T_\sigma \ni t_e : s_\sigma \to s_\sigma$ for all $s_\sigma \in S_\sigma$. Sometimes there is a subset $S_\sigma^{(pr)} \subset S_\sigma$ such that for any $s \in S_\sigma$ there are unique $s^{(pr)} \in S_\sigma^{(pr)}$ and $t^{(s)} \in T_\sigma$ such that $s = t^{(s)}(s^{(pr)})$. $S_\sigma^{(pr)}$ is called the *set of prototypes* and the $t^{(s)}$'s are the *paradigmatic transformations*.

Why are prototypes important here? A very crucial aspect of patterns in data is that we conceptually distinguish them from the (usually unavoidable) noise. Furthermore, one and the same pattern, may show slightly different features in different sets of data. The intuition behind prototypes is that they are used to describe a certain raw or idealised version of the structure, the actual pattern, and the paradigmatic transformations are applied to describe the relation between this pattern and the actual data. Take, for instance, an English text and for our pattern analysis it might be completely sufficient to consider only Fregean references[1] of the words. The sentences

(a) `My car is red.`

(b) `My automobile is red.`

(c) `My sedan is red.`

denote the patterns. However, performances of, for instance, the Goldberg variation by Glenn Gould or by Igor Levit are considered to differ significantly regarding its artistic content. That is why we cannot simply translate Bach's notes into signs to explicate the relevant artistic patterns of these two pianist's performances.

[1]References (Ger. *Bedeutung*) of nouns are discussed in Frege's (1892) classical paper.

are surely different sets of data. However, since given the information that I own only one car ‘`car`’, ‘`automobile`’ and ‘`sedan`’ are completely synonymous here, it is very reasonable to identify these two sentences for the purpose of an analysis of patterns in a text. One could chose, $s_1 := \texttt{automobile}$, $s_2 := \texttt{sedan}$, $s^{(pr)} := \texttt{car}$ with $t^{(s_1)}(\texttt{car}) = \texttt{automobile}$ and $t^{(s_2)}(\texttt{car}) = \texttt{sedan}$. In this example S_σ is the set of all words that refer to cars. We could even extend the example to German words (*e.g.* ‘`Auto`’) very naturally, since the grammar of the exemplary sentence is similar. However, note that $\hat{s}^{(pr)} :=$  would not be a proper choice, because ‘’ is not a word in Latin letters and therefore $\hat{s}^{(pr)} \notin S \supset S_\sigma$.

To mention an example that is even closer to what we intuitively have in mind when talking about patterns, I refer to biometric identification algorithms for human faces. In a first step, the image is transferred into greyscale (because the colour tones of an individual differ too much in different circumstances like health or lighting conditions). If we explicate the patterns on a pixel level, then the prototypes are greyscale pixels, whereas S contains all visible colours.

For an entire configuration $\vec{c} = (s_1, s_2, ..., s_n)$ any transformation $t \in T_\sigma$ can be applied as $t(\vec{c}) = (t(s_1), t(s_2), ..., t(s_n))$ with the natural extension $t(s) = s$ for all $s \notin S_\sigma$.

Finally, an equivalence relation R over configurations is introduced to define *images* I, which are the equivalence classes of R. The idea is that R defines configurations that are indistinguishable according to the interest of the observer. For the example of English sentences with Fregean references above the further sentence

(d) `I have a red car.`

can be introduced with (`I have a red car.`, `My car is red.`) $\in R$. This example illuminates the general difference between the transformations $t \in T_\sigma$ and the images defining equivalence relation R. Transformations separate the signs that are relevant for pattern detection, the prototypes, from other signs, whereas R describes the structure of the pattern itself. The given example already indicates how extremely hard it is to really explicate R in realistic examples.

The described theory focusses on transformations of signs, and the general aim of this theory is to define patterns. The image characterises a pattern, but all the laid out vocabulary above is a preparation to define patterns *constructively*. As Grenander (1970) puts it, “an array whose entries describe the history of formation of I” (p. 179) gives an epistemic route to a pattern. The *pattern* itself is defined as an image I with a construction $A(I)$ that determines the signs, the grammar for the configurations, the prototypes, the paradigmatic transformations and the equivalence relation R. All these mathematical objects that are part of $A(I)$ can be explicated, but the actually used mathematics for this depends on the cognitive and sensory capabilities of the relevant agents. Even if the relevant patterns are quite similar, the *analysis algebra* A_{Chess} of Garri Kasparow and Deep Blue differs vastly due to the very different cognitive and sensory capabilities of these two chess players.

This description is a crude simplification of general pattern theory, but it is, in my view, sufficient to show its merits. To make it more accessible I provide a very simple and visual step-by-step example that makes use of all the notions developed above:

1. We want to develop a pattern theory for geometric objects. For epistemic reasons the signs are rectangles of various forms colours and scales. An easy way to denote this is

$$S = \left\{ \boxed{1}, \boxed{2}, \boxed{.87}, \boxed{47}, \boxed{14}, \ldots, \phi \right\}$$

with the little number in the rectangles being a size scaling factor that I introduce for merely presentational purposes. S is obviously very large.

2. The grammar for the configurations states that the signs can be compiled to create a gapless surface without any overlaps like the following examples: , , , whereas I ignore any scale indication here. An exemplary pattern is a thick rectangle edge, which is illustrated by figures 4.9 and 4.10.

3. We consider only the distinction between white and non-white areas as relevant for geometrical patterns. That is why we introduce the two paradigmatic classes accordingly:

$$S_{\sigma_w} = \left\{ \boxed{1}, \boxed{2}, \boxed{.87}, \boxed{47}, \boxed{14}, \ldots, \phi \right\}$$

and

$$S_{\sigma_c} = \left\{ \boxed{1}, \boxed{2}, \boxed{.87}, \boxed{47}, \boxed{14}, \ldots, \phi \right\}.$$

The grammar restricts the cases in which we can replace one sign with another from the same paradigmatic group. Furthermore, we cannot replace a sign from S_{σ_w} with a sign from S_{σ_c} and vice versa without changing the pattern.

4. Since all non-white colours fall into one paradigmatic class, we can chose one colour as a prototype. I chose gray mainly for optimised visibility:

$$S^{(pr)}_{\sigma_c} = \left\{ \boxed{1}, \boxed{2}, \boxed{.87}, \boxed{47}, \boxed{14}, \ldots, \phi \right\}$$

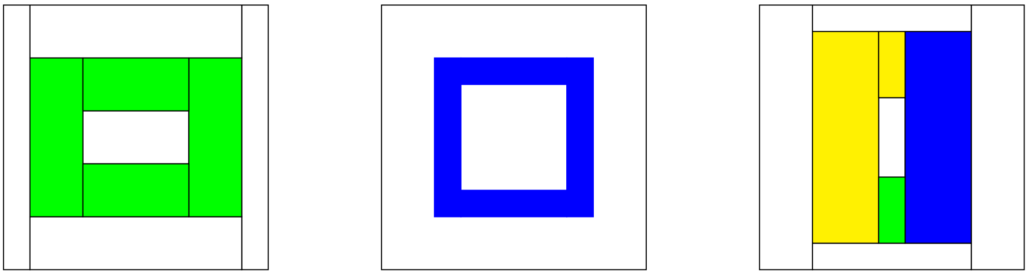


Figure 4.9: Three of the many possible ways to construct a thick rectangle edge with indication how they are composed from signs (left, right) and without (mid).

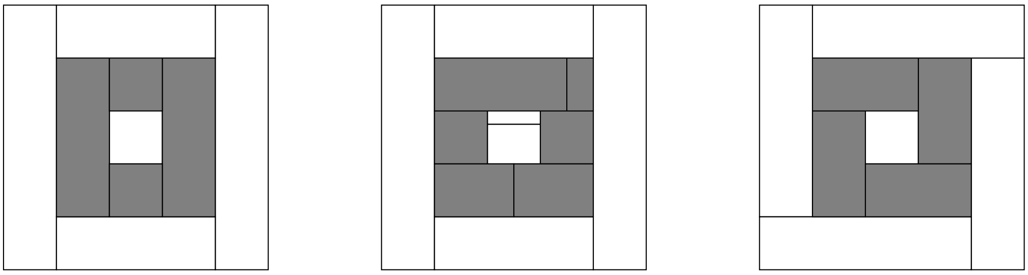
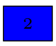
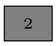


Figure 4.10: Three of the many possible ways to construct a thick rectangle edge from filled rectangles.

with $S^{(pr)}_{\sigma_w} = S_{\sigma_w}$.

5. The paradigmatic transformations are the inverses of the natural projections of the signs and configurations to $S^{(pr)}_{\sigma_w} \cap S^{(pr)}_{\sigma_c}$. Examples of $t \in T_{\sigma_w} \cap T_{\sigma_c}$ include $t^{-1}(\boxed{1}) = \boxed{1}$, $t^{-1}($$) =$  and $t^{-1}($$) =$ .

6. The hardest explication in this task is the one for the equivalence relation R. Figure 4.10 illustrates that there are many different ways to construct the exemplary geometric object from rectangles. However, if we restrict all the sizes and numbers involved to countable sizes and numbers (*e.g.* by restricting it to rational numbers $\mathbb{Q}$), which we should do for exactly this epistemic reason, then the class R is perfectly constructable. One could do this construction by, firstly, starting with the finite list of ways to construct the rectangle edge with the minimum number of rectangles and, secondly, iterate a further division of the used rectangles into arbitrarily small rectangles.

Note that this example helps to understand why the upper left shape from figure 4.6 is considered by human agents to be a simpler pattern than the other three examples. The exemplary shapes in figures 4.9 and 4.10 are chose to make this more apparent.

To summarise, the merits of general pattern theory is its constructiveness that can directly adapt to any peculiarity of the relevant epistemic agents. Other than Dennett's unexplicated version of real patterns and Kolmogorov complexity, this approach to patterns provides the ground to not only define patterns, but also test them based on the paradigmatic transformations and equivalence relation over images.

4.5 Conclusion

According to some influential accounts of anthropocentric epistemology and even metaphysics, what counts as a ("real") pattern depends on the human being's sensory and epistemic capabilities, as well as human motives and interests. But since scientists make use of various kinds of technical auxiliaries, human sensory and epistemic capabilities cannot serve to restrict the class of patterns in science. To make our notion of patterns in science adaptable to past, contemporary and future developments in science in principle, we cannot claim much more than that patterns *are* (in an ontological and epistemological sense) mathematical properties.

This claim is justified by the historical fact that computers are more and more able to precisely explicate various patterns in data, including very non-trivial cases, such as human facial recognition. However, if I accept such an abstract notion of patterns, which I do, I can still claim that a pattern must be *constructive* for epistemic reasons. Even if we use our fastest computers or some help from an extraterrestrial colleague, there must still be a mathematical way to test or explicate the pattern feasibly. This cannot be achieved with the general framework from the information theory of Shannon, Kolmogorov and Chaitin. Grenander's general pattern theory provides a well-suited approach, due to its constructiveness.

To clarify the connection between patterns and phenomena, which is not an identity, I distinguish between general and concrete patterns. Further scientific assumptions that are additional to the observations and data in question, determine what class of concrete patterns make a general pattern that corresponds to a phenomenon. In principle, these assumptions can be mathematically explicated, but the significant structural dissimilarities between many different concrete patterns of the same general pattern justify my differentiation between concrete and general patterns.

I raise some introductory questions in section 4.2 (p. 108) about the demarcation between parts or features of the data that are relevant to the pattern in it and the other parts or features of the data that are irrelevant to the pattern in it (*e.g.* noise). To answer these questions I want to emphasize that every constructive mathematical property of a set of data can be regarded as a pattern in this data. Every mathematically explicable feature of a picture that shows the Dome of the Rock (figure 4.2) or the Sierpinski triangle (figure 4.4) can be a pattern or be excluded from a pattern by an agent. *Multifractal geometry*[1] shows that self-similarity can be mathematically explicated. A written text can be a visual pattern and its semantic content can be a pattern of references or meanings. The separation from a pattern and the noise in the data is purely stipulated and based on the scientific interests, too. If a scientist, for instance, investigates empirical data to evaluate a hypothesis which's validity makes him to expect only *white noise* in the data (*e.g.* molecule movement follows a Brownian motion), then the white noise is the scientific pattern, whereas any deviation from the white noise counts as noise *regarding* the scientific pattern detection in such a scenario. Aesthetic criteria of human agents play a defining role for phenomena (cf. 5) but we cannot refer to them to narrow down the large class of possible patterns by any manner. The important point is that a (not necessarily human) agent defines patterns on the basis of his specific scientific interests and assumptions that guide the phenomenon selection. Indications of colours can play a role (*e.g.* for the phenotype of albinism) but they do not have to. In the facade image 4.2, only the very coarse grained pattern of sub-pattern changes can be regarded as a pattern, but quite contrary to this the whole configuration of pixels can be regarded as patterns, as well.

[1]For an introduction into multifractal geometry see Mandelbrot (1983).

Chapter 5

Conclusion about Phenomena

Chapter Abstract

The distinction between phenomena and non-phenomenal features of the empirical world is not realistic, but pragmatic and based on the shared body of background assumptions, as well as the agents' sensory and cognitive capabilities. A phenomenon is a feature of the empirical world and it is described by its general pattern, which can be fully known only by an epistemically unbounded agent. Therefore, a phenomenon has an investigator-independent reference in the empirical world (whatever that may be) and it is not theory-laden in a strict sense. A phenomenon is not supervened and also not represented by its general pattern or by one of its concrete patterns. The notions of evidence and of phenomena have some similarities, but they are not identical, due to the different roles they play in processes of scientific inference.

The overall aim of this thesis is to, firstly, explicate phenomena, data and patterns, and, secondly, to explain the relation between these concepts in science based on these explications. In this chapter I aim to explain what phenomena in science are and how they are related to data and patterns. Much of what is discussed in this chapter makes significant use of the concepts that I describe in earlier chapters.

In section 5.1 I introduce my explication of a notion of phenomena. The other sections of this this chapter focus each on one specific aspect related to phenomena (the section titles give clear indications).

In a large part of this chapter I discuss the relation of phenomena to some widely discussed notions in philosophy of science, namely natural kinds, structures, properties, utterances, scientific explanation, scientific representation, supervenience, theory-ladenness and evidence. I introduce and discuss these notions by reference to very general and widely shared philosophical explications of them. Naturally, due to the scope and volume of the thesis, I do not aim to provide a thorough explicative discussion on these notions, but I want to show that the relation of them to phenomena can straightforwardly be decided upon on the basis of my sparse anti-realistic and pattern based notion of phenomena.

5.1 Explication of Phenomena

Section Abstract

I aim to develop an empirically adequate descriptive notion of phenomena. I defend an anti-realistic, pragmatic distinction between phenomena and non-phenomenal features of the empirical world. Phenomena are real features of the empirical world, but the distinction between phenomena and non-phenomenal features of the empirical world is based on the shared body of scientific background assumptions, as well as the agents' sensory and cognitive capabilities. Ones metaphysical position towards phenomena is closely related to ones metaphysical position towards natural kinds. The notion of phenomena is closely related to the of notion scientific explanation, because one criterion for the selection of a phenomenon is that a theory about a phenomenon has good explanatory power over other observations.

I introduce phenomena separated from the introduction of data (chapter 2) and patterns (chapter 4). The reason for this is the strict conceptual distinction between the explication of phenomena, which are features of the empirical world, and the one of data, which are mathematically explicable information. More precisely, data are mathematical objects about which we need further background information regarding its origin to make scientific inferences with it. However, as I discuss below, even if phenomena are features of the empirical world, the selection between phenomena and non-phenomenal features of the empirical world is an epistemic act of a scientific agent, which can be a human scientist or an AI (in the broadest meaning of "artificial intelligence", cf. 3.3) or any other intelligent (cf. p. 90) agent.

Most importantly, I propose a notion of phenomena that incorporates an anti-realistic distinction between phenomena and non-phenomenal features of the empirical world. I claim that this notion is an empirically adequate descriptive explication of the scientists' use of the term 'phenomenon'.

Phenomena are a well established concept in the philosophy of science. Therefore, I want to briefly introduce the concept as far as necessary to prepare the ground for the explanation of the relation between phenomena and patterns in data. Phenomena are hard to define due to their specific heterogeneity; they play an essential role in every field of science and are in some cases formulated in mathematical terms (*e.g.* in some cases in physics) and in other cases not very mathematical (*e.g.* most cases in psychology). Hence, the concept of a phenomenon needs to be sufficiently flexible. In this section I also introduce the problem of identification of phenomena with patterns in data, which is often overlooked, in particular by (Bogen and) Woodward.

Phenomena for (Bogen and) Woodward are features of the world that play an important role in science, but they are neither a part of theories, nor can they be identified with the data. The authors introduce them as "features of the world that in principle could recur under different contexts or conditions." (Woodward 2011, p. 166) If the concept of phenomena is introduced in this way, then notorious problems in the explanation of science can be avoided. These problems are described by the following questions: Why do scientists edit or modify raw data before using them for confirmation (or falsification) of a theory? Can an observation be *theory-laden*, if the observation is a process in which the theory itself plays no role?

The mere identification of phenomena with patterns in data might be conclusive to some prima facie. Let us consider the simple example of the melting point of lead.[1] Let the data from experiments be a time series of numbers that denote temperatures. These numbers from different repetitions of the experiment slightly deviate from each other. The phenomenon is detected in the data by the clustering of the numbers around a certain discrete number. Further background assumptions also play a role in the detection of the phenomenon, such as a specific probability distribution (*e.g.* normal[2]), which is assumed to describe the samples from the measurement. These background assumptions influence the resulting description of the melting point (*i.e.* what number it exactly is). In this example, the discrete melting point is the expectation value of the probability distribution that is fitted to the samples gathered by the repetitions of the experiment. This expectation value can change, if another probability distribution is assumed to describe the samples (in many cases in science the number of available or producible samples is not sufficient to determine a *real* distribution, if such one exists).

What exactly is the phenomenon and the corresponding pattern(s) in this example? The phenomenon is the fact that there is a melting point. More precisely, the phenomenon is the constitution of what appears to us as an outside world that can be measured as the *general* pattern (cf. 4.1) of the melting of lead. That a unique melting point $\Theta/C^{o} \in \mathbb{R}$ under a presupposed air pressure really exists is a

[1] This is a common example from Nagel (1961, p. 79), and Bogen and Woodward (1988, pp. 307–310).

[2] It is not a trivially valid assumption that a normal distribution statistically describes the outcome of an empirical sample. This assumption already implies strong (in)dependency properties of the underlying processes on lower scales. The mathematical theorem that describes these implication of the normal distribution is the *central limit theorem* (See Klenke, 2008, ch. 15, for details). The common mistake of ignoring these implied (in)dependency properties has lead to serious issues in financial risk management (this a conclusion I draw from my work experience in the field).

theoretical assumption that may not be fully verifyable or adequate. A probabilistic or vague description of a melting point (or: melting *area* or melting *situation*) could be more adequate. Nevertheless, this is a question of physics itself and not of philosophy of physics. The phenomenon is a *property* of lead in the sense that there is very convincing evidence that there is what we call melting without the knowledge how it should be described perfectly, for example as a specific discrete temperature.

What is the corresponding pattern in the data in this example? Roughly speaking, the pattern is the clustering of the numbers in the results from the repeated measurements. This pattern is prima facie accessible, but a precise mathematical explication is no simple task. One problem for the explication is that every series from a new measurement shows different numbers. Therefore, the explicated pattern must be defined statistically or with integral vagueness.

Since the data is a list of rational numbers and therefore a comparably simple mathematical object, the pattern must be definable in mathematical terms as well. For this example, we need to clarify what *clustering* is. This definition must be applicable to all possible versions of series of measurements from experiments that are intended to detect the melting point of lead. Naturally, for more complicated examples of phenomena the definition of the corresponding pattern is much more complicated and may not be directly available in mathematical terms. Since clustering is in this example explicated by being a series of numbers that can be tested positive as being a sample from a certain distribution (*e.g.* normal) with an expected value, the pattern is exactly this mathematical description. But it is important to note that this pattern depends on the experimental design and could be different if we chose another experimental design for the investigation of the same phenomenon. According to my terminology, this pattern is one *concrete* pattern of the phenomenon (see section 4.1).

Can a general explication of phenomena be derived from this illustrative example? Phenomena are just features or supposed features of the part of the empirical world under investigation by the science. A phenomenon does not need to be described or understood completely by a scientific theory. That is why Bogen and Woodward (1988) write that scientists are more concerned with "claims about phenomena" than with "claims about data" (p. 314). Empirical data can hint at a phenomenon by showing a certain pattern, but the data and the information about its origin alone cannot provide any indication of whether the corresponding feature of the empirical world is non-phenomenal or whether it counts as a phenomenon. As will be further commented on in section 5.2, phenomena themselves are metaphysically completely independent of any knowledge of scientists and they are most often not completely describable by them in practice. With this I mean, trivially, that lead has a melting point and there are albinistic animals without any scientist discovering it. Quite obviously a newly discovered jungle frog was a biological species already before the first zoologist recognised it.

However, the *notion* of a specific phenomenon is to some degree observer dependent, because the demarcation between a phenomenon and a non-phenomenal feature of the empirical world is determined by the scientist's body of theoretical background assumptions and his available capabilities. This observer dependence is not restricted to human individuals and not to human societies. Quantitative macro economics is a good example in which the positive dependence between two variables (*e.g.* interest rate and GDP growth) may be regarded as a phenomenon, if it occurs in a roughly linear form, which is statistical *correlation*. If the two variables are statistically dependent in a non-linear way that cannot be detected by the common analysis methods (including the ones that are executed by software) the underlying empirical feature does not count as a phenomenon, even if it could, in principle, be detected by a pattern in the data. The reason for this is that the scientific community cannot detect it with their available body of statistical knowledge.

Some may object at this point that the melting point of lead perfectly counts as a phenomenon, even before any human or machine may have discovered that metals melt at high temperatures. But in my view, this objection implies a wrong descriptive conception of a phenomenon, which may not be obvious. Here are my reasons. Human beings are biologically restricted by their sensory (*e.g.* visible light, non-dark matter) and cognitive (*e.g.* slow at calculations, finite life time) abilities. We can describe features of the empirical world only under these constraints. We have access to the empirical aspects of world that can be described under these constraints, but we are inherently blind to the rest (*e.g.* overly complex features of the empirical world).

This may remind some readers of the classical distinction between *observable* and *unobservable* (see chapter 1). Humans expand their sensory (*e.g.* microscope) and cognitive (*e.g.* software) abilities for scientific endeavours. However, the causal works of such an expansion device counts as part of the body of scientific knowledge—these devices are not *magical* and devices that are not sufficiently understood with the body of scientific knowledge (*e.g.* magical crystals) are not scientifically accepted as devices for the expansion of our capabilities for scientific inferences. To say it more directly, even if we constantly expand our body of scientific knowledge and use this knowledge to *leverage* our capabilities for further observations, we are still restricted by our human capabilities. Science happens in a historically grown network of *trust* in expert knowledge and technical auxiliary devices. Our academic textbook education for young researchers that inroduces the methods and topics of the field may aim to a large extend to establish criteria for trust rather than rationality alone, since the rational scrutiny of the full body of a field's knowledge is not feasible for a human agent.

These agents' restrictions also include artificial intelligences, because we design them in a non-magical way, even if the programmers may not be able to understand the vastly complex decisions that a trained AI with machine learning makes. (cf. 3.3) Since humans and also artificial intelligences are blind to certain parts of the empirical world (*e.g.* dark-matter realm, gods, ghosts, a possible computer simulation we live in)[1] we cannot distinguish between phenomenal features of the empirical world and all the non-phenomenal features. We simply cannot have a sufficiently comprehensive and accurate epistemic notion of features of the empirical world.

However, scientific agents make explicit or implicit distinctions between phenomena and non-phenomenal features of the empirical world all the time. In my view, phenomena are of specific interest, because they play, metaphorically speaking, the role of conceptual *anchor points* in the body of knowledge of the relevant agents, which socially form a scientific field. Albinism is more interesting for genetics than many other mutations, because its phenotype is very easily detectable with our human sensorial equipment and culturally pragmatic connotations, such as witchcraft or racial implications of skin tones, make it a subject of increased interest for a broader community, too. In the periodic table chemical elements are sorted by their *atomic number*, which is the number of protons. Chemical elements can be distinguished via several properties from a very long list (*e.g.* aggregate state at a certain temperature, reactivity, degree of toxicity to a cockroach, taste to a human, sound of a flute made of it). The number of protons were chosen as the distinguishing feature because Dmitri Mendelejew and Lothar Meyer derived the masses of one atom from measurements and grouped the elements roughly into subgroups according to very noticeable criteria (*e.g.* aggregate state at room temperature). Later, atomic models lead to the conclusion that the order is determined by the number of protons.[2] Isotopes, which are versions of the same chemical element with differing numbers of neutrons, on the other hand, are not distinguished in the periodic table despite their differing chemical properties. These are examples in which the *relevance* of certain features of the empirical world are measured by their accessibility to human agents with their specific background knowledge, assumptions and stipulations (*e.g.* room temper-

[1] It is obviously hard to find useful examples about things we are inherently blind to.

[2] A neutron and a proton have approximately the same mass. An element occurs with different numbers of neutrons. These non-standard atoms are called *isotopes*. Since elements in the period table are ordered by the number of protons alone, the atomic masses do not increase monotonically with the atomic number if we include all possible isotopes into this list.

ature is a somehow normal state for matter; weight is a very noteworthy specific property of matter). *Laws of nature* are phenomena under a very basic level of epistemic capabilities without many more theoretical background assumptions.

The role of phenomena as conceptual anchor points can also be seen the other way round: we formulate our body of theoretical background assumptions to reduce the number of anchor points for the sake of epistemic sparseness. Friedman (1974) suggests exactly this sparseness of phenomena for the predominant goal of scientific explanation:

> I claim that this is the crucial property of scientific theories we are looking for; this is the essence of scientific explanation – science increases our understanding of the world by reducing the total number of independent phenomena that we have to accept as ultimate or given. A world with fewer independent phenomena is, other things equal, more comprehensible than one with more. (p. 15)

The criteria that we have to explicate to explain why a phenomenon is considered a phenomenon are the background knowledge, the sensory and cognitive capabilities that a certain scientific community *shares*. I give a vivid and very simple example: a blind person in a laboratory can reliably distinguish liquids by the sound of their boiling (due to his trained hearing senses), but this distinctive feature is very *unpractical* for his non-blind colleagues. It may only be promoted to the rank of an accepted observation routine, if the shared body of scientific knowledge in the field implies a pragmatically favourable use for it. These *pragmatic* and *social* aspects of phenomenon selection seem very obvious to me, but to my knowledge they are usually neglected or fully ignored in the philosophical literature on phenomena.

Adapting their terminology, Kuhn's (1962) *paradigms* and Lakatos' (1978) *hard core* may play a predominant role in the body of background assumptions that guide the phenomenon selection. This may explain why some evidentially sufficiently supported phenomena were or are not accepted as such and tabooed by the majority of scientists (*e.g.* continental drift; strong architectonic similarities in (pre-)ancient buildings on different continents; UFOs in military reports)[1].

Are Phenomena *Structures*, *Properties of* - or *Utterances about* the Empirical World?

If, as (Bogen and) Woodward suggest, phenomena *are* patterns in data and we accept my notion of patterns, then phenomena would be mathematical objects.

[1] For a historical and contemporary survey on the demarcation between science and *pseudoscience* see Regal (2009).

But in contrast to that, an adequate descriptive explanation of phenomena needs to take into account that a scientist refers with 'phenomenon' to a feature of the empirical world about which he may not have sufficient knowledge.

If someone wants to defend (Bogen and) Woodward's *pattern view of scientific phenomena* (terminology suggested by Apel 2011, p. 24), then he may refer to a construction of the corresponding *general pattern* from an epistemically unbounded perspective. This epistemically unbounded agent is able to define the general pattern as the class of all of its *concrete patterns* from all possible experiments, which can obviously be listed only on a very (often infinitely) large list. This general pattern would be a mathematical object, but can we *identify* it with the phenomenon? Regarding its epistemic *role* for scientific inferences, the phenomenon and the general pattern are indistinguishable, because we (*e.g.* human scientists, AIs) will never be able to find out more about the phenomenon than what the general pattern already implies. However, *ontologically*, we can still make the case that there may be an actual empirical world *out there*, which can ontologically not be identified with a mathematical structure. In case this is false, as *ontic structural realists* (cf. Ladyman and Ross 2007) claim or Bostrom's (2003)[1] Cartesian simulation thesis implies, phenomena would in fact be identical to the general pattern.

At this point, I want to remain agnostic regarding the existence of an external world in the sense of a Kantian *Ding an sich*, but I am a sufficiently convinced *anti-realist* regarding our actual scientific knowledge to claim that it is impossible or at least extremely rare that the set of all scientifically discussed concrete pattern of a phenomenon (*e.g.* for albinism) that were actually explicated by the scientific experts of the field, implies all the scientific knowledge about it that the general pattern exhibits. This epistemic reason is sufficient to reject a conceptual identification of phenomena with patterns.

If ontic structuralists are right, then phenomena are, of course, structures, too. In this case, our description in our science textbooks of the phenomenon might be very different from its actual structural constitution, due to different scalings and pragmatic concerns. Furthermore, in this case, it seems, a normative demarcation between phenomena and non-phenomenal structures can possibly be made precise by the number of occurrences of a structure, or a vague version of it, in the actual world. The reasoning is the following: if the empirical world is a giant structure, then the accurate mathematical description of this structure allows for the search

[1] One of Bostrom's claims in this paper is that there is a non-negligible chance that we humans are living in an ancestor-simulation of a posthuman civilization. Such a civilization implies a level of technology, "where humankind has acquired most of the technological capabilities that one can currently show to be consistent with physical laws and with material and energy constraints" (p. 245).

of patterns in the empirical world itself and not only in descriptions of it. If certain patterns occur more often than others (however we define the rules for such a counting in detail), these may rightfully be regarded as phenomena. Such an approach would fit quite well to the realists' intuition that phenomena are somehow more often or more prominently manifested in the actual fabric of the real world.

However, in my view, we cannot rule out that the empirical world is not merely structural. Furthermore, *our* structural scientific description of it is (in most cases) very different from any actual structural constitution of the world due to epistemic limitations. In this case, there is no *correct* language for reference to parts of the empirical world. Then, our *descriptions* of phenomena, its concrete patterns, as well as the best possible description, its general pattern, are rather *utterances* about the empirical world than anything else. Under an utterance I understand a somehow grammatically formed unit from signs from a language, which also includes mathematics.

The difference between a proposition and an utterance is, in my terminology, that a proposition is grammatically more restricted. Since descriptions of phenomena are patterns, we do not have sufficient grammatical structure (*i.e.* subject; predicate; object) to classify it as a proposition.

Descriptively, a phenomenon is a property of the empirical world. We cannot claim anything more about phenomena. Note that we consider many singular events as phenomena that are worthy of scientific investigation, examples include the Holocaust or Bloop[1]. That is why a phenomenon is not well explicated by recurrence.

Why am I not an ontic structural realist?—The general pattern is restricted only to agent independent epistemic access; even if a super-AI formulates, for instance, an empirically perfectly adequate atomic model, we still do not know whether the Cartesian demon is feeding to us and to the AI the illusion of atoms. The general pattern is the best description of a phenomenon that can be gained via the epistemic access of any agent.

Phenomena and Natural Kinds

According to a realists view on objects to which scientific theories refer our best theories describe *natural kinds* or, at least, natural kinds exist and are to be discovered by science. As Bird and Tobin (2017) put it, this idea "corresponds to a grouping that reflects the structure of the natural world rather than the interests and actions of human beings". A possible relation to phenomena, as (Bogen and)

[1] *Bloop* was an unique and extremely powerful ultra-low-frequency underwater sound detected by the *U.S. National Oceanic and Atmospheric Administration* in 1997.

Woodward describe them, suggests itself: if there are natural kinds, then should not phenomena be features of natural kinds? This serves as a strategy to explain why they recur via corresponding concrete patterns in different sets of data. Then, a realist view on phenomena could be supported.

However, in my view we do not have sufficient reason to believe that any of our scientific descriptions refer to natural kinds, and we can also not defend the much weaker claim that there are natural kinds (due to the classical Cartesian-Bostrom argument). If one takes this stance, it is as unclear how to separate phenomenal features of the empirical world from other features of the empirical world, as it is unclear to separate alleged natural kinds from stipulated theoretical objects, which do not qualify as natural kinds. That is why, overall, ones metaphysical view towards phenomena is closely intertwined with ones view towards natural kinds. But as stated in this thesis, I defend a neo-Kantian anti-realistic position.

Assuming we are realists about natural kinds and about phenomena (such as Woodward), then how would these classes of things and properties be related to each other? An obvious difference between natural kinds and phenomena is that natural kinds are mostly discussed as kinds of *things* (*e.g.* lead), whereas phenomena are *properties* (*e.g.* melting point of lead). I cannot think of a good reason why a natural kind view and phenomena realism are not consistent. In the same way in which natural kind realism allows for an agent to have insufficient knowledge about a real natural kind, it allows for an agent to have insufficient knowledge about the phenomenon. If there are natural kinds, they must have properties. Are some or even all of these properties phenomena? All of them would be, because the selection criterion about what is real and what not is manifested in the distinction between natural kinds (*e.g.* lead) and classifications due to "human interests" (*e.g.* a flute made out of lead). That is why, according to realists, the melting point of lead is a phenomenon (since it is a property of the natural kind lead), but the sound that a lead flute makes is not (since a lead flute is not a natural kind). That is why natural kind realism is not only closely related to phenomena realism, but also implies it.

However, briefly summarised, the history of science (*i.e.* pessimistic meta-induction) and my intentional inclusion of social sciences (and other fields), make a realists' position about our best scientific knowledge seem extremely naive.

Phenomena as Explanatory Key Properties

The discussion of McAllister's account (cf. 6.1) points to an important question: what qualifies a specific deterministic function that can be detected in a set of data as a pattern corresponding to a phenomenon in (Bogen and) Woodward's sense? McAllister argues that no principal criterion could be given to identify a deterministic function as a pattern of this special kind. I summarise my answer briefly. The reason for this is that such a detailed account demands a vast study of the empirical phenomenon of science as a (human) activity. To answer more detailed which background assumptions, cognitive and sensory capabilities exactly guide our phenomenon selection is obviously a much larger task than this thesis is able to tackle. Aspects not only from general and special philosophy of science, but also from sociology and psychology of science, from neurology and various other fields need to be considered to tackle this enormous task.

I want to address this important point in a very general way: apart from any data, what qualifies a property of the world as a phenomenon? Obviously, every part of the empirical world has many different properties, but no scientist would be willing to say that every part of the empirical world exemplifies arbitrarily many phenomena. What scientists call phenomena are *interesting* properties of the empirical world in the sense that their explanation inherits a lot of explanatory power that can be applied to answer further related problems to the empirical part of the world in question. For example, if a biologist can explain albinism as being a specific gene defect, she does also find evidence for explaining the different human appearances by the colour of the hair and eyes by or partly by mutations at this specific chromosome. Albinism is, so to speak, an indicative extreme case and therefore a scientific phenomenon. Due to the strong contrast in their appearance, albinos play a noticeable role in parts of African sub-Saharan mythology and superstition.[1] This is a further clue in favour of the claim that albinism is something for which humans show a stronger demand for explanation than most other genetic mutation.

In the example of the melting point of lead this phenomenon is critical[2] in the sense that something more informative seems to be measured here compared to the measurement of other properties of lead under a specific arbitrary temperature. Why is it more interesting to investigate the exact melting point of lead and not, for example, its electrical conductivity under various different temperatures[3], or the sound that a solid lead recorder makes? The aim of such experiments is to characterise the metal and the critical phenomenon of the change of the aggregate

[1]For a survey that focuses on legislative and medical implications see Cruz-Inigo, Ladizinski and Sethi (2011).

[2]Under *critical phenomena* I understand phenomena that stand in connection to phase transitions as studied in statistical physics and introduced by, for example Sornette (2006, ch. 9). When a substance melts, this phenomenon influences the behaviour of it on many scales, as opposed to other properties of a substance that may be described and relevant either on only a macro level or on only a micro level. This scale invariance makes a critical phenomenon very interesting for physics.

[3]Note that this property becomes a phenomenon in the case of ultra cold superconductors.

state seems more informative for this purpose than, for example, the electrical conductivity at 300 degree Kelvin.

Overall, we can make the connection between phenomena selection and the explanatory power that a theory of the phenomenon would have over other observations. Does this mean that the notion of phenomena can somehow be reduced to a notion of explanation? At least not in a very direct way. Accounts of explanation try to explain or to explicate why a *theory* provides a good explanation *for* a phenomenon. I investigate what makes a phenomenon a phenomenon. So it seems a lot of additional thought has to put into the project of binding an account of phenomena closer to an account of explanation. However, I do not pursue this idea further, because the philosophical discussion about scientific explanation is already very intricate and vast on its own.

5.2 Relation between Patterns and Phenomena

Section Abstract

A phenomenon is a feature of the empirical world and is *described* by its general pattern. Two different phenomena have different general patterns and a general pattern can describe only one phenomenon; otherwise, the principle of identity of indiscernibles would be violated. General patterns and therefore phenomena are investigator-independent, because they can only be fully defined by an epistemically unbounded agent. An epistemically bounded investigator cannot fully define a phenomenon, but only some of its concrete patterns.

We are now prepared to explicate the relation between patterns and phenomena, which needs a further explanation that is based on the mere identification that (Bogen and) Woodward suggest. To do this, we have to put the theoretical pieces together that were developed so far. The relation between patterns and phenomena is a relation between phenomena and general patterns. Let us consider the example of albinism again (cf. 4.1). We need to clarify: what is the general pattern in this example? And what is the phenomenon? How are they related to each other?

What is the general pattern? As already mentioned, the two exemplary sets of data concerning an individual with albinism are the following. At first, the results from measurements of the colour of the skin, of the hair and of the eyes. At second, the records of a gene analysis. The general pattern is a class of concrete patterns and includes the ones recognised in the mathematical representations of each of these two sets of data plus possible patterns that we do not know about from possible measurements that we do not know. Due to the very different nature of the two sets of data, these two concrete patterns do structurally not have much in common. It does not make sense to look for any kind of structural similarity between these two concrete patterns. Thus, it is very hard to actually define this class, which is the general pattern of albinism, with appropriate generality.

What is the phenomenon? In the case of albinism pre-theoretical criteria based on visual impressions of human beings are available. The phenomenon of albinism is that there are some people having exceptionally light coloured hair and reddish eyes. But not for all phenomena in science are such simple criteria for observation with the naked eye available. Many interesting phenomena can be detected only by very specific measurements or experiments. Obviously, we distinguish between the phenomenon itself and its appearance to a layperson and to different scientists who are specialised in the respective field.

For the explication of the concepts of patterns and phenomena a philosophical decision has to be made as an assumption. To illustrate, let us assume human beings with albinism would have a slower rate of growth of the hair on their left leg directly under the knee, but no scientist ever discovered this fact. Let us assume this effect is so significant that it makes it possible to distinguish human beings with albinism and without it solely by this empirical criterion. Thus, clearly, in a time series of measurements of the hair length in the respective spot on the body a concrete pattern of albinism can be detected. Is this undiscovered fact of hair growth also part of the biological phenomenon of albinism?

A scientific theory is constructed to explain claims about phenomena; data is gathered to detect phenomena. Seen from this perspective, it seems reasonable to exclude facts from the concept of a phenomenon that are scientifically not yet discovered. The same holds for the concept of a general pattern. The peculiarity of this specific growth of leg hair from human beings with albinism—that I fictionally assume for toy reasons—could be a scientifically *uninteresting side effect* of albinism (whatever that means in detail). To say it in Aristotelian metaphysical terms: phenomena could have *accidental* effects as opposed to *substantial* ones. What is the different between albinism itself and properties that are caused by albinism? On the other hand, such a distinction between accidental and substantial concrete patterns forces scientists to update their understanding of a specific phenomenon or a specific general pattern with development of science, due to new discoveries (*e.g.* motion of stars in the night sky after discovering that they are large and very distant celestial bodies with own gravitational forces). This is not a reasonable approach to explicate phenomena or patterns. Scientists just say that there is the phenomenon of albinism and our knowledge of it does not have anything to do with it. There is no reasonable distinction between substantial and accidental effects of phenomena. Even phenomena realists have to agree that our knowledge about phenomena changes substantially over time with scientific progress (or decline).

I conclude that, in fact, it is most often not feasible for human scientists to give a full description of a phenomenon in science or of its general pattern in data, due to the notorious lack of knowledge, as illustrated for the simple example of albinism above. Therefore, it can only be assumed, but not constructively shown that the general pattern, which fully describes a phenomenon, is a class that is in principle definable, but not in practice by the human scientists' community. Only an epistemically unbounded agent with access to all empirical knowledge of the world may be able to give such a precise description of every phenomenon that is discussed in science.

A general pattern that is associated with a specific phenomenon, is the class of all concrete patterns, which can be recognised in all possibly producible data from the part of the world under investigation and which are associated with the respective phenomenon. This is, of course, a circular definition. How does a biologist know that different concrete patterns in very different data without any form of reasonable similarity between them can be used to detect the same phenomenon of albinism? That the appearance of a human being is related to its genes is a theoretical assumption itself and therefore, again, only the epistemically unbounded agent is able to describe a phenomenon to its full extent. It knows that every human being with albinism has the mentioned bright optical appearance and the genetic mutation of the 11$^{\text{th}}$ chromosome. All descriptions of a phenomenon from real scientists are more or less sketches from a specific scientific perspective. To say it in other words: the property that defines a general pattern as a class of concrete patterns is usually unknown and subject or even in the centre of the scientific speculation around the phenomenon itself.

Every possible concrete pattern that results from a measurement, if the phenomenon occurs, is element of the general pattern. This takes into consideration that patterns from confounding factors are excluded. Where should we conceptually locate the information about the experimental design and the measurement? As I elaborate on in 2.4, for an inference from data background assumptions of the data's origin are necessary. In the same way, a concrete pattern of a phenomenon can only be *discovered* by a scientist with the help of proper background assumptions about the origin of the data in which the pattern is detected. However, the general patter is not restricted to any specific form of measurement and is therefore extremely large and contains a lot of scientifically not pragmatically interesting concrete patterns.[1] The most important property of general patterns

[1] An analogy to DNA as an information carrier might be helpful. Bananas and humans have large parts of identical DNA, which is *junk*, when it comes to distinguishing different life forms on earth.

is that two different phenomena have two structurally different general patterns.

Can a phenomenon be *identified* with the respective general pattern? This question states a more refined version of the original position of (Bogen and) Woodward, which is the mere identification of phenomena with *patterns* (without a distinction between general and concrete patterns). An objection against the mere identification of phenomena with general patterns is that, from an ontological perspective, a pattern is rather a *description* of the phenomenon than the phenomenon itself. If we assume that data and patterns are always mathematical—and this is what I suggest—and phenomena can be identified with patterns, then it follows that phenomena have to be mathematical objects as well. Not many philosophers (of science) are willing to say that the biological phenomenon of albinism or the physical phenomenon of a melting point of a metal are mathematical objects or properties. They are just properties of the empirical world that may be describable by a mathematical model, they would state intuitively. And we do not possess sufficient knowledge about the empirical world to make the bold metaphysical claim that *everything* that causes phenomena is mathematical.

A phenomenon is (mathematically) *described* by its associated general pattern and is not identical to it. This description is a one-to-one relation; every phenomenon is associated with one and only one general pattern. But a general pattern cannot be constructed as an arbitrary set of concrete patterns. A phenomenon is an empirical feature of the world and its ontic properties therefore depend on our ontological attitude towards the empirical world. The crucial distinction at this point is that a phenomenon is a feature of the world, but patterns are (mathematical) properties of data, which are measurement results from some part of the empirical world and not a part of this part of the empirical world itself.

The assumption that all pattern recognition procedures that are used in science are reducible to the execution of mathematical algorithms is not as adventurous as it may sound prima facie to some. The application of formal systems in epistemology, and in particular in Bayesian epistemology, to model the rules of gaining knowledge and justified beliefs shows analogies to this assumption: reasoning is modelled by a formal calculus. Why then should it not be possible to model human beings' recognition of patterns in photographs, sound tracks or texts with the formal calculus, which is mathematics?

With this overall picture the discussed problems with the identification of phenomena with patterns (*i.e.* ontological difference and non-isomorphic patterns) vanish. That albinism is scientifically defined rather as a pattern in genes than by the optical appearance of a human being by biologists is not of relevance for the description of the phenomenon of albinism. But in this picture the crucial distinction between the general pattern, its concrete patterns and the phenomenon itself must be clear.

To summarise, the relation between phenomena and patterns in data is explained as a one-to-one description, but the concept of general patterns and its representations needed to be introduced. General patterns are classes of its concrete patterns in the data. Some of these representations may be known pretheoretically and some are revealed only after further scientific insights. Anyhow, due to the notorious lack of knowledge about all possible concrete patterns in all possible data by the actual scientific community, a phenomenon, or more precisely: the general pattern that describes it, can in many cases virtually not be defined by a scientist or a group of scientists.

5.3 Supervenience and Representation

Section Abstract

A phenomenon is not supervened by its pattern(s), because we have epistemic access to and a reference language for only patterns, but not for the world *an sich*. Phenomena are not represented by its corresponding pattern(s), because the patterns are all there is to describe phenomena without us having any epistemic access to a possibly real target system in this alleged relation of representation. The most adequate view is: a phenomenon is described by its general pattern and patterns that correspond to phenomena are what is scientifically represented by models.

A common way out of the ontological problem of the phenomena/pattern demarcation is the claim that a phenomenon is supervened by its corresponding pattern. Supervenience has its roots in the philosophy of mind and bridges the alleged ontological difference between a physical brain process and the corresponding mind process, like having a feeling or a though. In my terminology this idea might be adapted in the way that as long as a phenomenon, in an alternate world, would be different, its pattern would be different, as well, but phenomena and its patterns can still fall into different ontological classes.

If lead would have a different melting point, then this would clearly show in the patterns in measurement results from respective experiments. It is inherently true that, if the phenomenon would be somehow different, then its general pattern would be different as well. However, the problem with a supervenience explanation of the phenomenon/pattern relation is that we do not have epistemic access to both sides of the supervenience bridge. The problem with the mind/body duality is that we can physically measure the brain processes on the one hand, but we are also absolutely certain that at least we have conscious feelings. For the phenomenon/pattern relation, we can measure the relevant part of the empirical world under investigation to observe its structural properties, which are the patterns, but we are completely blind to the other side of the bridge. We simply cannot know what causes the patterns in the data. In other words, we have a well established language for the patterns of this world, but we cannot talk about the empirical world *an sich*, due to the lack of epistemic access. We cannot make any claim about the empirical world itself; we can refer to only empirical adequateness of our best theories in van Fraassen's (1980) sense.

That is why supervenience does not seem to provide a satisfying solution. We would need to stipulate a strong form of realism that implies an ontology of phenomena with the exact same demarcation lines between phenomenon A and phenomenon B that the patterns in data suggest. This is exactly the assumption that I reject throughout this thesis.

Another idea to explain the relation between phenomena and patterns is representation, however we want to explicate it in detail.[1] As Frigg and Nguyen point out, there are many philosophical explicative accounts for representation. Since I am mostly concerned about phenomena and patterns in scientific discourse, I focus on *scientific* representation, which is the representation of a *target system* by a *model*. My notion of a scientific model here is as broad as possible (cf. Frigg and Hartmann 2017), which implies that the notion of scientific representation is very broad as well.

Is representation a valid solution? I do not think so. The reason is that representation rather refers to, roughly speaking, a relation in which an object, system, proposition or property is represented by another object or proposition based on its observable features.

What exactly characterises the nature of representation is subject of the discussion about scientific representation? Common candidates are similarity (Giere 1988; 2004; 2010), morphisms between structures (Ubbink 1960; van Fraasen 1980, 2008; Bueno and French 2011), inferential purposes (Huges 1997; Suárez 2004, 2015; Contessa 2007), fictionalism (Frigg 2010a, 2010b; Godfrey-Smith 2006) and representation-as (Goodman 1976; Elgin 2010). Be that as it may, the crucial aspect is that all accounts refer to a target system or object on one side of the relation of representation.

The crucial point is that patterns play, in some sense, a more important role for phenomena than representation does for what it represents. What is albinism, if the phenotype and the genotype would only be representations of it? Without the patterns it corresponds to, albinism would be an epistemically empty notion, because there is nothing we can refer to except its patterns. The reason for this

[1] Scientific representation received a lot of specific philosophical attention from the 1980s on until now. I refer to Boesch's (2015), and Frigg and Nguyen's (2016) encyclopaedic surveys.

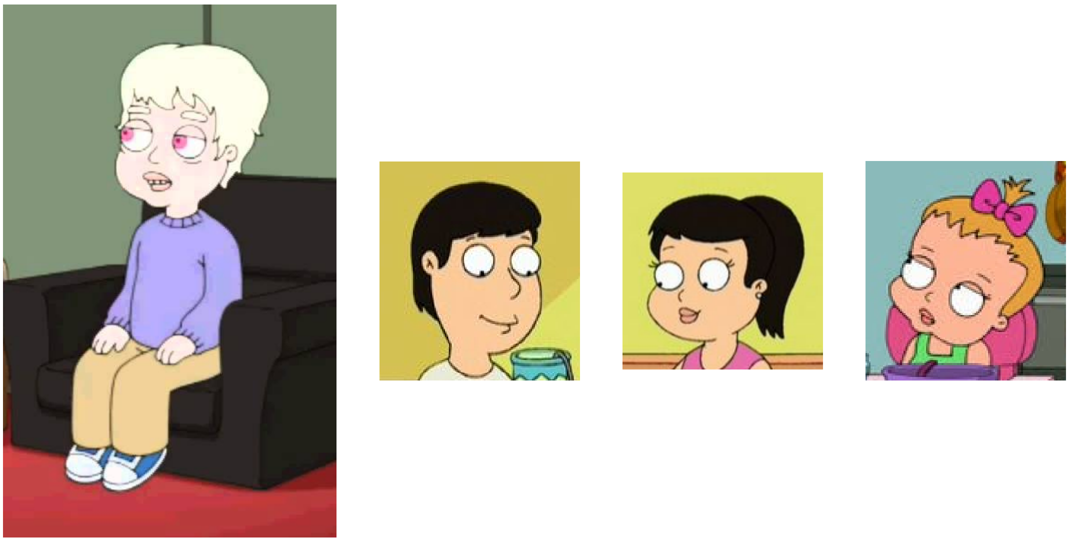


Figure 5.1: Left: Representation of a child with albinism. Note that, due to the cartoon character, the phenomenological concrete pattern of albinism is represented, but not exemplified as in the photograph 4.1, right side (p. 103). Right: Other cartoon children from same cartoonist for stylistic comparison. (Source: screenshot cuts from TV cartoon series Family Guy)

lies in the very nature of the concept of phenomena: it is a pragmatic and anti-realistic notion and we can refer only to specific descriptions of a phenomenon, but in no reasonable sense to a phenomenon itself. We can reasonably assume that there is something that causes the observation of a phenomenon (*e.g.* a certain part of real matter; Descarte's demon; a part of the program code from Bostrom's simulation), but we cannot make any claims about it, since for that we have to have direct epistemic access to it. In summary, a phenomenon is described by its general pattern and this general pattern, a subset of it or one concrete pattern can be represented by a model. Figure 5.1 shows an example of a representative model of albinism by a cartoonist impression. Figure 4.1, right side (p. 103) illustrates the respective concrete pattern.

In the case of scientific representation the representing models are empirically meaningful on their own and we can understand their empirical meaning quite isolated from the actual representation target. This is not so much the case for patterns, which receive their only empirical meaning (in addition to the mere mathematical structure, which they are) by stipulating the existence of the phenomenon. Here is an example to make more precise how I use '(empirical) meaning': ether theory is an attempt to represent the laws for the propagation of light and the theory turned out to be empirically wrong. However, the theory and the model of a cosmos with ether is still meaningful, because we roughly understand how the phenomenon of light propagation is represented by this model, even if it fails to represent the patterns of light propagation in some scientifically crucial regards.

If we assume that there is in fact no extrasolar planet in Apel's (2011) example, then the detected pattern of a sine curve is still a mathematical object, but we have to look for another interpretation, to give it any empirical meaning. In other words, representing models are, in contrast to mere patterns in data, empirically meaningful in isolation because we can understand them as being embedded in a *broader semantic context* than possibly rather arbitrary patterns.

Furthermore, as the example in figure 5.1 also exemplifies, a representation is in some way distorted in comparison to the pattern. However, one explicative account of representation comes close to the phenomena/pattern relation. Hughes (1997) claims that representation by physical models *denotes* their targets, which then serve for demonstrational and interpretational purposes. This idea of representation comes very close to my account of the phenomena/pattern relation. The difference, however, is that Hughes' models are designed to be epistemically very accessible (*e.g.* Bohr's atomic model), whereas a general pattern and also many concrete patterns can be very complex and cannot be used by a human agent to make inferences. Additionally, as Frigg and Nguyen point out Hughes' account "is unsatisfactory because it ultimately remains unclear what allows scientists to use a model to draw inferences about the target" (sect. 5.1). That is why I cannot use Hughes' account to tie the problem of representation any closer to the problem of the phenomena/pattern relation.

5.4 Are Phenomena Theory-Laden?

Section Abstract

Phenomena are nor theory-laden in a circular sense, neither do they have to be real in the realists' sense. Scientists share basic assumptions, interests and aesthetic criteria about the part of the empirical world under investigation that leads them to agree on whether patterns in data correspond to a phenomenon or not. Phenomena are not theory-laden, because phenomena selection is guided by a scientists' community's whole body of shared theoretical assumptions and also non-theoretical criteria.

An observation is theory-laden, if theoretical background knowledge is necessary to infer from the data evidence for or against the hypothesis in question. Woodward (and Bogen) refer to the notion of theory-ladenness often. Bogen describes it:

> For example we can challenge the use of a thermometer reading, e, to support a description, prediction, or explanation of a patient's temperature, t, by challenging theoretical claims, C, having to do with whether a reading from a thermometer like this one, applied in the same way under similar conditions, should indicate the patient's temperature well enough to count in favour of or

> against t. At least some of the $C[\,]$will be such that regardless of whether an investigator explicitly endorses, or is even aware of them, her use of e would be undermined by their falsity. All observations and uses of observations evidence are theory laden in this sense. (2017, sect. 4)

Since I believe that our phenomena selection is based on a vast body of shared background assumptions, the notion of theory-ladenness is interesting to get a clearer grasp of the topic. Regarding phenomena and the discussion of theory-ladenness, Bogen and Woodward (1988) write that those are “recent attempts to cast doubt on the possibility of objective, non-circular tests of competing theories” (p. 304). As phenomena realists, they, of course, argue against any theory-ladenness of phenomena selection.

Even if our phenomena selection is based on shared background assumptions, does this imply the circularity to which Bogen and Woodward refer? No, it does not. We need a lot of theoretical background assumptions to understand that a DNA record has something to do with how a person’s hair and skin appears. But this does not imply that a detected genotype of albinism corresponds with the phenotype of albinism. According to my account of phenomena, everything that we can investigate about a phenomenon is its general pattern. The only relevant scientific test with relation to albinism is whether the concrete patterns, which are elements of the general patterns, are correct in the sense that a person with albinism exemplifies them. If it turns out that a certain concrete pattern is not shown in all measurements of persons with albinism, biologists have to reconsider their notion of albinism, which is a selection of the known concrete patterns of albinism’s general pattern.

The reference to circularity is in general not substantial, because every human proposition is based on various sorts of assumptions and capabilities. Another, more technical example are LIGO detectors for the observation of gravitational waves. With them, gravitational waves can be detected in a non-circular way despite all the theoretical knowledge that is necessary to interpret the data (*e.g.* laser beam propagates with speed of light; tectonic activity shown in data). The important aspect is that the occurrence of gravitational waves and the theoretical knowledge that is necessary to interpret the data are to a crucial amount theoretically independent—at least according to today’s relevant physical knowledge that was considered at designing LIGO. In a trivial sense every observation is theory-laden (except for the claim that I have a certain mental state, as Descartes already pointed out), but the non-trivial aspect of theory-ladenness is whether it undermines the role that empirical evidence plays in scientific reasoning, which would be manifested by occurrences of circularities regarding the actual evidential support of a hypothesis.

In the same way as the LIGO example Bogen's initial example is not circular, because our background assumptions imply that a feverish patient shows a temperature $> 37°$ on a proper thermometer and there is no need to commit ourselves to any further realistic interpretation.[1] The thermometer is proper if our best theoretical knowledge is applied to make it work (*e.g.* empirically often corroborated and never falsified heat expansion of mercury) and the risk for false measurement results is minimised by the application of this best knowledge.

Another problem with the notion of theory-ladenness is its reference to *theoretical* assumptions. Regardless of whether you prefer a *syntactic* or a *semantic* view towards scientific theories, not every criterion for phenomenon selection can rightfully be called theoretical. *Aestetical* criteria or human cognitive and sensory capabilities play a profound role, as the exmple of albinism exemplifies. Furthermore, pragmatic concerns are very important in science—how does it help us in our society to understand the phenomenon?

One may oppose that all these criteria can be explicated in theoretical terms in principle. I agree, but we do not now these explications and otherwise there would be no need for explicative philosophy.[2] Therefore, it does not seem right to me to claim that our phenomena selection is *theory-laden*, it is rather laden with all sorts of human interests and ideas.

5.5 Evidence, Phenomena and Patterns

Section Abstract

The difference between phenomena descriptions (*i.e.* general patterns) and evidence is that evidence shows in many cases only one concrete pattern but not the whole general pattern. Like phenomena, evidence has to be epistemically accessible to the scientific agent. Evidence plays the role of confirmation, corroboration or falsification in the epistemic process of scientific inference and is therefore formulated in terms according to the hypothesis in question. Phenomena are formulated in more general terms of background assumptions. That is why the notions of evidence and of phenomena are different.

The concept of *evidence* plays an important role in various epistemological discussions, in particular the discussions on confirmation[3], explanation and explanatory

[1]One may object that we know that fever has the aim to kill viruses by overheating them or the like. *Heat* is a very successful pragmatic notion and we should use it. However, Descartes' demon may have planted the idea of actual heat very thoroughly.

[2]Explications of relevant but unprecise concepts is one philosophical task.

[3]I refer to Crupi's (2015) encyclopaedia entry for a comprehensive historical introduction into confirmation theory and my discussion concerning confirmation theory is in most parts based on his exposition of the topic.

strength[1]. Due to the general aim of this thesis, I do not accept the notion of evidence as unproblematically given. Since we want to explicate the distinction and relation between phenomena and patterns in data, it is helpful to this discussion to make clear, how evidence fits into this picture. However, most authors, who will be further specified below and contribute to the mentioned discussions do not put a lot of effort into explicating the notion of evidence; they rate this notion as unproblematically given.

As I explain more thoroughly in another paper (2018), for both Bayesian and deductive explications of confirmation, explanation and explanatory power, a distinction between data on the one hand and evidence on the other hand is important. Evidence has to be propositional to make logical (probabilistic or deductive) inferences from it possible. Data is often given as a vast amount of numbers or pixels or the like. Data cannot enter a propositional logical calculus in its original raw form, instead, a presupposed propositional evidence has to be confirmed by the data.

In this section, I, firstly, present a brief history of the common understanding of the notion of evidence among philosophers. On the background of this history I can support my claim that a distinction between data and evidence is neglected by the vast majority of influential authors. Secondly, I elaborate on my concept of evidence and how it is related to data. Thirdly, I discuss in which regard exactly my concept of evidence is problematic for the mentioned historical notions of evidence and in which it is unproblematic. Lastly, I describe how exactly evidence are related to phenomena and patterns.

Hempel and Oppenheim (1948) state in their introductory essay on the deductive-nomological model ('DN model') that under the explanandum regarding a scientific explanation they understand "the sentence describing the phenomenon to be explained (not that phenomenon itself); by the explanans, the class of those sentences which are adduced to account for the phenomenon" (p. 137). Thus, their *logic of explanation* is in fact a logic, which is a purely syntactical calculus on logical sentences.

In my view, a logical sentence that plays the role of the explanans in an application of the DN model is in many realistic cases rather a certain description of observed data and not the observed data itself. This idea becomes more apparent by investigating the following simple examples about black ravens and a patient with measles. Figure 5.2 shows possible observation data according to the three explananda black raven, red rash, gravitational waves signal (only a part of the actual data), which would have to be expressed in purely logical terms. These im-

[1]For a brief historical discussion of Bayesian explanatory strength see the introductions of Schupbach and Sprenger (2011) and Schupbach (2011).

time (seconds)	strain ∗ 1.e21
2.500000000000000000e-01	2.454791884395226415e-02
2.500610351562500000e-01	1.529268268197186628e-02
2.501220703125000000e-01	6.372337209458739903e-03
2.501831054687500000e-01	-2.075009545624776318e-03
⋮	⋮

Figure 5.2: Examples of data (Ströing 2018)

ages may still not represent the actual observations perfectly; for technical reasons we are restricted to a finite number of colours and pixels, these pictures are not animated videos and so on.

The point that I want to raise is that we have to *interpret* the data, that is, we have to infer an explanandum from a set of data *before* we can consider an explanation. Why can a set of data not be the explanandum in the DN model? An explanation can explain only a very specific aspect of the data (*e.g.* why the raven's feathers are black), but most information of the data is necessarily ignored (*e.g.* the contrasts of the image). In some cases of inference from data to evidence, the irrelevant aspects of the data have presupposed features and to one sort of these irrelevant aspects is usually referred to as *noise*.

However, for Hempel and Oppenheim (1948) evidence comes into the picture to support or confirm an explanans:

> The sentences constituting the explanans must be true. That in a sound explanation, the statements constituting the explanans have to satisfy some condition of factual correctness is obvious. But it might seem more appropriate to stipulate that *the explanans has to be highly confirmed by all the relevant evidence available* rather than that it should be true. This stipulation however, leads to awkward consequences. Suppose that a certain phenomenon was explained at an earlier stage of science, by means of an *explanans which was well supported by the evidence* then at hand, but which had been highly disconfirmed by more recent empirical findings. In such a case, we would have to say that originally the explanatory account was a correct explanation, but that it ceased to be one later, when unfavorable evidence was discovered. This does not appear to accord with sound common usage, which directs us to say that on the basis of the limited initial evidence, the truth of the explanans, and thus the soundness of the explanation, had been quite probable, but that the ampler evidence now available made it highly probable that the explanans was not true, and hence that the account in question was not—and had never been—a correct explanation (...) (p. 137–8, my emphasis)

Later in the text they refer to evidence in the context of *confirmation*:

> The requirement of truth for laws has the consequence that a given empirical statement S can never be definitely known to be a law; for the sentence

> affirming the truth of S is logically equivalent with S and is therefore capable only of acquiring a more or less high probability, or *degree of confirmation, relatively to the experimental evidence* available at any given time. (...) (footnote 18a, p. 152–3, my emphasis)

Confirmation is a well explicated notion for Hempel in his work (1943; 1945a; 1945b) in which evidence is modelled by a logical sentence as well. At no point in their text Hempel and Oppenheim identify evidence with an explanandum in the DN model. But since evidence in the model seem to be logical sentences, the difference between an explanandum and evidence is the role they play in a specific case of scientific explanation rather than any metaphysical aspect. But this implies that the discussed problem for the explanandum applies to evidence, too: we need to distinguish between some evidence and the observed data from which this evidence is inferred from.

Why do I focus so much on Hempel's notions of evidence and explanandum? Influential works from other authors, most importantly Popper (1935) and Carnap (1966), use these notions in very similar ways without focussing much on specific explications for them.[1] Hempel serves as a proxy for *modern classics* of confirmation and explanation theory, due to the noticeable influence he seemingly had, or at least due to the strong commonalities of his conceptual ideas about this subject with the other mentioned authors.

Horwich (1982), in his Bayesian essay on scientific knowledge, subjective probabilities and evidence, identifies evidence with sets of data. Note that he introduces his text as being about "particularly the concept of evidence" (p. 1) and titles a chapter 'Evidence' (p. 118–129). Figure 5.3 illustrates his concept of evidence by the use of an example. Horwich's text is a further example of the negligence of the distinction between data and evidence.

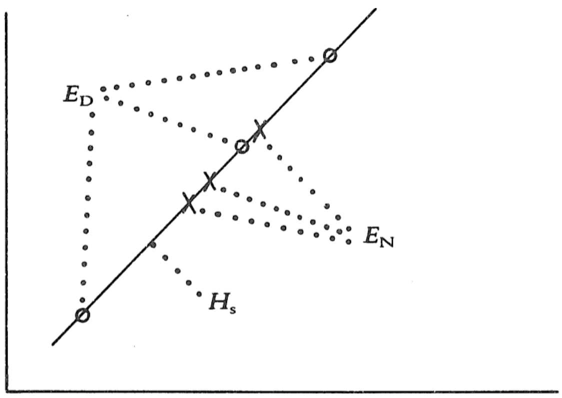


Figure 5.3: Illustration from Horwich (p. 119) for more diverse data (E_D) and narrow data (E_N). In fact, as the labels 'E_D' and 'E_N' indicate, he refers to these data as *evidence*.

Apart from this section, I use only my notion of evidence throughout this thesis

[1] Popper and Carnap provide carefully edited subject indices at the end of their books. Carnap does not include "evidence" or some other reference to it in the subject index. Popper refers more explicative to evidencel only in a later appendix of his book (1959, appendix IX), which consists of a reprinted series of papers (1954; 1957; 1958) with minor corrections.

and not the one from the influential authors that I mention above. Interestingly, whereas scientists often use phrases like 'the data shows', 'evidence' is a term that occurs more prominently in reasoning by lawyers in court. This may hint to the fact that in science it is more common to apply statistical routines on the data,[1] whereas lawyers are usually not very familiar with statistical methods.

What can we say about the relation between evidence and phenomena? If evidence play the role of explananda in explanatory relations and if scientific theories "explain claims about phenomena", then, it seems, evidence and phenomena must at least be closely related. Is there any substantial difference between these two notions?

Evidence are propositional in the sense that it, as opposed to data, must be epistemically accessible to a human agent—this holds at least for the *classical* idea of science as an endeavour for the creative human mind. A much larger class of possible evidence is epistemically accessible to an AI scientific agent in principle. However, even an AI or alien scientific agent, a neural network for instance, implies some narrowing structural criteria for whatever can play the role of evidence according to its methods of inference. Figure 5.4 illustrates an example of scientific inference by a human scientific agent via a Bayesian network.

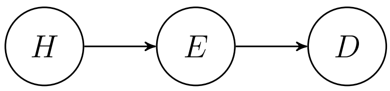


Figure 5.4: Bayesian network for
H : *the earth rotates around the sun*,
E : *pattern of an elliptic orbit* (basic geometry),
D : *records of the sun position* (lots of numbers).
(Ströing 2018)

In the following, I provide a more detailed analysis of the different roles of evidence, phenomena and patterns in the epistemic process of scientific inference.

The pragmatic description of a phenomenon, which is its most discussed concrete pattern(s), is based on the full body of the scientist's background assumptions and knowledge. It plays the role of describing epistemically crucial properties in this theoretical and sensorial framework. Evidence, on the other hand, is formulated in the terms that are linguistically predetermined by the hypothesis in question. There is no concept of evidence, when there is no hypothesis in question for or against which the evidence has to be formulated.

A phenomenon is fully described by its general pattern and a hypothesis is a claim about one or more phenomena. A hypothesis cannot be a claim about a feature of the empirical world that is not considered as a phenomenon due to the points raised by my descriptive account of phenomena; non-phenomenal features

[1] However, note that the pragmatic quality of applied statistics varies greatly among different scientific fields and working groups. Derived from personal experience I dare to claim that statistical tests are often conducted without proper mathematical understanding of them in, for instance, some branches of social sciences.

are not considered worth investigating. Otherwise, we are forced to accept them as phenomena.

Since evidence plays the role to confirm, corroborate or falsify a hypothesis, it is something different than a pattern. More specifically, the class of all patterns (see chapter 4 for explication and discussion of patterns) is very broad and a pattern can occur in complete epistemic isolation, meaning that it does not have to play any role in any inferential process of science (*e.g.* an aesthetic pattern for decorative purpose). This implies that, for instance, we can observe Venus' movement on the horizon as a pattern, but remain completely puzzled by it, if we do not have any theory of planetary motion at hand. Quite contrary to this, evidence is always evidence *for* or *against* a hypothesis. That is why evidence has to be formulated in the proper linguistic terms of the hypothesis. An example: if the hypothesis in question is that Venus revolves around the sun in an elliptic orbit, then the evidence cannot be only the observation of Venus' motion seen from an arbitrary point of the earth's surface. We need the appropriate compilation of information to accept it as evidence or counterevidence. In this case, this includes the position of observation on earth and knowledge about the earth's own motion, but it does not include the observers aesthetic opinion about nail polish or the latest winning Lotto numbers at the observation date.

Chapter 6

Replies to some Articles about Phenomena

I comment on some selected articles about phenomena in science.

6.1 McAllister's Criticism and the Example of Albinism?

Section Abstract

McAllister (1997) criticises (Bogen and) Woodward's account of patterns in scientific data. He stresses that every set of scientific data can be decomposed into a sum of an arbitrarily chosen pattern and the remaining noise. I reject his objection by stating that his accounts of patterns in data, as well as of empirical phenomena in science are generally misleading. Patterns that correspond to phenomena are a specific subset of all arbitrarily selectable patterns, but this subset cannot be fully defined in isolation of our shared background assumptions. Patterns that correspond to scientific phenomena can be non-discrete mathematical objects that express vagueness or uncertainties. An agent's knowledge of one phenomenon's concrete pattern may be insufficient to describe it precisely. Investigators make decisions on whether a pattern can be detected in a set of data, but the pattern itself is investigator independent.

McAllister (1997) criticises (Bogen and) Woodward's account of the relation between patterns and phenomena in a very general way. I start with reconstructing his arguments.

(Bogen and) Woodward claim(s) that phenomena are real in the ontological sense. McAllister stresses the conclusion that, given the relatively few scientifically discussed phenomena, a list of all the phenomena of the empirical world needs not only to be finite, but also not very long. Brown (1994, p. 124–5) stresses exactly this point, too. This means that there are only relatively few phenomena in contrast to the infinite variety of all the possible patterns in all the possible sets of scientific data. Different sets of scientific data are as various as possible measurement designs and further external influences on the experiments are, including noise. The relatively low number of scientific phenomena follows from the reality assumption regarding them and the philosophical desire to deal with an ontologically sparse universe. If a list of all phenomena of the empirical world would be infinitely long, then the very fundamental naturalistic claim that the empirical world can be, at least approximately, described by a finite catalogue of *natural laws*—including physical, psychological and social laws—seems misleading.[1] From the reality assumption also follows that phenomena, and therefore patterns, must be independent from an investigator.

McAllister's criticism is based on the idea that every set of scientific data can generally be described by a deterministic function plus some random noise, whereas there are infinitely many options for the deterministic function and the noise part. In more rigorous mathematical terms, a *deterministic* function

$$f : I \to E$$

is a function that maps every data point indexed by an $i \in I$ to a value from the set of possible data results E without any modelling of probabilities or inherent vagueness. McAllister identifies patterns with the deterministic part of this sum, which is in accordance with (Bogen and) Woodward's wording that "phenomena are patterns" and that I criticise throughout this thesis. McAllister's idea that every deterministic function counts as a pattern is also in accordance with my definition of a pattern (see chapter 4).

He infers from this description that patterns are far from being only a few very distinct cases in actual scientific data and could therefore be identified with the very few phenomena of the empirical world. It is rather the case that infinitely many patterns can be detected in data and therefore infinitely many more or less resembling patterns that each correspond to a certain phenomenon. Even worse, infinitely many patterns can be detected in only one set of data by properly adjusting the noise term. In an easy example one can always find a slightly different deterministic function and tolerate a higher proportion of noise in the sum to describe one and the same set of scientific data.

[1]For a further discussion of the assumption of only relatively few phenomena see Brown (ch. 7) and McAllister (p. 218).

This all leads to the problem that either a phenomenon cannot correspond to only one single pattern, or that there are as many different phenomena (or 'structures' as McAllister suggests in a later text 2011) in the empirical world than different patterns can be found in scientific data. In this later case the notion of phenomena would be deflated in the sense that one says nothing more about a pattern than that it is a pattern. It remains an open question in (Bogen and) Woodward's account, how either many patterns in data disqualify for corresponding to a phenomenon or how different patterns can correspond to one and the same phenomenon.

In fact, *noise* is a very common concept in statistical modelling and is in some contexts precisely defined (see appendix A.1 for details). In the context of mathematically explicated patterns in time series noise is usually modelled as a specific stochastic process, for example a Brownian motion. These precise definitions may count as a counterexample to McAllister's description of patterns and noise in data, since a precise definition of noise leads to a unique solution or at least to a vastly restricted space of solutions for the deterministic function in the decomposition of data. However, examples to strengthen McAllister's position regarding this objection can easily be found by questioning the appropriateness of a specific definition of noise, in particular for an explanation of patterns in science in more general without a restriction to the established mathematical framworks of statistical modelling.

McAllister (1997) suggests the solution that phenomena are in fact not investigator independent, as (Bogen and) Woodward imply, since the investigators are the ones who decide which level of noise is tolerable to detect a phenomenon via a pattern in data. I believe this critique is not very substantial, because McAllister's interpretation of patterns is a misconception. As he points out, (Bogen and) Woodward are realists. For them phenomena are features of the empirical world that can be studied via data. It seems, for them, the notion of patterns has a Platonic feature to it in the sense that a phenomenon has a single corresponding ideal pattern (or a finite list of ideal concrete patterns), which can be detected in data in approximation to find evidence of the phenomenon in question. Measurement errors, side effects, insufficiently isolated objects under investigation and other technical circumstances make gathered data most often non-ideal in this Platonic sense. Conclusively, almost every pattern that corresponds to a phenomenon and occurs in an actual set of data is non-ideal in this sense as well.

To further illustrate my criticism of McAllister's account, note the following implication: if two different sets of data from two different measurements show the same pattern in (Bogen and) Woodward's terminology with which the same phenomenon is detected, these two patterns may still be different in McAllister's understanding of differing patterns. A simple example for this could be two measurement runs to detect the melting point of lead. If two time series of 100 measured melting points in degree Kelvin are given, then the two sets of data are very likely not completely identical. But every physicist is willing to admit that the same pattern was detected by the two measurement runs under exactly the same experimental conditions. The crucial point is that the pattern of a phenomenon does not have to be a completely deterministic function. In my view, as well as for—seemingly—(Bogen and) Woodward, a certain level of vagueness, which may even be mathematically expressed in statistical theory, can be a genuine part of a concrete pattern of a phenomenon.

A specific deterministic mathematical explication of a pattern, as McAllister suggests, that was detected in real data is not the pattern (Bogen and) Woodward refer to. It is undoubtedly mathematically very similar to the ideal pattern corresponding to the specific phenomenon, but in fact only similar in some of its mathematical properties. In McAllister's terminology every deterministic function falls under the term 'pattern', whereas for (Bogen and) Woodward only a few ideal versions, from which some deterministic functions are approximations of, fall under this label. This is the confusing misconception in McAllister's account. In my account, every deterministic function counts as a pattern, but only relatively few possible patterns in data count as patterns that correspond to phenomena. As elaborated on in chapter 5, the body of our shared background assumptions, cognitive and sensory capabilities determine our phenomenon selection and not the fact that there are some patterns in data. Most patterns in scientific data are irrelevant for scientific investigations of phenomena.

What then are patterns in data that correspond to phenomena? They are the mathematical correspondences of the empirical phenomena. And as such they have a precise mathematical form, but it is another question whether a scientific agent actually explicates this form or not. I elaborate on this in chapter 4. A scientist may even define a pattern sharply, but this definition can simply be wrong, which may often be the case in actual scientific practice. Furthermore, we know that an animal with albinism has white hair and reddish eyes, but we do not fully know what biological abnormalities it has apart from these obvious facts. In the same way, patterns corresponding to a phenomenon have a precise form, but may be only defined imprecisely. To say it in other words, phenomena are empirically real and patterns (that correspond to phenomena) are mathematical ideal objects, but knowledge about them must be gained and can be imprecise or wrong in both cases.

A precise definition does not imply discreteness of the mathematical object or determinism of some process, which is the pattern (*e.g.* the wave function in quantum mechanics is a non-deterministic pattern). A pattern can be defined as an analytic object, *i.e.* with continuities and uncountable sets, or an object in a probability space, *i.e.* with uncertainties, or with the use of some other mathematical version to model vagueness or uncertainties, such as fuzzy logic. For albinism, a corresponding concrete pattern is the gene defect in the *TYR* gene in the eleventh chromosome, but the exact form of the mutation may be unique for each individual with albinism. So, at this point the concrete pattern needs to be open for vagueness concerning the specific form of the gene defect.

But this vagueness is not the noise that we discussed above based on McAllister's description of patterns. Noise is the deviation of the data, or parts of it, from the concrete pattern that corresponds to the phenomenon in question. Noise is caused by measurement errors and disturbing influences to the measured part of the empirical world, whereas vagueness is a property of the concrete pattern and the phenomenon in question itself. This distinction is a fundamental one and does not depend on the knowledge of any scientist. This distinction is exactly what McAllister does not take into account accordingly. To say it in other words, two forms of uncertainties are involved at pattern detection in scientific data: one of it origins from the vagueness as a property of a concrete pattern itself, the other one is caused by the lack of experimental isolation of the phenomenon and other pragmatic imprecisions regarding the measurement procedure.

A criterion for a deterministic function in data to be a pattern in (Bogen and) Woodward's sense cannot easily be given in purely mathematical terms. The reason for this is that a pattern in this sense must correspond to a phenomenon. But a phenomenon can in many interesting cases not be characterised in terms of a developed theory and also not as a pattern in data from measurement results. A phenomenon in science is, such as the ancient Greek origin of the word indicates, a property of the empirical world that obtrudes itself to the observer by means of the observer's interests and epistemic capabilities. A melting point of a metal or the occurrence of albinism force us to find explanations before any theory construction has started and may be the motivation to start a theoretical endeavour to explain them. Measurement procedures are very uniquely designed for the explanation of specific phenomena and phenomena may in many cases not be fully understood with the available theories. Therefore, no *simple* mathematical criterion to identify patterns that correspond to a phenomenon is possible.[1]

McAllister reasons that, since different scientific investigators may not fully agree on the acceptable level of uncertainty in pattern recognition, the phenomenon

[1] An epistemically unbounded agent may be able to provide such a mathematical criterion in principle. It would imply a full description of human and, depending on its scope, non-human sensory capabilities and background assumptions.

itself depends on the investigator. If two scientists can disagree on the occurrence of a phenomenon based on analyses of the same set of data, then the phenomenon cannot be something real that simply is there or is not, he infers. Woodward's recent response (2011, sect. 5) is a very natural one from the realists' perspective: of course, investigators can disagree over the detection of a phenomenon in the same set of data, but one of them can simply be false in his inference. If a coin is tossed twenty times in a row and it lands head twelve times, then investigators may disagree whether the coin passed a statistic test of fairness or not. But this disagreement does not change anything about the phenomenon of a coin being fair (*i.e.* the occurrence of head in repeated tosses converges to 50%)[1]. Phenomena are real in this sense. The data is not our subject of study in science, but the empirical world is. A set of data may—and often does—give only insufficient evidence of the occurrence of a phenomenon. A phenomenon may not even be correctly described by science due to, for instance, bad noise filtering, but it may still exist.

However, the role of the noise needs to be explained. The ideal pattern in data for this phenomenon is that 50% of the tosses show a head as a result. The experiment of a coin toss can only be repeated finitely often. Even two tosses can fulfil the criterion to exemplify the ideal pattern, if a scientist decided that this evidence is sufficient (which would reveal a poor intuition regarding statistics). Therefore, an investigator's decision on whether a phenomenon is detected via a pattern in a set of data does not depend only on the set of data itself, but also on his personal attitude towards statistical confidence or on the number of data points. But this fact still does not change any properties of the general or the concrete pattern that corresponds to the phenomenon in question. It is not the pattern corresponding to a phenomenon that depends on an investigator, only the interpretation whether it is detected in a set of data or not does. What McAllister introduces as the noise is, in parts, the acceptable level of deviation from the ideal, but can also, in parts, be a vagueness that a phenomenon may naturally has, when it comes to detecting it in scientific data. For the first case, the occurrence of the noise does not have any metaphysical implications whatsoever. In the second case, it has, since here a pattern is assumed to be a non-discrete mathematical object.[2]

[1]This explanation of probability is a frequentist's one. Convergence is a mathematical concept applied to infinite series $(x_i)_{i\in\mathbb{N}}$ with, for example, $x_i \in \mathbb{R}$ for all $i \in \mathbb{N}$ and the metric $d : \mathbb{R}^2 \to \mathbb{R}_{\geq 0}, d(x,y) \mapsto |\, x - y\,|$. Therefore, this explanation of probabilities can be objected by stressing the problems with modelling finite empirical objects and time-frames with infinite mathematical models. However, my argumentation is open to alternative interpretations of probabilities and does not crucially depend on any specific choice in this regard.

[2]The uncertainty principle in quantum mechanics and its standard interpretation, as opposed to the Bohmian view, may serve as an example, here.

I introduced the term 'pattern similarity' (cf. 4.1) to denote concrete patterns that differ by mere simple mathematical transformations, in, for example, time series from asset price data in returns or, alternatively, in log-returns. McAllister's description of noise is not very much related to this concept, even if it may seem so. In cases of pattern similarity the patterns are not identical or in any way related to each other according to McAllister's account. He chooses the maximum possible version of unrelatedness among concrete patterns: every two patterns that are not identical mathematical objects, differ from each other in the same general way. To me it seems much more reasonable to distinguish between two rather similar concrete patterns (*e.g.* melting point in Celsius or in Kelvin) on the one hand, and very different concrete patterns (*e.g.* genotype and phenotype of albinism) on the other hand.

To sum up, McAllister's notion of noise is not problematic for (Bogen and) Woodward's concept of a pattern, as well as for my further specification of concrete and of the general patterns. In fact, it describes what I further specify as either vagueness or non-determinism of a pattern or the approximation of a pattern in data. Investigators decide whether a phenomenon can be detected via a pattern in a set of data, but all properties of a pattern are investigator independent. My criticism to McAllister is that he misinterpretes patterns that correspond to phenomena in one sense too strict as deterministic, discrete mathematical objects, but in another sense too broad as an arbitrary object so that every feature of the empirical world becomes a phenomenon. Her infers that the concept of a phenomenon is empty by stressing that there is no property of a pattern to qualify it as being a pattern that corresponds to a phenomenon. I believe that there is no simple mathematical property fulfilling this criterion, but there are real empirical phenomena that are selected on the basis of the agents' shared background assumptions, cognitive and sensory capabilities.

6.2 Glymour's Approach to Weaken the Distinction between Data and Phenomena is Misleading

Section Abstract

Glymour claims that the distinction between data and phenomena as unnecessary, since, in his view, it does not add anything to the distinction than concepts that are already known from basic statistics. This view is misleading, because, firstly, not all phenomena are recognised by applying explicated statistical routines and statistical pattern recognition does not well represent all pattern detection routines. Secondly, in such an account it needs to be explained why certain statistical routines for pattern detection are employed, whereas infinitely many others are not employed.

Glymour (2000) aims to challenge the distinction between data and phenomena, which is the essential distinction in (Bogen and) Woodward's account. The review of his position helps us to further sharpen our concepts of phenomena and data.

In a nutshell, his arguments are as follows: firstly (i), according to Bogen and Woodward, data and phenomena are distinguished ontologically and with respect to their epistemic roles in science. Secondly (ii), scientific data should be understood as statistical samples (*i.e.* one- or more-dimensional series of numbers $(r_t)_{t\in\{1,2,\ldots,N\}}, N \in \mathbb{N}, r_t \in \mathbb{R}$). At least, statistical samples are a common case of scientific data and can play the role of sufficiently representative example case for us. Thirdly (iii), patterns in data are nothing more than *statistical* features of these samples (*e.g.* autocorrelations) and do not differ from the data itself in respect to their "ontological and epistemological status." (p. 30) Therefore, Glymour states, "the distinction between data and phenomenon simply gives a new name to a distinction which is already deeply embedded in the literature on statistical inference." (p. 34)

I believe that the second point is too narrow and therefore inadequate to describe the general concept of data in science: in my view, data cannot be boiled down to statistical samples and from a philosophical analysis of statistical samples we cannot gain much sufficiently general knowledge about the concept of scientific data. Consider the simple example of recognising a specific animal in a picture. Please note that McAllister's (1997) example of data and patterns in it is one of a simple statistical sample with statistical features in it, as well.[1] But only a minor fraction of what counts as scientific inferences are applications of well explicated

[1]More specifically, see the introduced example of a pattern as a finitely described function $F : \mathbb{R} \to \mathbb{R}$ at pages 219–223, more specific $F(x) \mapsto (a \sin \omega x + b \cos \omega x) + R(x)$ with $a, b \in \mathbb{R}$, $\omega \in \mathbb{R}_{[0,2\pi]}$ and R describing the noise part.

routines of statistical pattern recognition. I claim that even if scientific inferences can be mathematically explicated in principle, the interesting aspect is what makes one pattern more interesting (*i.e.* corresponding to a phenomenon) than others.

The third point, I believe, is also misleading due to the fact that the general, as well as the concrete patterns in a set of data are ontologically not equal to the phenomena, as I elaborate on in 5.2. A property of the empirical world cannot be equal to a mathematical object or property. This is the same error that (Bogen and) Woodward make. And patterns are mathematical objects or properties and Glymour's restriction of data to statistical samples even strengthen this position. I reconstruct Glymour's arguments with comments on them more thoroughly.

For the second (ii) point, we want to introduce some basic vocabulary to discuss the arguments raised by Glymour with a sufficient precision. Glymour's aim for the use of these terms is to further specify data and phenomena on the basis of statistics to present an exemplary account for data and phenomena.[1] A *sample* is a series of real numbers or vectors of real numbers that is gathered by measuring in observational or experimental contexts. The *statistical structure of the sample*, *sample statistics* or *sample structure* are statistical properties of the sample as, for example, mean values, variances or correlations between different dimensions of or autocorrelations in the sample. A *population* is a set of empirical items that is subject to the statistical analysis. The *structure of a population* or *population structure* are structural properties of a population; a population structure is investigated with statistical analysis by gathering samples and specifying its sample structure.

Glymour speculates that statistical analyses present a common and useful example for scientific use of data. He follows as the third step (iii) in his argument concerning the epistemic status:

> If we are certain, at least in the relevant sense, of the observations comprising a data set, then the mean value of a variable[[2]] in the data, or its variance, the shape of the distribution, correlations between variable values, and so on, are no less certain. So sample statistics have the same epistemic status as the observation reports comprising the data in the sample. But it is precisely this sort of statistical feature of data sets that are explained by scientific theories. (Glymour 2000, p. 33)

As we will see in the following passages, this argument is the core of Glymour's critique on (Bogen and) Woodward's concept of data and phenomena: there would

[1] Glymour's use of statistical vocabulary, including 'sample', 'population', 'sample statistics', 'sample structure', 'population structure' is in fact not sufficiently well introduced in his article.

[2] Apart from Glymour's use I want to avoid the use of the term 'variable' since a set of data, *e.g.* a series of real numbers, does not have anything to do with variables. According to the strict use of vocabulary I want to follow, variables come into play only if some model for the data is introduced.

be no substantial epistemic difference between the set of data, if it is exemplified as a sample, and the sample statistics. This may be true from our perspective as well. Let us take a view on a very simple example of a sample being a set of numbers

$$X := \{2, 3, 3, 2, 4, 1\},$$

its mean

$$\mu(X) := \frac{\sum_{x_i \in X} x_i}{|X|} = \frac{15}{6} = 2.5,$$

its variance[1]

$$\begin{aligned} \sigma^2(X) := \mu([X \mathbin{.-} \mu(X)]^{.2}) &= \mu(\{-0.5, 0.5, 0.5, -0.5, 1.5, -1.5\}^{.2}) \\ &= \mu(\{0.25, 0.25, 0.25, 0.25, 2.25, 2.25\}) \\ &= \frac{5.5}{6} = 0.91\bar{6} \end{aligned}$$

and the largest number in it divided by 2

$$\lambda(X) := \frac{\max(X)}{2} = \frac{4}{2} = 2.$$

These numbers, 2.5, $0.91\bar{6}$ and 2, express something about the sample X, but what is expressed is only some arbitrarily chosen mathematical property of it out of infinitely many mathematical properties. I introduce the statistically uncommon $\lambda(X)$ to illustrate this arbitrariness. Of course, 2.5, $0.91\bar{6}$ and 2 do not have any special epistemic status above the sample X itself. From a mathematical perspective, these numbers plus their mathematical origin, which are the definitions of μ, σ^2 and λ, simply tell us *some* information about X.

But this is not the point of interest concerning the distinction between data and phenomena. The point of interest is, why is the mean $\mu(X)$ of specific interest concerning this sample (*e.g.* it scientifically describes the melting point)? Why are μ and σ^2 more interesting and common in statistical analysis than λ? The answers to these questions can by no means be given on grounds of only the sample X or some arbitrary sample statistics $\mu(X)$, $\sigma^2(X)$ or $\lambda(X)$. This is the same misleading view McAllister shows in his inquiry. Both, Glymour and McAllister

[1] I use non-standard operations on sets according to the following definitions:

$$\mathbin{.-} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n, n \in \mathbb{N}, \{r_1, r_2, ..., r_n\} \mathbin{.-} r \mapsto \{r_1 - r, r_2 - r, ..., r_n - r\},$$

$$^{.2} : \mathbb{R}^n \to \mathbb{R}^n, n \in \mathbb{N}, \{r_1, r_2, ..., r_n\}^{.2} \mapsto \{r_1^2, r_2^2, ..., r_n^2\}.$$

interpret phenomena as being detectable via purely mathematical properties of purely mathematical data. But in fact the central question for the explication of the notion of phenomena is why certain mathematical properties are more interesting to the scientists than other properties. And my answer to this question is that the shared body of background assumptions and human capabilities make a pattern to be a pattern of a phenomenon. Statistics is a suitable example to defend this view, since notions like variance σ^2 are to some extent chosen due to pragmatic concerns, but also to some extent arbitrarily (in particular the exponent $^2 : \mathbb{R} \to \mathbb{R}_{\geq 0}$).[1]

However, let us take a more detailed look at Glymour's idea how samples and sample statistics are related to data and phenomena:

> We can either take the distinction between data and phenomena to correspond exactly to the distinction between sample and population structure, or not. Suppose we take the distinction between data and phenomena to involve something over and above the distinction between sample and population structure. Then statistical inference procedures, and methodological justifications for them, will not require the distinction between data and phenomena, and hence the distinction will be unnecessary. Suppose we deny that the distinction between data and phenomena involves something over and above the distinction between sample and population structure, *i.e. we take the statistical structure of data samples, i.e. sample statistics, to correspond to data and population structure, i.e. population parameters, density functions and conditional independencies, to correspond to phenomena.* Then the distinction between data and phenomenon simply gives a new name to a distinction which is already deeply embedded in the literature on statistical inference. (Glymour 2000, p. 33–34, my emphasis)

Let us first discuss the position that I deny and which is discussed secondly in this quoted passage: phenomena are nothing more than the population structure. Concerning the highlighted (in italics) part, I do not see any reason to identify the data with a sample structure, that is: sample statistics; if any, the sample corresponds to the data and the sample structure corresponds to a pattern in this set of data. Note that it is our specific choice for what "structure" we look for in a set of data. But since 'population structure' is a term that denotes properties of a set of empirical items and is investigated via statistical analysis it does not seem to be very useful to grasp a better understanding about the distinction between data and phenomena after all. Glymour's philosophical analysis does not tell us more about our riddle of what exactly counts as a phenomenon. We need to further investigate what would count as a *population structure*.

[1] The variance operator is applied to measure a form of dispersiveness of a random variable or a time series. A measure with this aim should be positive but the square seems to be established for historical reasons and due to the pragmatic aim to provide an epistemically well accessible stochastic calculus.

According to the common use of terminology I understand under a population the set of empirical items under statistical analysis, but, on the other hand, Glymour implies that "[o]ne treats the sample as a sample from a *population of data* with a particular statistical structure" (Glymour (2000), p. 32, my emphasis). With this use of the terminology it is hard to see how the term 'population' is used and what is implied with it. If there is a population of data, then 'population' seems to denote something purely mathematical.

The firstly discussed option of a solution to the riddle what data and phenomena are in our statistical setup is the one that I prefer: the statistical objects introduced, as well as statistical inference, cannot fully describe what a phenomenon is in this setup. Glymour suggests that in this case the distinction between data and phenomena would be unnecessary, because "statistical inference procedures, and methodological justifications for them, will not require the distinction". He goes on about the proclaimed unnecessity of the distinction:

> The terminological reform is unnecessary and in some respects misleading, and hence should be avoided. Moreover, since on some statistical inference procedures, *e.g.* Bayesian scoring procedures, one infers directly from sample structure to theory, the distinction between data and phenomena will not play any essential role in these sorts of inferences or the justification of these inferential methods. (Glymour 2000, p. 34)

He deems the distinction between data and phenomena *unnecessary*, in case it involves something over and above the distinction between statistical samples and the population structure. But the point of interest for us is that scientific detection of phenomena cannot be exhaustively explained by the explicated body of stastitical pattern detection. There is more to most phenomena than an arbitrarily chosen pattern in data, as McAllister suggests, or theory-laden interpretation of some data. As I elaborate on in chapter 5, the distinction between phenomena and non-phenomenal features of the empirical world is based on the shared body of scientific background assumptions, as well as the agents' sensorial and cognitive capabilities.

6.3 Massimi's Kantian Approach to Phenomena

Section Abstract

Massimi stresses the important role that reliability of data plays for the defence of (Bogen and) Woodward's realists' notion of phenomena. I agree with her conclusion that reliability depends too much on the agents' understanding of involved causal mechanisms to defend phenomena realism by reference to it. In my view, her Kantian stance towards phenomena selection with reference to Kant and Suppes neglects other important influences for phenomena selection that are important for a descriptive account. These are non-Kantian background assumptions that are better described in Kuhnian terms and related approaches.

How do (Bogen and) Woodward defend their phenomenon realism? Massimi (2008; 2011) rightly focusses on "a distinct feature of reliabilism [which] is that it licenses theory-free data-to-phenomenon inferences" (2011, p. 103). This contrasts McAllister's observer-dependent notion of phenomena. What makes data reliable? Massimi cites Woodward (1989, p. 403–4): "in order for data to be reliable evidence for the existence of some phenomenon (...), it is neither necessary nor sufficient that one possesses a detailed explanation of the data in terms of the causal process leading to it from the phenomenon." The crucial aspect is that reliability of the data does not have to be epistemically explained by reference to scientific knowledge of causal mechanisms. Confounding factors may occur here and there, but overall the non-idiosyncrasy of data in which the patterns of a phenomenon can be detected make—somehow—sure that these data reliably provide evidence for the occurrence of a phenomenon.

I agree with Massimi that "reliability cannot be entirely detached from the causal knowledge of the mechanism that generates true beliefs with high frequency." (p. 106) Why is that?

> Unless we somehow know already how the phenomena that we are searching for should look like, how can we appraise whether data production and data reduction provide *reliable* evidence for them? In a way, this problem is a re-elaboration of what Harry Collins (1985/1992) has described as the *experimenter's regress*: in order to prove that an experimental process is reliable, we have to show that it identifies the phenomenon correctly. But in order to identify the phenomenon correctly, one has to rely on the experimental process whose reliability is precisely at stake. So reliability seems to fall back into a justificatory circle. (p. 108, original emphasis)

The crucial point is that data and patterns in them are to some degree idiosyncratic. Even if we discuss (Bogen and) Woodward's example of the detection of weak neutral currents, some idiosyncrasies (*e.g.* exact angle of the trace) occur and to decide whether these idiosyncrasies are relevant for the phenomenon detection or not we need to know how the measurement apparatus works, or in other words, we need sufficient "causal knowledge of the mechanism that generates true beliefs" about weak neutral currents in bubble chambers. In the more extreme exemplary case of albinism the two very distinct concrete patterns are integral parts of the scientific description of the phenomenon of albinism, but agents need a lot of knowledge about the involved data gathering techniques to be able to properly refer to the phenomenon of albinism. A reference to a certain mutation of the 11th chromosome is not a reference to albinism; all known concrete patterns are necessary to properly describe the phenomenon of albinism according to the relevant scientific community.

Conclusively, since reliability of data cannot be detached from causal knowledge about the data gathering routines, phenomena realism cannot be defended by reference to reliability.

In other parts of this thesis I already stated that Kantianism provides a fruitful approach for our analysis of phenomena.[1] Massimi follows this route in claiming that "[p]henomena are not ready-made in nature, instead we have somehow to make them. And we make them by first ascribing certain *spatiotemporal properties* to appearances (...), and then by subsuming them under a *causal concept*." (p. 109) This is an accurate brief summary of the cornerstone of Kantian epistemology. Massimi also discusses Suppes' (1962) hierarchy of models (cf. 2.1) to derive from these approaches that "phenomena scientists investigate are often the end product of these intermediate steps, at quite a distance from the original data" and these steps "may also require a significant amount of *conceptual construction*." (p. 110)

However, in my view, pure Kantianism too much neglects the role of empirical assumptions that are not so deeply grounded by reasons of basic human epistemic capabilities or even science. Phenomena are not *global* in the sense that at every historical phase of science every scientist of a specific branch applies the same criteria for the distinction between phenomena and non-phenomenal features of the empirical world. I want to add a specific focus on background assumptions that are neither rational nor justified by natural human epistemic restrictions.

One well-known approach to add further criteria to any descriptively adequate[2] account of phenomena selection is Kuhn's (1962) idea of paradigmatic science. If

[1]The prominent role of mathematical objects for Kant's *transcendental aesthetics* support my claim that data and patterns are epistemically nothing more than mathematical objects. (cf. ch. 2 and 4) Furthermore, Kant's focus on fundamental epistemic preconditions for human perception support Grenander's pattern constructivism that I follow (cf. 4.3).

[2]Massimi explains that she "want[s] to endorse an approach to the data—phenomena distinction which is normative and naturalised at the same time" meaning that she takes "the natural sciences as paradigmatic of scientific knowledge" (p. 102).

science is organised in paradigms, which I believe it is, then a proper account of paradigms need to be a part of or need to be referenced for any descriptive account of phenomena. Furthermore, even a normative account of phenomena needs to take criteria of paradigmatism into account. According to Kuhn, not only hypothesis, but also scientific methodology and criteria for relevant scientific questions are part of a paradigm. Therefore, a paradigm may also guide the distinction between phenomena and non-phenomenal features of the empirical world. This seems at least valid to me for the science of human agents.

Chapter 7

Conclusion

I provide the results of my thesis that I deem to be the most crucial ones. I start with replies to central questions and concerns first, and then I correct (Bogen and) Woodward's notion of phenomena in accordance with my conclusions. Chapters and sections are referenced in brackets.

Answers to Central Questions and Concerns

- In the discussion I made a very fundamental epistemological assumption: every pattern recognition technique that is used in science can in principle be reduced to the execution of a mathematical algorithm. This position is, of course, subject of discussion and opposed by some in the philosophy (of mind) and to a lesser degree in neurobiology. I argued in favour of this position by reference to some classics, in particular Russell (2.4 and 3.1), and to the *optimistic induction* (3.3). Mathematical structuralism shows how broad an adequate notion of mathematics is (3.1) and Grenander's *general pattern theory* (4.4) is an approach for a mathematical pattern epistemology. Nevertheless, rejecting or accepting this assumption leads to very different philosophical views towards science and, in particular, of phenomena in science. It even leads to different views towards human cognition in general, since the whole variety of pattern detection procedures that human beings carry out in science may already reflect many different features of human cognition. However, I focus on scientific data and inferences that have to be intersubjectively accessible (2) and explicable to other agents in principle. Therefore, possibly inexplicable criteria of pattern selection (if such exist) are genuinely *unscientific* and play no inferential role in science.

- Historically, science is a *human* endeavour. This is due to the fact that, at least until very recently, humans are the only agents that we are aware of being able to *do* science to the full extent of this activity. Concerning this point I make two claims. First, laymen's everyday reasoning and scientific inferences from scientific experts does not differ in principle. (1.2) Both have the same normative grounds but can be executed erroneously. Second, science cannot in principle be restricted to human agents.—This statement is meant to have normative as well as a descriptive implications. Artificial intelligences, in particular deep learning AIs, exemplify that computers can make inferences that are *creative* in the sense that a human agent is not capable of *explaining* how this inference was made. (3.3) The *optimistic induction* implies that this will extend to scientific applications of AI agents in the future. Furthermore, possible aliens (from other planets, physical realms or periods) may be able to formulate theories with predictive empirical success and, therefore, we would have to accept them as scientific agents.

- Concrete patterns are the elements of the classes that are the general patterns and therefore *describe* the phenomena (4.1). In most cases, scientific agents know only some concrete patterns and even those usually not in a fully explicated way due to pragmatism and epistemic restrictions. Almost without exception only an *epistemically unbounded agent* (4.1) is able to know the full general pattern of a phenomenon. In the example of an animal with albinism there are at least two concrete patterns in very different data: (a) in the measurements of the colour of the hair, skin and eyes and (b) in the records of genetic analyses. Since the knowledge about (b) is itself part of a theory, the *description* of a phenomenon seems to depend on theories, but the phenomena themselves are not theory-laden in the strict sense. (5.4) They are real features of the empirical world that do not depend on theories or knowledge and they *cause* observations. (5.1) But we need to remain agnostic to what exactly this real world is; actual physical particles, Descartes' demon, Bostrom's simulation or the like (1.3). Concerning phenomena that are described in the context of a scientific field that makes use of references to many very theoretical entities, such as theoretical physics, there may be only theory-dependent descriptions of respective patterns possible. But this theory-dependence is nothing substantially different than the agents' natural epistemic restrictions due to shared background assumptions, sensory capabilities and epistemic limitations. (5) One can say that the phenomena of our science are not theory-laden but they rather are *human-laden*, since human epistemic and sensory capabilities play the most important part for the historically changing list of scientific phenomena. That is why phenomena are real features of the empirical world, but the selection between phenomena and non-phenomenal features has no parallel part in the empirical world.

Conclusive Answer to (Bogen and) Woodward

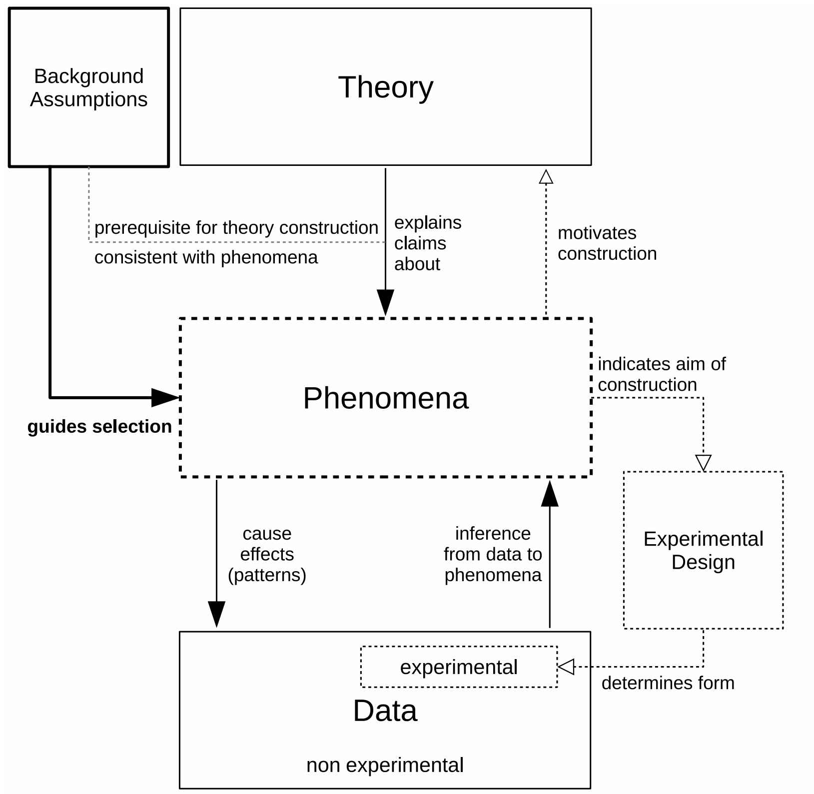


Figure 7.1: Schematic view of my ideas about the role of scientific theories, phenomena and data. Box and line styles are used with the same meaning than in figure 1.2 on page 3. Double line width indicates my adaptions with regard to figure 1.2 but no different meaning than regular line widths. The box for the phenomena is dotted because the phenomenon selection depends crucially and in principle on the background assumptions and epistemic capabilities (which I here subsume under the background assumptions for illustrative simplicity).

In the introduction I expressed a need for further clarification on how (Bogen and) Woodward's view on phenomena can be maintained or criticised. According to their view, phenomena are best described as being (i) not idiosyncratic to the different ways that are used to detect them, (ii) real in an ontological sense, and (iii) identical to patterns. These points can be answered as follows.

Regarding (i), the not idiosyncratic occurrences of a phenomenon, which are the occurrences of concrete patterns from the respective general pattern in proper data, can be explained by, firstly, the relation of representation for data (2.3) and for patterns (4.1) and, secondly, by the classification of *non-similar* concrete patterns under one general pattern (4.1) on the basis of shared theoretical background assumptions (5). The general patterns describe the phenomena one-to-one (5.2).

Regarding (i) and (ii), phenomena are real features of the empirical world, but the demarcation between phenomena and non-phenomenal features of the empirical world is based on the body of shared background assumptions, sensory and epistemic capabilities of the relevant agents. Regarding (iii), phenomena are not ontologically identical to patterns, but they are fully described by their corresponding general pattern and pragmatically described in the context of the agents' assumptions and capabilities by the scientifically discussed concrete patterns. Figure 7.1 provides a schematic view with reference to the schemes that are given in the introduction (1).

To summarise, the explanatory gap between the idiosyncratic data and patterns on the one side and the not idiosyncratic phenomena on the other side of this dichotomy is closed by the introduction of representations for data and for patterns, and the account of a description of phenomena by general patterns.

Appendix

A.1 Decomposition of Data into Patterns and Noise from the Perspective of Statistical Modelling

Section Abstract

McAllister's view on the (statistical) decomposition of data into patterns and noise is empirically inaccurate. As part of the scientists' shared background assumptions a body of worked out and established statistical methods is applied in scientific applications of pattern analysis. Therefore, any descriptive notion of noise in data needs to incorporate actual notions of noise from the field of statistics.

McAllister (1997) states that sets of data can, in principle, be decomposed into infinitely many combinations of a pattern and of a noise term. The intuition behind this decomposition is very simple: a set of data, or at least a proper representation of it, can be understood as a mathematical time series or signal. A pattern in this set of data is, according to McAllister, a deterministic statistical model that describes this time series or signal in approximation. The noise are the statistical *residuals*, which are not modelled and are of no further crucially significant purpose for the scientific analysis of the data. Without any scientific preconditions, that is if the data, pattern and noise are interpreted as purely mathematical objects without having to fulfil any further (mathematical) criterion, McAllister's problem occurs.

He stresses the argument that, if we cannot define any satisfactory criterion to explain what properties of a set of data belong to the pattern part and which belong to the noise, then the choice of the pattern in data is arbitrary. Therefore phenomena cannot be explained as patterns in data, because in this case phenomena are necessarily arbitrary properties of the empirical world, which would no proponent of phenomena in science be willing to admit. He further infers that, if patterns in data are arbitrarily defined and occur infinitely often, then phenomena must so, too. I argue against this criticism in part 6.1.

The plots in figure A.2 show a very common textbook example (that I already mentioned in 4.2) on mathematical time series analysis and they are very illustrative to describe the idea of the decomposition of data into a pattern and the residual noise, whereas the noise carries no statistical *information* in terms of the specific statistical theory applied.

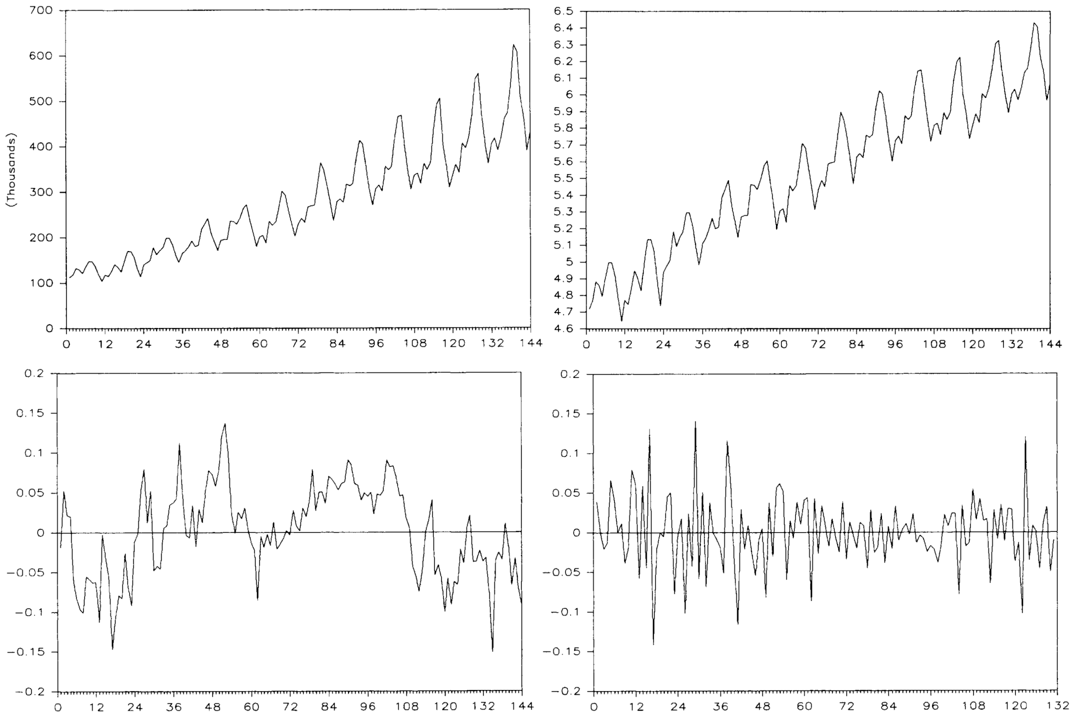


Figure A.2: Monthly international airplane passengers from January 1949 to December 1960 as an illustrative example of data for statistical data analysis and modelling. (plot from Brockwell and Davis 1991, data from Box and Jenkins 1976) A more thorough discussion of the classical *decomposition model* applied on these data is given in Brockwell and Davis (§1.4 and §9.2).
Top left: plot of totals in thousands of passengers $(P_t)_{t\in\{1,2,\ldots,144\}}$.
Top right: natural logarithm $(L_t)_{t\in\{1,2,\ldots,144\}} := (\ln P_t)_{t\in\{1,2,\ldots,144\}}$ of the data.—The application of the natural logarithms is the application of a *variance-stabilising* technique, as discussed by Box and Cox (1964).
Bottom left: *residuals* $(Y_t)_{t\in\{1,2,\ldots,144\}}$ after removing the linear *trend* $(m_t)_{t\in\{1,2,\ldots,144\}}$ and *seasonal component* $(s_{\hat{t}})_{\hat{t}\in\{1,2,\ldots,12\}}$ with period 12 from $(L_t)_{t\in\{1,2,\ldots,144\}}$. That is the following equation holds:

$$L_t = m_t + s_{t \bmod 12} + Y_t \text{ for all } t \in \{1, 2, \ldots, 144\}.$$

Bottom right: differenced series $D_t = \Delta_1\Delta_{12}L_{t+13}$ for all $t \in \{1, 2, \ldots, 131\}$ with Δ_n, $n \in \mathbb{N}$ being the difference operator defined by $\Delta_n X_m := X_m - X_{m-n}$ for any series $(X_t)_{t\in\mathbb{Z}}$.

However, despite McAllister's view that patterns in statistical data can be arbitrarily chosen I want to point out that the notion of noise is for actual scientific inferences in fact not that arbitrary. In the field of mathematical statistics, which is a straight-forward application of McAllister's description of the decomposition of data into a pattern and residual noise, the notion of noise is pecisely specified. More importantly, an established statistical framework that includes precise definitions of noise is widely applied in everyday science of various quantitative fields (*e.g.* social studies; particle physics; astrophysics). In the mathematical theory concerning the modelling with stochastic processes solutions are available to carefully avoid the scenario presented by McAllister for many practical cases. This is achieved by simply constraining noise with very specific properties that are reasonable from a modelling perspective.

Following McNeil *et al.* (2005, p. 127), we can define (strict) white noise: a real valued, covariance stationary[1] stochastic process $(X_t)_{t\in\mathbb{Z}}$ with autocorrelation of zero for all lags,

$$\operatorname{corr}(X_t, X_{t+n}) = 0 \ \forall\ t \in \mathbb{Z}, n \in \mathbb{N},$$

is called **white noise**. A stochastic process $(X_t)_{t\in\mathbb{Z}}$ is called **strict white noise**, if it is a series of identically and independently distributed random variables with finite variances.

Please note that strict white noise is a special case of white noise since identical and independent distribution is a stronger constraint than covariance stationarity.[2] According to these common definitions, (strict) white noise has a constant volatility for every time $t \in \mathbb{Z}$, but this parameter is not fixed, that is, every (strict) white noise process can have a different constant volatility.

In statistical time series analysis it is a usual approach to describe the empirical time series as a composition of a deterministic process and white noise or strict white noise. In this setup the deterministic part plays the role of McAllister's version of a pattern and the (strict) white noise plays the role of McAllister's version of noise. A simple example of this approach is the AR(1) model discussed by McNeil *et al.* (2005, p. 129).

The **AR(1) model** or **- process** is defined as a stochastic process $(X_t)_{t\in\mathbb{Z}}$ satisfying the following condition:

$$X_t = \phi X_{t-1} + \varepsilon_t \ \forall\ t \in \mathbb{Z},$$

[1] Be I a set that is closed under addition $+ : I \times I \to I$. A stochastic process $(S_t)_{t\in I}$ with values in Ω' ist called **covariance stationary** (or **weakly stationary**), if

(cs1) $\mathbb{E}[S_t^2] < \infty$ for all $t \in I$,

(cs2) $\exists\ m \in \Omega' : \mathbb{E}[S_t] = m$ for all $t \in I$ and

(cs3) $\operatorname{cov}(S_t, S_s) = \operatorname{cov}(S_{t+r}, S_{s+r})$ for all $r, s, t \in I$.

[2] One can define stochastic processes that are covariance stationary, but is not a series of independently distributed random variables. For an example see Ströing (2012, 3.1.4).

with $(\varepsilon_t)_{t\in\mathbb{Z}}$ being white noise and $\phi \in \{r \in \mathbb{R} : -1 < r < 1\}$.[1]

'AR' is shortened for *autoregressive*, '1' denotes the number of variables, and the process is an easy example to model dependencies in empirical time series. It is intended to model autocorrelation and, in most applications to empirical data, positive correlations are produced by a choice of $\phi \in \{r \in \mathbb{R} : 0 < r < 1\}$. Figure A.3 illustrates the produced autocorrelation in an example.

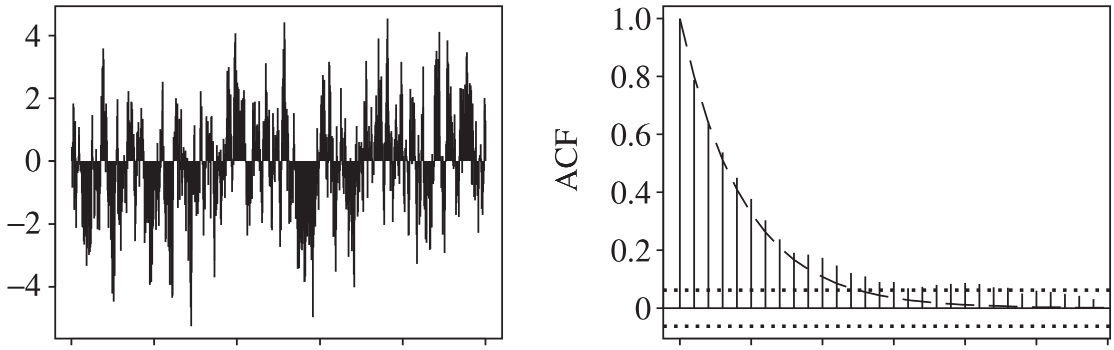


Figure A.3: From McNeil *et al.* (2005, p. 130). A series sampled from an AR(1) process with $\phi = 0.8$ (left plot) with theoretical autocorrelations over the different lags (right plot, dashed line), autocorrelations of the sampled series (right plot, vertical bars) and 95% confidence interval according to the sample size (right plot, dotted lines).

The statisticians' motivation behind the AR(1) model is representative to a standard approach that aims for a separation of empirical time series into a deterministic and a noise component by statistical modelling. Several further models, such as the MA -, ARMA -, ARCH - and GARCH processes follow this approach and are empirically applied to risk models for liquid financial markets.[2] As opposed to McAllister's description, the pattern and the noise term are precisely defined. To model an empirical time series with the AR(1) model the parameter $\phi \in]0,1[$ can be uniquely estimated to describe a unique pattern and the volatility of the white noise $(\varepsilon_t)_{t\in\mathbb{Z}}$ can be uniquely estimated to describe the unique noise part.

If an AR(1) model does not fully describe a specific set of empirical data we would not admit that there is something wrong with the noise part or the pattern,

[1]The constraint $\phi \in \{r \in \mathbb{R} : -1 < r < 1\}$ is stricter than most definitions of the AR(1) process in the literature in which often only $\phi \in \mathbb{R}$ is demanded. However, with our stricter constraint the AR(1) model is always a *causal* process that converges with a solution

$$X_t = \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i}.$$

See McNeil *et al.* for details.

[2]'MA' denotes *moving average* and this model is very similar to the AR, but provides a easier control of the autocorrelations for higher lags. ARMA is a model that is defined as a sum of a AR process and a MA process. (G)ARCH models are further described and discussed in section A.2.

via the deterministic part, it describes. A scientist would interpret such a result as an approximation for which he has to decide whether he considers the pattern he assumed as detected or not. But there is not much room to alter the pattern except for choosing or defining another statistical model.

Another approach to mathematically specify noise of an empirical time series or signal that is is the decomposition into *independent random drivers*, which is a very common technique in statistical analysis. I introduce this approach first in an outline and discuss the consequences for a concept of noise. The general idea behind this modelling approach is that the data points of the time series or signal are described by a random variable X in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and can further be described as the result of a function

$$f : (Y_1, Y_2, ..., Y_n) \to X$$

on $n \in \mathbb{N}$, $n \leq 2$, independent random variables $(Y_i)_{i \in \{1,2,...,n\}}$ on the same probability space or on subspaces (*i.e.* $(\Omega, \mathcal{F}_i, \mathbb{P})$ with $\mathcal{F}_i \subset \mathcal{F}$ for all $i \in \{1, 2, ..., n\}$).[1]

The most common approach of the described kind in various scientific and engineering contexts is the so called *principal component analysis* ('PCA') as intuitively introduced by Pearson (1901). PCA models the random variable X as a linear combination of its independent random drivers $(X_i)_{i \in \{1,2,...,n\}}$

$$f : (Y_1, Y_2, ..., Y_n) \to X, f : (Y_1, Y_2, ..., Y_n) \mapsto \lambda_1 Y_1 + \lambda_2 Y_2 + ... + \lambda_n Y_n$$

with $\lambda_i \in \mathbb{R}$, $i \in \{1, 2, ..., n\}$. Due to the linearity, PCA can make use of two comparatively simple mathematical theories, which are numerically worked out for the purposes of application in science and engineering: covariances[2] of random variables or statistical samples on the one side and linear algebra on the other side. In the PCA the components $(Y_i)_{i \in \{1,2,...,n\}}$ are called *factors* (motivated by the linearity, of course) and are defined as pairwise orthogonal (*i.e.* $\langle Y_i, Y_j \rangle = 0$ for all $i, j \in \{1, 2, ..., n\}$, $i \neq j$, for a scalar product $\langle \cdot, \cdot \rangle : \{Y : Y$ random variable $\}^2 \to \mathbb{R}$) according the specific scalar product

$$\langle \cdot, \cdot \rangle_{PCA} : \{(Y_1, Y_2) : Y_1, Y_2 \text{ real-valued random variable}\} \mapsto \operatorname{cov}(X_1, X_2).$$

[1]Note that $(\mathcal{F}_i)_{\{1,2,...,n\}}$ is not necessarily—and in fact in all useful cases—a filtration (*i.e.* $\mathcal{F}_s \subset \mathcal{F}_t$ for all $s, t \in \{1, 2, ..., n\}$ with $s < t$), but fulfils another modelling purpose: whereas a filtration is used to model the progress of knowledge along, for example, the time, the subspaces $(\Omega, \mathcal{F}_i, \mathbb{P})$ are introduced to model different risk drivers that are independent.

[2]Note that the correlation $\operatorname{corr}(X, Y) \in \{r \in \mathbb{R} : -1 \leq r \leq 1\}$, which is the variance-normed covariance, of two random variables X and Y describe the strength of the linear dependency between these random variable. Even uncorrelated (*i.e.* $\operatorname{corr}(X, Y) = 0$) random variables can be strongly dependent in various ways. For a discussion and examples see Ströing (2012, ch. 3).

PCA is most often applied to multivariate empirical time series, empirical signals or defined random variables $\vec{X} = (X_1, X_2, ..., X_n)$, $n \in \mathbb{N}$, $n \geq 2$, and the covariances are applied to the time series, signals, or random variables of different dimensions. Certain univariate empirical time series, empirical signals or stochastic processes $(X_t)_{t\in\mathbb{Z}}$ can, however, be modelled with a independent random drivers analysis, such as PCA, for which the orthogonality is not defined for different dimensions.

For both cases, PCA as an approach to model a multidimensional or an one-dimensional empirical time series or signal, simple examples may help to explain the general idea of PCA and the approach of analysis of independent risk drivers in general. I introduce examples from the modelling of financial market data due to the very intuitive nature of dependencies in these data.

For the multidimensional case consider the price time series of two stocks that are in some aspect considered to be related and are therefore significantly non-independent concerning their price evolution. This relation can be that they share a geographic home market (*e.g.* both are mainly active in France), a business model (*e.g.* both are airlines) or their well-being depends on certain other market influences (*e.g.* both depend on the prices of a certain commodity that is subject to heavy price fluctuations itself). Due to this simple form of dependence, a positive and significant correlation between the two price time series is detectable. The dominant random driver according to a PCA on these data results in the modelling of all uncertainty that both time series have in common. The second, orthogonal random driver in the PCA models the remaining risk from all uncertain influences, except this most driving factors. Another, more complex, but very illustrative example of a PCA model for a multidimensional time series would be a bigger portfolio of stocks, as the German DAX, which consists of the 30 biggest German public limited company fulfilling certain criteria describing the tradability of their shares. A PCA with the choice of some $n \in \mathbb{N}, n < 30$, approximates the 30 real dimensions of risk by a lower dimension of risk drivers that may be interpretable (*e.g.* a German financial institutions factor, an global car sales potential factor, an energy price factor and so on).

A one-dimensional example of a decomposition into independent random drivers could be explained by modelling data of the financial market, too. Consider the price time series of an international conglomerate. A model can reasonably describe this time series as the sample from a stochastic process, which is a function on other stochastic processes that model the independent (or negligibly dependent) random drivers. Those random drivers are semantically assigned to foreign exchange rates, commodity prices, refinancing costs and others. Note that, strictly speaking, this approach os useless as a model in isolation; it only makes sense in a market with data from further assets that depend on the random drivers or these random drivers themselves as isolated prices.

Why is the statistical technique of decomposition into independent random drivers interesting for our analysis of noise in scientific data? In practice PCA is applied in the following way: first, a number of PCA random factors $n \in \mathbb{N}$ is chosen, second, the n most driving random factors are determined by eigenvalue decomposition of the covariance matrix from the empirical time series or signal that is aimed to be modelled; if n is not the maximum, that is the dimension of the time series or signal in the multivariate case, the remaining components of the empirical time series, empirical signal or random variable is left unexplained by the model and therefore interpreted as noise. The point of interest to us is that this noise component is uniquely determined after the choice of the number of PCA random factors.

The analysis of independent random drivers is a widely used statistical technique in which the noise and the pattern are very explicitly distinguished and this is another example against McAllister's problem of an arbitrary decomposition of scientific data into the pattern and into the noise part. The point I want to highlight is that the common use of modelling techniques for scientific pattern detection often implies a clear definition of the residual noise.

A proponent of McAllister's criticism may stress the point that no *principal* objection to McAllister is given by the argument that the *commonly applied* framework of statistical pattern detection precisely define the residual noise by their models. However, these two cases are examples of how a large proportion of pattern detection in science is based on a very restricted repertoire of established methods. In a similar way in which we can easily identify a person with albinism by his/her appearance with our senses, we can detect autocorrelations in a time series from a scientific measurement. Our well-trained sense for the detection of skin colour and our statistical knowledge are two sides of the same coin, because they are both part of the shared epistemic background that guide our phenomenon selection. But is the detection of a phenomenon not something very different than the mere detection of a pattern in accordance with these shared epistemic background assumptions? No, in my view there is no reason to believe that.

A.2 *Volatility Clustering* in Financial Market Data as an Example of Phenomenon Selection[1]

Section Abstract

Volatility clustering is an example to illuminate various discussed aspects of phenomenon selection. Volatility clustering can be observed with the naked eye in plots of raw data. It can be tackled with approaches from different scientific fields and has important implications for them. The mathematical explication of its pattern(s) is a highly regarded, Nobel Prize awarded scientific task.

I discuss the example of a phenomenon from quantitative finance with implications for economics and psychology: *volatility clustering* and ARCH type models that were rewarded with the *Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel* in 2003.[2] In part A.1 I discuss the common concept of decomposition of a set of statistical data into some pattern and residual *noise*. With the discussion in this section I aim to investigate a very specific example of a statistical analysis according to this standard statistical procedure of pattern detection. I focus on a scientific example of a widely discussed phenomenon to highlight the import characteristics of phenomena in general. Main questions, in accordance with our general discussion about phenomena, are: how theory-laden are phenomena? Can phenomena be identified only on the grounds of statistical pattern analysis? What makes scientific agents *pick* a phenomenon?

In a nutshell, ARCH type models extend the model of the (geometric) Brownian motion, as used by Bachelier (1900), Merton (1973), and Black and Scholes (1973), for returns of financial assets by an additional local volatility component for risk applications instead of option pricing.[3] Find a plot of the geometric Brownian motion at figure A.4. *Local volatility* models have parameters to fit the simulated absolute level of returns to market phases with high volatility due intensive trading activities, and market phases with low volatility due to less trading activity. These changes are observed in many financial data and are commonly listed under the

[1] I adapted plotting scripts from my diploma thesis (2012) to produce the plots in this section.

[2] See *nobelprize.org* (2003)

[3] These mentioned sources predominated the introduction of option pricing models and therefore established the (geometric) Brownian motion as an asset price model. Merton and Scholes received the *Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel* in 1997 (see *nobelprize.org* 1997).—Please note the very long time-gaps between the different publications and the awarding. Bachelier's text was only poorly received by the contemporary scientists, despite the fact that his instructor was Henri Poincaré. Paul Samuelson, receiver of the *Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel* in 1970, popularised Bachelier's text among economists in the 1950s. (cf. Taqqu 2001) The period of 24 years between the publication of the Black-Scholes model and the reception of the Nobel honours can be explained by the establishment of the model among practioners in the 80s and 90s.

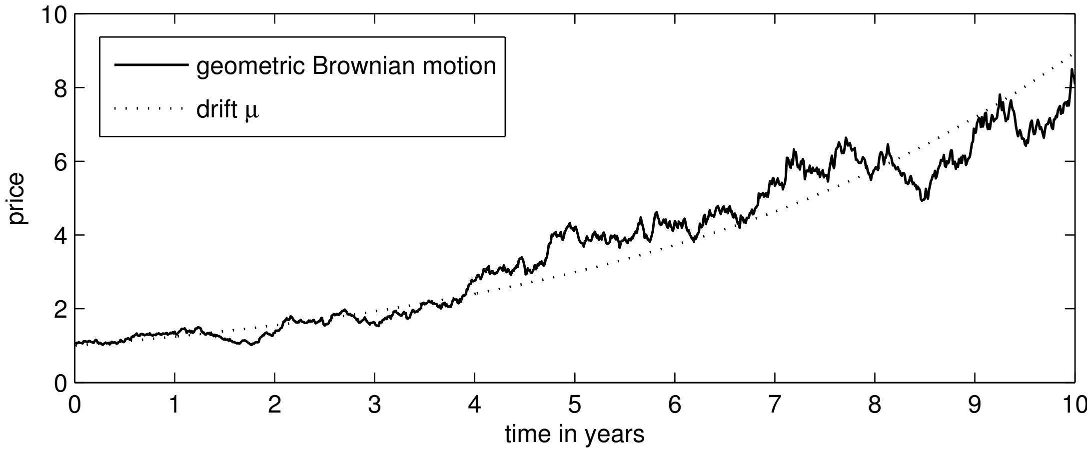


Figure A.4: Sample path of a geometric Brownian motion ('GBM') with drift $\mu = 0.08$ and volatility $\sigma = 0.20$. The GBM makes use of the normal distribution to model log-returns. The description of the axis indicate its use as an asset price model. See figure A.5 (bottom) for a plot of the probability distribution function of the normal distribution. A GBM $(S_t)_{t \in \mathbb{R}_{\geq 0}}$ is usually defined as a stochastic differential equation

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t$$

with *drift* and *volatility* $\mu, \sigma \in \mathbb{R}_{\geq 0}$ and $(W_t)_{t \in \mathbb{R}_{\geq 0}}$ a Brownian motion, that is a time continuous process with normally distributed, independent and stationary increments; see Klenke (2008, ch. 21) for an introductory discussion.

so-called *stylised facts* of financial time series in the literature.[1] The local volatility, often also called *volatility clustering*, is the phenomenon we discuss in this example from science. In a brief outline, volatility clustering is the phenomenon that financial markets undergo *calm* phases of low absolute returns and *hectic* phases of large absolute returns without any implication for the direction of the returns.

The scientific aim of ARCH type models is to explain and predict the return distributions of financial assets under the assumption that these assets can be described by a stationary stochastic process in sufficient approximation. Observations of real market data show that log-returns are not normally distributed, as suggested by the mentioned classical asset return models introduced by Bachelier or Black, Merton and Scholes. But the returns have a *leptokurtic* distribution which imply more return outcomes close to the mean and more return outcomes very far away from the mean than predicted by a normal distribution. Figure A.5 illustrates this statistical behaviours found in financial data. Figure A.6 illustrates how the most often discussed ARCH type model, the GARCH(1,1) process, explains leptocurtic heavy-tail behaviour of the financial data.

The Brownian motion as a model for financial price data has a theoretical justification. By application of the Central Limit Theorem[2] the Brownian motion

[1]See, for example McNeil *et al.* (2005, ch. 4.1.1).

[2] In a nutshell, the central limit theorem states that the probability distribution of the sum $\sum_i X_i$ of independent and identically distributed random variables $X_1, X_2, ...$ with real values and finite variance converges to a normal distribution. This theorem may not be intuitive, but

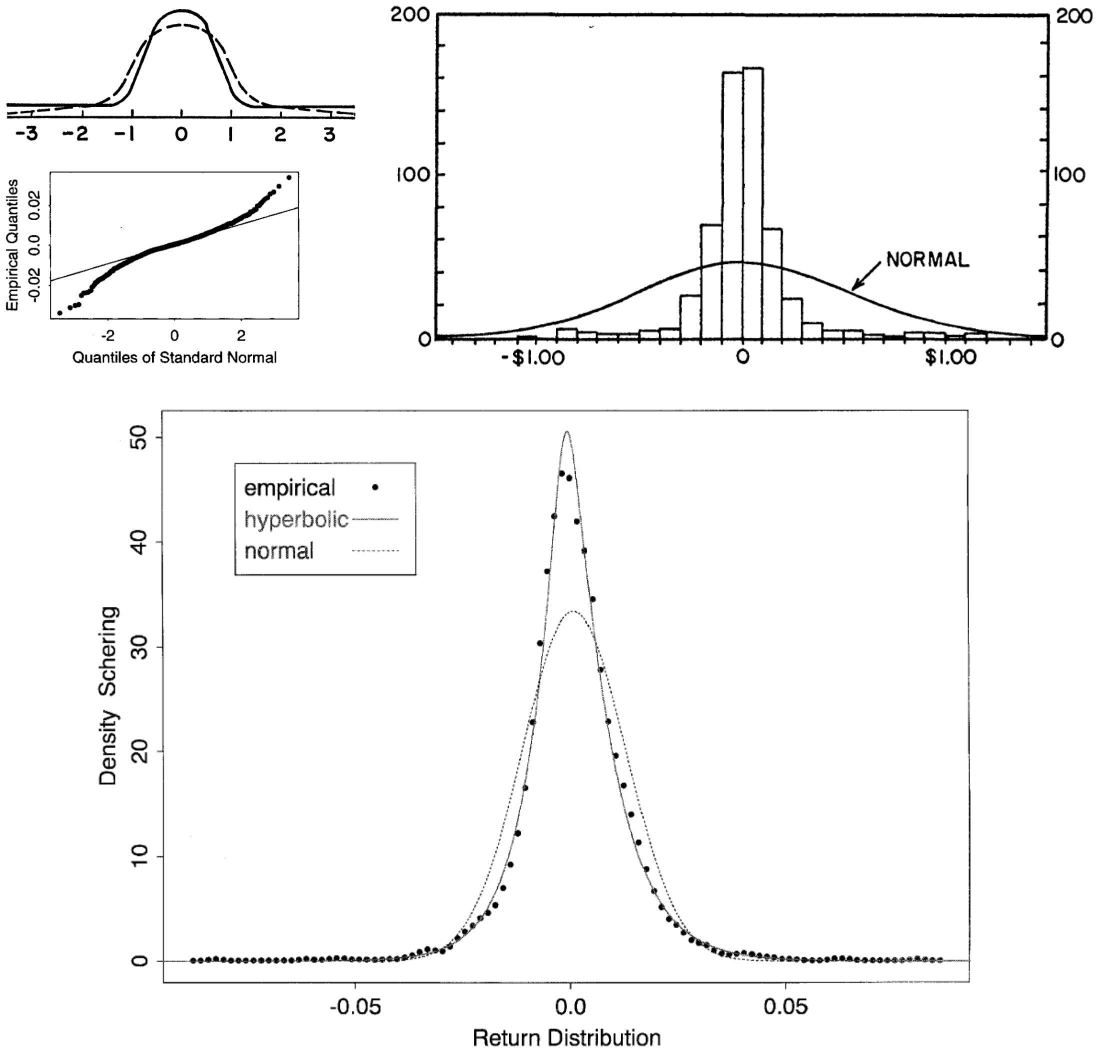


Figure A.5: Top left: simple sketch from Fama (1965) to illustrate the findings in stock market price data—the solid line shows the empirical probability distribution compared to a normal probability distribution indicated by the broken line.
Top right: study from Tintner (1940), quoted by Mandelbrot (1963), showing a histogram of the fifth difference of historical monthly wool prices from 1890-1937 indicating the non-normality features.
Mid left: The quantile-quantile plot from Eberlein (1998) based on daily NYSE return data from the 1990s confirms the leptocurtic behaviour of the empirical probability distribution far away from the mean—so-called *heavy tails* occur.
Bottom: study from Geman (2002) showing a fitted leptocurtic density of daily returns in the 1990s from the stock of Schering, a large German pharmaceutical company. The hyperbolic model is one of many non-normal models to fit empirical stock prices.

follows to be the appropriate model, if returns over a certain time interval (in many applicatory cases daily returns) can be assumed to be pairwise independent. Due to *market efficiency* pairwise independence—at least in approximation—is often considered a reasonable assumption.

However, this assumption is misleading according to the available empirical measurement data, the financial time series. As common properties of stock asset price data show positively autocorrelated volatilities (or absolute returns) can be found and are usually mentioned under the *stylised facts*. For the scientific aim of ARCH type models it is important to note, that they do not aim for a prediction of specific returns to aid routines for return maximisation. They explain and predict return distributions based on historical time series of the asset in question. And the crucial discovery by the development of ARCH type models is that the non-normality of asset log-return distributions is caused by a certain local behaviour of the volatilities.

For the sake of a thorough historical introduction we mention that the geometric Brownian motion as a model of asset prices (see figure A.4) is not a very good model from the perspective of the stylised facts. However, Merton (1973), Black and Scholes' (1973) aim for the discussion of this model does not only imply the prediction and explanation of the stylised facts. They rather discuss the fulfilment of practical requirements such as hedging routines that can be realised in the Black-Scholes model, but cannot be realised with other models that may show more empirically adequate modelling behaviour in accordance with the stylised facts.

After these precedent introductory notes I introduce the ARCH type models in further detail. ARCH is shortened for *autoregressive conditional heteroscedasticity* and these models aim to explain statistical asset price behaviour of, for instance, liquidly traded stocks. *Scedasticity* is a statistical term denoting the distribution of error terms from a signal to which a model is applied to. *Heteroscedasticity* denotes non-identically distributed error terms, and more particular error terms with varying variances.[1]

On the background of the classical Brownian motion model to describe statistical asset price behaviour in the case of ARCH this heteroscedasticity refers to

affirms the outstanding role of the normal distribution in many applications of statistics. It can be applied to price time series, if we assume that the market prices are driven by sufficiently small time atoms (for convergence) and in a world with bounded amounts of resources (to ensure finite variances). However, the empirical results of, among other contributions, the ARCH type models show that the assumption of the time atoms being driven independently of each other is wrong. It can be shown that current market volatilities depend on earlier volatilities in the market. See Klenke (2008, ch. 15.5) for a thorough mathematical discussion of the central limit theorem.

[1] Find a statisticians introduction to scedasticity in Armitage and Colton (2005).

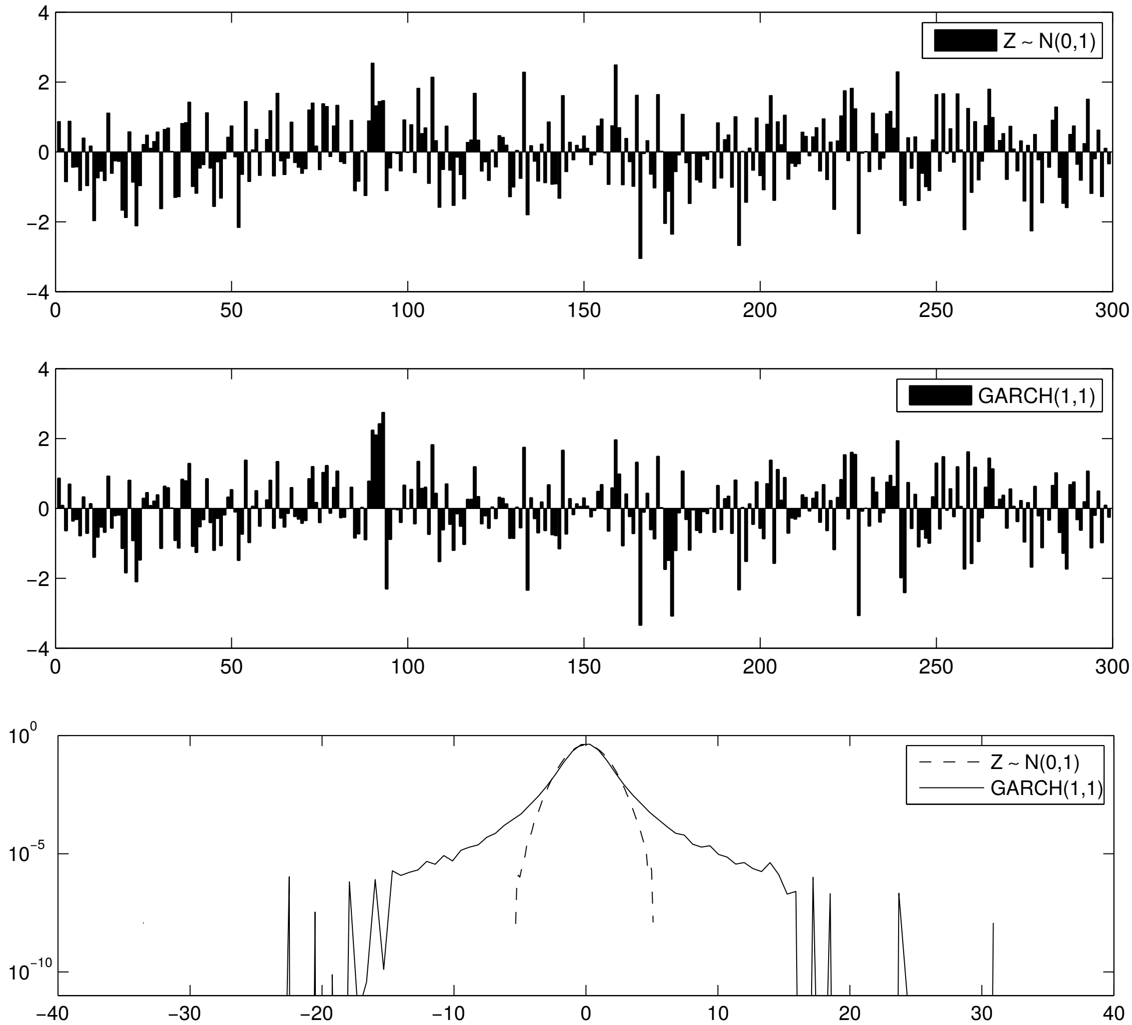


Figure A.6: Top: 300 realisations of $Z := (Z_t)_{t\in\mathbb{N}}$ normally distributed $Z_t \sim \mathcal{N}(0,1)$ for all $t \in \mathbb{N}$.
Mid: realisations of GARCH(1,1) process $X := (X_t)_{t\in\mathbb{N}}$ with $\alpha_0 = 0.3$, $\alpha_1 = 0.4$, $\beta_1 = 0.3$ and covariance stationary second moment based on the normally distributed realisations $(Z_t)_{t\in\mathbb{N}}$ above.
Bottom: Empirical probability distribution of Z and X from the mid plot based on 10^7 realisations. Please note the logarithmically scaled second axis.

The top and mid plots explain why the distribution of the GARCH(1,1) process in the bottom plot shows *heavy tails*. After an extremely large return in Z medium size returns of Z become relatively large returns in X as well (*e.g.* $t \approx 90$). After a series of consecutive large returns of Z blow up effects with very large absolute returns in X occur (*e.g.* $t \approx 230$).

the local changes in volatility that are subsumed under the error terms in the classical constant volatility model of a Brownian motion with its constant volatility parameter. Nevertheless, ARCH type models tackle the phenomenon of *volatility clustering* by modelling the volatility conditionalised on earlier realisations of the volatility.[1] That is, squared returns of the near future are stochastically modelled positively correlated to recent realisations of squared returns. These squared returns of ARCH are autoregressive in the sense that the expected value of future volatilities change with the realisations of the process.[2]

Engle (1982) introduces ARCH motivated by rates of inflation in the United Kingdom. Bollerslev (1986) expands the model to GARCH (*Generalized ARCH*) to define a more empirically adequate model process with respect to runaway values in realisations of the process, as will be explained more detailed below. We call ARCH and later specifications and enhancements based on ARCH *ARCH type models*. Further ARCH type models are defined to specify the autocorrelations of the volatilities of return processes.

I introduce the original ARCH model.

Definition (ARCH process)

Be $(Z_t)_{t\in\mathbb{Z}}$ a stochastic process of iid.[3] random variables with $\mathbb{E}[Z_t] = 0$ and $\mathbb{V}\mathrm{ar}[Z_t] = 1$ for all $t \in \mathbb{Z}$. The stochastic process $(X_t)_{t\in\mathbb{Z}}$ is an **ARCH(p) process**, if it is strictly stationary[4] and for all $t \in \mathbb{Z}$ and a strictly positive $(\sigma_t)_{t\in\mathbb{Z}}$

$$X_t = \sigma_t \cdot Z_t, \tag{A1}$$

$$\sigma_t^2 = \alpha_0 + \textstyle\sum_{i=1}^{p} \alpha_i X_{t-i}^2 \tag{A2}$$

holds with $\alpha_0 > 0, \alpha_i \geq 0$ for $i = 1, ..., p$.

The volatility σ_t scales with earlier realisations of the processes variance X_t^2 for all $t \in Z$ and therefore models the volatility clustering. In many discussions in the literature the parameter choice $p = 1$ is chosen. Plot A.7 illustrates the statistical behaviour of the ARCH process. Note the specific modelling results of the volatility clustering in form of the autocorrelation function in part (d) of figure A.8.

An enhanced version of ARCH, called GARCH, is widely considered to describe the empirical data more appropriately with sufficient pragmatic simplicity.

[1]Find a more elaborated mathematical treatment in Ströing (2012, ch. 3.2). For proofs of mathematical propositions I refer to this text.

[2]For another introduction see McNeil *et al.* (2005, ch. 4.3). For a more general introduction into autoregressive processes see Brockwell and Davis (1991, ch. 3).

[3]The common abbreviation 'iid' stands for *independent and identically distributed*.

[4]Be I a set that is closed under an addition operator $+ : I \times I \to I$. A stochastic process $(S_t)_{t\in I}$ is called **(strictly) stationary**, if $\mathcal{L}[(S_t)_{t\in I}] = \mathcal{L}[(S_{t+s})_{t\in I}]$ for all $s \in I$.

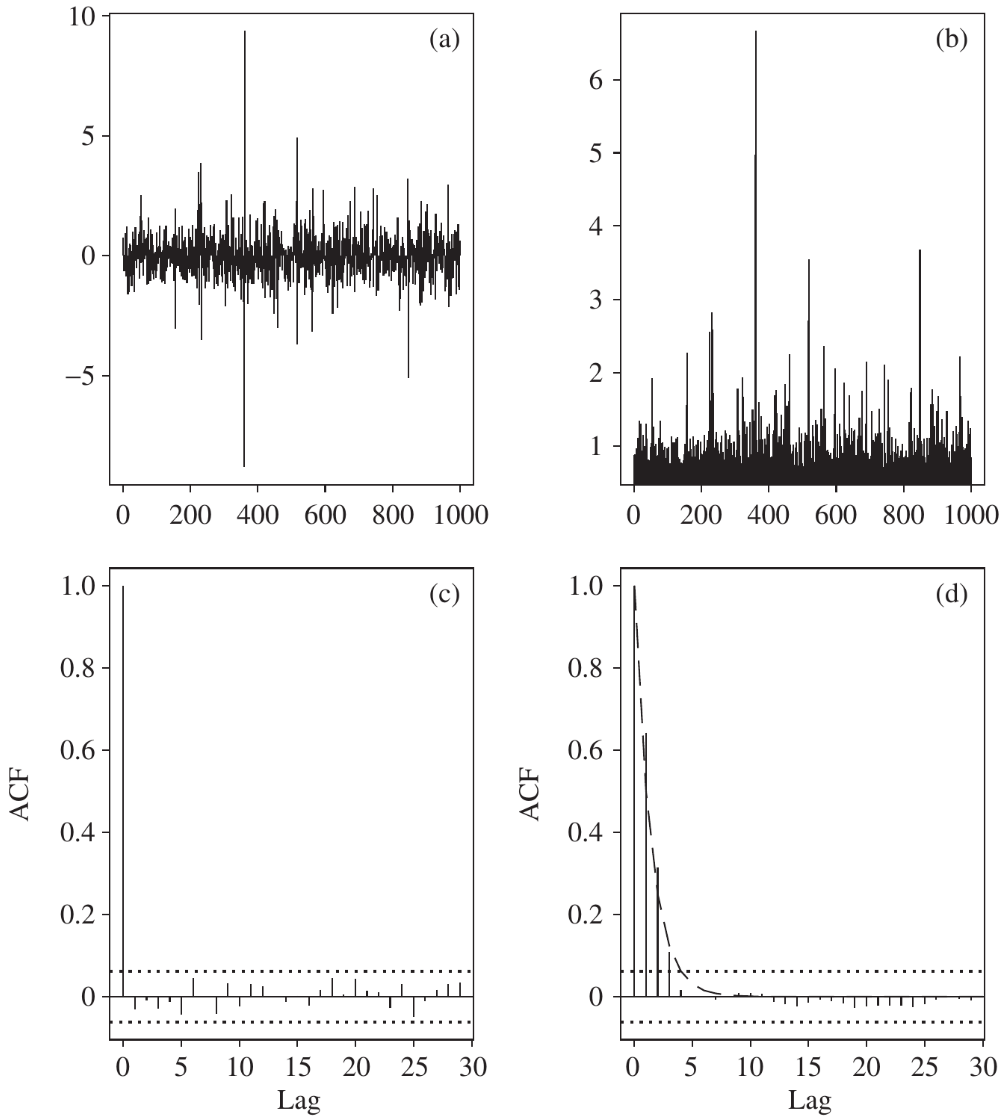


Figure A.7: Illustrations of an ARCH(1) process from McNeil *et al.* (2005, p. 140) with $\alpha_0 = \alpha_1 = 0.5$.
Top left (a): realisations X_t, $t \in \{1, 2, ..., 1000\}$.
Top right (b): volatilities σ_t, $t \in \{1, 2, ..., 1000\}$ of realisations.
Bottom left (c): autocorrelations of the realisations from (a) to lags 0 to 30.
Bottom right (d): autocorrelations of the squared realisations from (a) X_t^2 of lags 0 to 30. The dashed line shows the true form of the autocorrelations that was analytically inferred.

Definition (GARCH process)

Be $(Z_t)_{t\in\mathbb{Z}}$ a stochastic process of iid. random variables with $\mathbb{E}[Z_t] = 0$ and $\mathbb{V}\mathrm{ar}[Z_t] = 1$ for all $t \in \mathbb{Z}$. The stochastic process $(X_t)_{t\in\mathbb{Z}}$ is a **GARCH(p,q) process**, if it is strictly stationary and for all $t \in \mathbb{Z}$ and a strictly positive $(\sigma_t)_{t\in\mathbb{Z}}$

$$X_t = \sigma_t \cdot Z_t, \tag{G1}$$

$$\sigma_t^2 = \alpha_0 + \textstyle\sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{G2}$$

holds with $\alpha_0 > 0, \alpha_i \geq 0$ for $i = 1, ..., p$, $\beta_j \geq 0, j = 1, ..., q$.

Parameter choices are most often $p = q = 1$. Note the illustration in figure A.8 and in particular the autocorrelation function in part (d).

IGARCH (*Integrated GARCH*), coined by Franses (1995) but already indicated by Engle and Bollerslev (1986), is constrained by a pathological parameter choice—for clarification see the definition of GARCH process below—fulfilling

$$\sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{q} \beta_i = 1.$$

This parameter choice is motivated by empirical analysis of asset price data and the need for respectively strong autocorrelations of the squared log-returns in modelling, but the process has a unit root and therefore non-stationarity of the process can be shown.[1] Without the property of stationarity ARCH type models cannot be applied reasonable. The reason is that statistical estimators of probability distributions and auto-dependencies in time series can be defined reasonably only if the sample data is assumed to be described by a strictly stationary process. Otherwise the specific non-stationarity, that is the change of the probability distributions and auto dependencies in the stochastic model process over time need to be described specifically. This endeavour calls for another box of non-trivial statistical tools than the ARCH approach delivers.

As a further approach to model the observed data more adequately FIGARCH (*Fractionally Integrated GARCH*) is introduced by Baille, Bollerslev and Mikkelsen (1996). In the FIGARCH processes class autocorrelations of volatilities can be modelled with long-range dependence, that is an only hyperbolic decay of autocorrelation in time, as already defined (p. 115). This long-range dependency of autocorrelations is indicated by several empirical studies of market data[2].

[1] For the empirical evidence and parameter estimators to find these estimations see Franses (1995). For a statistician survey on unit roots at stochastic processes, see Box and Jenkins (1976, ch. 7.5).

[2] As also mentioned by Baille *et al.* (1996) examples of such studies are Baillie and Bollerslev (1989), Bollerslev (1987), Hsieh (1989), and McCurdy and Morgan (1988).

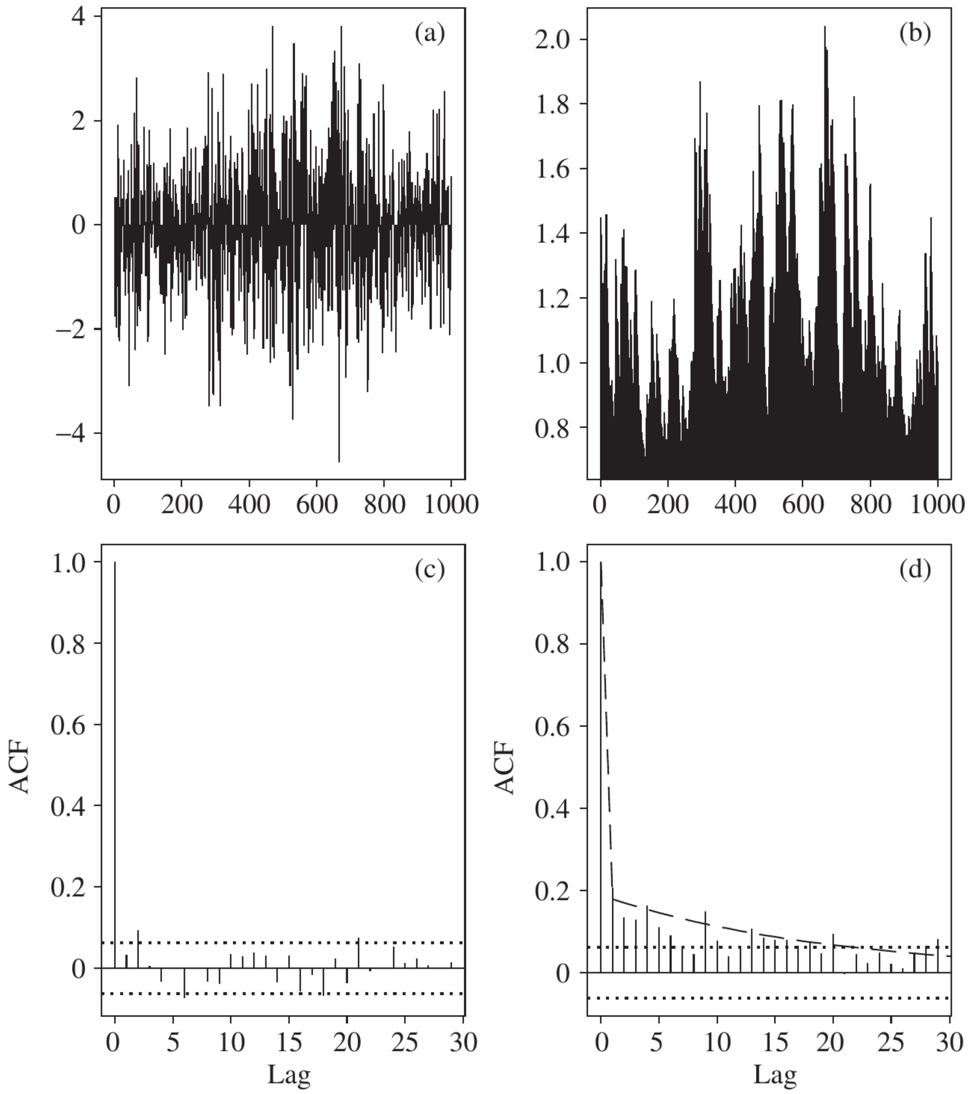


Figure A.8: Illustrations of a GARCH(1,1) process from McNeil *et al.* (2005, p. 147) with $\alpha_0 = 0.5$, $\alpha_1 = 0.1$ and $\beta = 0.85$.
Top left (a): realisations X_t, $t \in \{1, 2, ..., 1000\}$.
Top right (b): volatilities σ_t, $t \in \{1, 2, ..., 1000\}$ of realisations.
Bottom left (c): autocorrelations of the realisations from (a) to lags 0 to 30.
Bottom right (d): autocorrelations of the squared realisations from (a) X_t^2 of lags 0 to 30. The dashed line shows the true form of the autocorrelations that was analytically inferred.

It should further be mentioned, that the family of ARCH type models is more widely discussed and several further models in this family are introduced. For an overview see the glossary by Bollerslev (2008).[1] In addition to the development of stochastic processes to model volatility dependencies in financial time series by ARCH type models extensions to multivariate models are introduced, as for instance the *Multivariate GARCH* model to model markets of more than one asset that have positively, negatively or not correlated log-returns with each other. (cf. McNeil *et al.* 2005, ch. 4.6)

I choose the ARCH type models to discuss phenomena in science due to the very apparent relation between the phenomenon, which is the volatility clustering, and the statistical patterns, which are mathematically explicated above. We find some very standard characteristics of phenomena and patterns in this example:

(a) The phenomenon of volatility clustering can be described without the use of vocabulary that is connected to some empirical theory. It is solely based on the data of market prices as a time series in any form. That is why it is mentioned under the common stylised facts of financial data. Simple illustrations of return time series show this stylised fact to the naked eye without any mathematical explication.

(b) The concrete patterns corresponding to the phenomenon in different sets of data are not completely identical. In the case of different asset classes or markets, the patterns may slightly differ. To say it in other words, the estimated parameters α_0, α_i, $i \in \{1, ..., p\}$, (and β_j, $i \in \{1, ..., q\}$) of the (G)ARCH model are significantly different in sets of data from different asset classes, different historical phases or different markets.[2] These differences are too large to be just explained by measurement inaccuracies. However, in the economists view all these cases still count intuitively as instances of

[1]To say it more frankly, the development of ARCH type models is promoted by many researchers, as the introduction of the referred glossary indicates. The topic may be appealing due to the relative simplicity of model construction, the empirical explanatory power and empirical applicability to log-return time series, as well as the non-trivial mathematics, that comes into need for proving crucial properties of the models. To proof for instance the existence of second and forth moments, and covariance stationarity (see p. 195) under certain constraints in the parameter choices stochastic recursion equations are used. These were explicitly discussed by Brandt (1986) and goes back to earlier works by Kesten (1973). For details on this point see Ströing (2012, proposition 3.2.6) and McNeil *et al.* (2005, p. 141–142).

[2]Chou (1988) investigates weekly return data of a *New York Stock Exchange* value-weighted index from 1962 to 1985. His estimations of the α and β of the GARCH(1,1) model significantly differ for different time intervals before and after 1974. Malkiel's (1979) and Pindyck's (1984) hypothesis about the rising investment uncertainty in the overall market around 1974 is a scientific explanation for this change of market behaviour in general, and volatility clustering in particular. Franses and van Dijk (1996) show estimation results for GARCH(1,1) for weekly returns from main indices from different European countries from 1986 to 1989. The parameters α and β differ significantly, too.

the one phenomenon of volatility clustering. Thus, unlike other phenomena, such as the melting point of the elements, we face clear indications that the phenomenon of volatility clustering is detected by a concrete pattern that is not discrete in the mathematical sense and therefore vague to a certain degree. In opposition to the following point (c), this vagueness is, in principle, not caused by the scientists' specific model choices, instead it is caused by the data itself.

(c) Scientists developed a variety of different ARCH type models to further specify the pattern that corresponds to the phenomenon of volatility clustering. Therefore, the actually defined concrete patterns have approximative features, which is not generally uncommon for many mathematically specified concrete patterns in science. This is an easily comprehensible feature of modelling results in scientific fields in which scientists face a principal lack of data to define more refined models based on solely empirical grounds, such as in social and economic sciences. However, the heterogeneity of the defined models is a matter of practising science and producing conveniently applicable results rather than of the pattern in question in general. In the case of ARCH type models one may be able to define an "overall", very cumbersome, model with many parameters to incorporate all statistical effects as descibed by each of every ARCH type model. If completely successful, this pattern would be the general pattern of volatility clustering.

Based on these facts I want to further discuss the example. If (a) is true, how then can volatility clustering be a phenomenon and many other arbitrary patterns in time series of asset prices do not correspond to any phenomenon?—This is the question McAllister raises (6.1). Even if (a) is true, asset prices have an obvious empirical interpretation that is shared by every scientist of the relevant field: prices are caused by investors' decisions, stock market infrastructure, macroeconomic circumstances, implemented algorithms and other influences; all these influences take place in a very large and complex world of many market participants. This shared empirical interpretation of the data implies certain expectations on the time series of asset prices, such as the level of tolerance against arbitrary statistical effects and noise that is deemed insignificant, stationarity of a modelling process, and boundaries of statistical values like the drift and volatilities.

Against the background of these shared expectations about the empirical behaviour of asset markets volatility clustering is a phenomenon because it is neither an obvious implication of the common empirical interpretations nor do these empirical interpretations exclude the observed volatility clustering. This phenomenon is something that has to be explained in addition to the statistical behaviour of asset price data that can usually be explained be the mentioned common empirical interpretations that are shared by all scientists of relevant fields.

Furthermore, the other way round, volatility clustering brings interesting implications for the different relevant scientific fields with it. Some of these can be outlined by the following questions: is volatility clustering consistent with *market efficiency*, as introduced by Fama, in the field of economics?[1] What does volatility clustering, which is in most parts caused by the decisions of human traders, tell us about herding behaviour and applied rationality among human beings in the field of psychology? To conclude, volatility clustering is a phenomenon because it is interesting due to its causational relations and implications by empirical influences and for different scientific theories and even different scientific fields. Furthermore, its pattern is easily detectable with the naked eye in plots of market data. However, the analysed set of data has to be appropriately modified to log-returns[2]

The characteristics of volatility clustering mentioned with (b) and (c) indicate that the concrete patterns that correspond to it are defined by certain boundaries, which can be mathematically specified, or are even vague. These boundaries can be expressed in certain parameter choices of the ARCH type models. The definition of acceptable boundaries for a certain financial time series to explicate the pattern of volatility clustering is no simple task. To do this may even not be possible in an observer-independent way. To use the mathematical vocabulary proposed by the ARCH process, one scientist may consider a time series at which we can estimate, for example, $\alpha_1 = 0.1$ showing volatility clustering and another scientist may not be willing to say so. However, to clarify philosophical aspects of boundary cases

[1] *Market efficiency* is a widely discussed economical hypothesis for which a proper definition is a subject of debate itself. Fama's (1965) introductory explanation may be sufficient to outline the general idea:

> In an efficient market, competition among the many intelligent participants leads to a situation where, at any point in time, actual prices of individual securities already reflect the effects of information based both on events that have already occurred and on events which, as of now, the market expects to take place in the future. In other words, in an efficient market at any point in time the actual price of a security will be a good estimate of its intrinsic value.

The literature about the efficient markets hypothesis usually refers to Fama's (1970) later and specific discussion of it. Concerning volatility clustering one can raise the question whether significantly different temporal volatility clusters can occur in efficient markets in which "actual prices" already "reflect the effects of information". Critics, including Shiller (1981; 1981), state that the volatility in financial markets tends to be too high to be explained solely by changes in the "intrinsic value" and available information about it. Keynes (1936) criticises the idea of a stock price reflecting its *intrinsic* or *fundamental value* by comparing stock markets with casinos and traders' preferences for certain stocks are driven by criteria that are better comparable to beauty than to the application of a rational calculus on available information.

[2] The returns of a geometric Brownian motion in the Black-Scholes model are log-normal distributed and the log-returns of a geometric are normally distributed. Thus, ARCH type models model log-returns of financial assets. See figure A.9 for illustration. See Ströing (2012, ch. 2.3.3) for a brief discussion on log-returns with example.

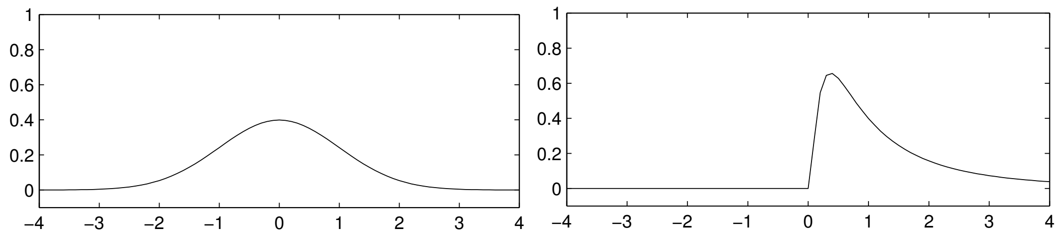


Figure A.9: Probability distribution functions of the normal distribution (left) and log-normal distribution (right).

and is concerning phenomena is another question we do not want to discuss here. But at already stated above, that volatility clustering is a phenomenon is not an observer-dependent proposition, but borderline cases may are so.

Against the background of characteristics (b) and (c), and the mentioned vagueness, there is still no problem in talking about *the* phenomenon of volatility clustering. All the patterns that are defined by different mathematical models according to (c) with differently estimated parameters according to (b) have to be put together in one class to define the general pattern of volatility clustering. No scientist would disagree on this.

We already discussed the case of the—mathematically rather simple—concrete pattern for the exemplary phenomenon of the melting point of lead. (p. 149 f.) I argued that an *ideal* pattern in the realists' sense is a reasonable approach to explain the ontology of phenomena based on this example. The vagueness of volatility clustering, as indicated by characteristic (b) seems to undermine this interpretation of phenomena prima facie. However, as already elaborated on, vagueness is not a problem for the ontological classification of a phenomenon. (cf. 6.1)

In our example of volatility clustering, we may not interpret *data reliability* as a critical subject prima facie. If we consider time series of asset prices $(S_t)_{t\in\{1,2,....,T\}}$, $T \in \mathbb{N}$, to be given in a reliable way via market quotes, then we can simply calculate the respective log-returns $R_t = \ln\left(\frac{S_t}{S_{t-1}}\right)$ for all $t \in \{2, 3, ..., T\}$ and check for volatility clustering by estimating the parameters of an ARCH type model.

On the other hand, one can argue that the market quotes themselves do not have to be reliable in every case. Due to market incidents the *liquidity* of a traded asset can be insufficient in certain time windows or, to say it in other words, not enough sell and buy activities take place to produce sufficiently reliable prices. Another scenario is that a stock can be subject to extreme or *untypical* price movements by sudden strategic actions of investors, managers or politicians, or a rapid change of available information; figure A.10 shows two exemplary stock price time series that motivate this discussions. Concerning quantitative risk management for banks, funds and investors the interpretation of a *typical* and *untypical* price movement, and whether this distinction is useful or even justified in principle is subject to discussions of practioners in the field.[1]

So, even in our example that does not require measurements comparable to those usually designed for various fields in experimental physics and where the pattern is fully mathematically specified, data reliability is a point of concern in our discussion. Volatility clustering can simply be detected by the application of statistical analyses to reliable sets of data, but the data reliability for asset prices is a subject of a—not very intricate—discussion on its own.

I briefly summarise the results from the discussion about the phenomenon of volatility clustering. To classify volatility clustering as a phenomenon does not depend on a specific theory an observer may favour, but it is not completely observer independent either. The community of scientists from the relevant fields share some very basic assumptions and interpretations, in our case the influences on asset price volatilities, and volatility clustering is accepted as a phenomenon since it is an empirical finding observable with the naked eye that cannot be explained by the assumptions and interpretations that are shared by the scientists. Another aspect we focussed on is that the reliability of the relevant sets of data is not a trivial matter in our example regardless the absence of experimental design and measurement routines. The mathematical explication of volatility clustering is a non-trivial task on its own and is aimed to further specify the concrete patterns and the general pattern of volatility clustering.

[1] This is a subject matter often criticised by econophysicists. Bouchaud (2008), and Bouchaud and Potters (2003, ch. 4 and 6) present a general critique of the economists and mathematicians lack of formulating models for the data as they really are. The natural scientists aim to build a proper model for the available market data is often corrupted by economists fundamental theoretical principles that may not and even should not be empirically justifiable. Another problem is the interest to fulfil certain criteria of elegance or rigour that is common in mathematical communities instead of choosing the best empirical model. Furthermore, practioners interests to discuss models that artificially lower risks and raise potential for financial gain comes into play, too. The econophysicists' criticism of the specific custom to neglect certain *untypical* data points in time series from market data is expressed by Mandelbrot:

> When the weather changes and hurricanes hit, nobody believes that the laws of physics have changed. Similarly, I don't believe that when the stock market goes into terrible gyrations its rules have changed. It's the same stock market with the same mechanisms and the same people. (2004)

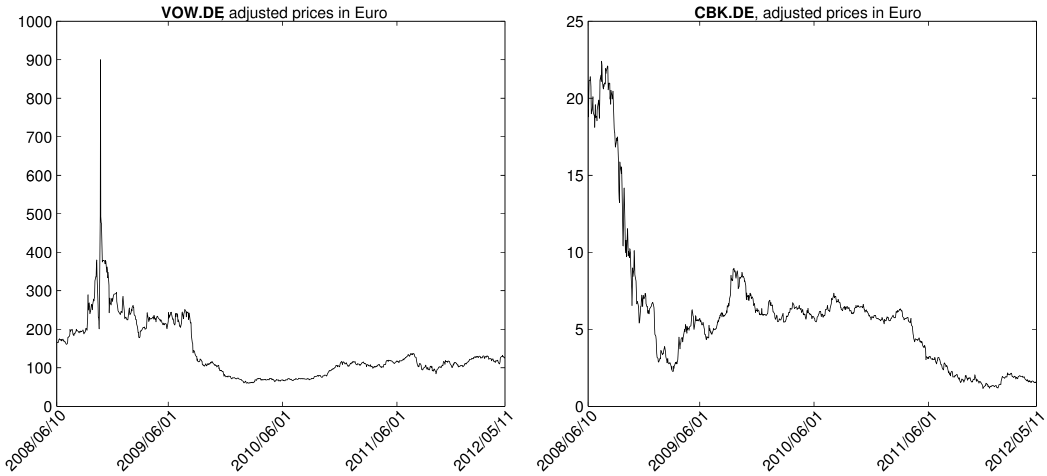


Figure A.10: Daily closing prices of stocks as quoted by Yahoo! Finance. Prices are adjusted for dividend payments and stock splits.—This is a common technical routine to gain more comparable data. Both stocks are listed in the German DAX since July 1988 and are therefore suitable examples of liquidly traded stocks.
Left: Prices for *Volkswagen AG* from mid-June 2008 to mid-May 2012. Due to a large amount of shares that were hold by *Porsche Automobil Holding SE* (up to 74.1%, cf. Schrinner 2008), including call options) and the fact that the federal state of Lower Saxony held 20% of all shares, short sellers were forced to pay extremely high prices for the remaining available stocks at the market in October 2008. The management of Porsche aimed for an acquisition of Volkswagen but later failed due to financing issues. (cf. Waldermann 2009) The resulting market scenario is referred to as *short squeeze* and, in this occurrence, reached its peak at October 28$^{\text{th}}$ 2008.
Right: Prices for *Commerzbank AG* from mid June to mid-June 2008 to mid-May 2012. These stock prices are heavily influenced by the *financial crisis of 2007–2008* and the subsequent *European debt crisis*. The financial crisis resulted in many defaults in the private and in the banking sector. The first heavy decrease of the price in autumn 2008 can be explained by the default of the American investment bank *Lehman Brothers Holdings Inc.* in September 2008. The US Government decided against the bail-out of the institution. Investors were alarmed that financial institutions are not safe from defaulting. The second heavy decrease of the price in mid-2011 can be explained by the European debt crisis. Market participants raised their believed probabilities that some European states, including Greece, Ireland and Portugal, default. Since the Commerzbank is a debtor of these states to a certain amount, a decrease of its value is reasonable.

A.3 A summarising illustration

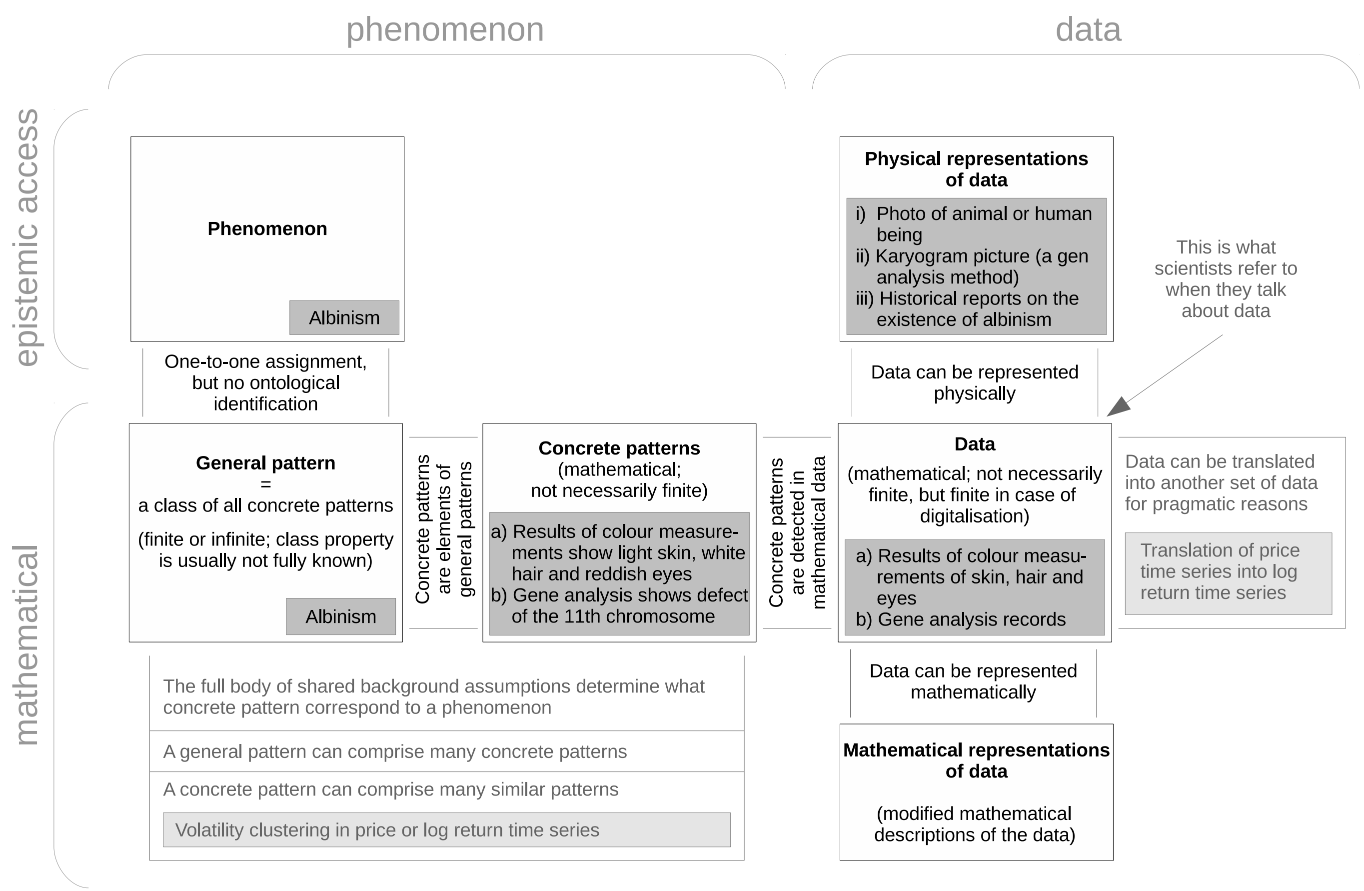


Figure A.11: Illustration of the relation between phenomena, data and patterns in science as introduced in this thesis. Examples of the illustrated objects or propositions or utterances are given from the phenomen of albinism (dark grey boxes) and from asset price data (light grey boxes). Black bordered boxes denote the objects or propositions or utterances and its properties. Grey bordered boxes denote further descriptions.

Bibliography

Editor's Comment

Non-English sources are listed in the original language with a suggested translation into English, if available. Clickable URLs are provided for PDF usage. Articles from journals or anthologies, as well as books with online access have stable URLs from DOI; in some rare exceptional cases without DOI entries (*e.g.* some older journals) direct URLs to the article at the Journal's homepage or at another licensed publisher are provided, if available.

Sources

[1] Abbott, B. P. *et al.* (2016). Observation of gravitational waves from a binary black hole merger. *Physical Review Letters*, **116**:061102.

[2] Aigner, M. and Ziegler, G. M. (2004). *Proofs from THE BOOK*. Springer, Berlin.

[3] Akins, K. (1986). *On Piranhas, Narcissism, and Mental Representation*. PhD thesis, Tufts University, Medford (MA).

[4] Alama, J. (2017). The lambda calculus. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Summer 2017 edition. URL `https://plato.stanford.edu/archives/sum2017/entries/lambda-calculus/`.

[5] Apel, J. (2011). On the meaning and the epistemological relevance of the notion of a scientific phenomenon. *Synthese*, **182**:23–38.

[6] Aquinas, T. (1265–1274). *Summa Theologiæ*. Unpublished. Several publications and translations posthume. Reprint 1947, Benziger Brothers, New York, translated by Fathers of the English Dominican Province.

[7] Armitage, P. and Colton, T. (Eds.) (2005). *Encyclopedia of Biostatistics*. Wiley, 2nd edition.

[8] Ashworth Jr., W. B. (1999). Brahe, Tycho (1546–1601). In Lankford, J. (Ed.), *History of Astronomy: An Encyclopedia*, pages 97–100. Routledge, New York.

[9] Avner, R. (2010). The Dome of the Rock in Light of Development of Concentric Martyria in Jerusalem: Architecture and Architectural Iconography. In Necipoğlul, G. (Ed.), *An Annual on the Visual Cultures of the Islamic World*, volume 27, pages 31–50. Brill, Leiden (NL).

[10] Awodey, S. (1996). Structure in mathematics and logic: A categorical perspective. *Philosophia Mathematica*, **4**(3):209–37.

[11] Awodey, S. (2004). An answer to Hellman's question: Does category theory provide a framework for mathematical structuralism? *Philosophia Mathematica*, **12**(1): 54–64.

[12] Azzouni, J. (2004). *Deflating Existential Consequence: A Case for Nominalism*. Oxford University Press, New York.

[13] B. Ubbink, J. (1960). Model, description and knowledge. *Synthese*, **12**(2):302–19.

[14] Bachelier, L. (1900). Théorie da la spéculation. *Annales de l'Ecole Normale Supérieure*, **17**:21–86. English translation by A. James Boness in Cootner (1964), pages 17–78, and with historical commentary by M. Davis and A. Etheridge (Eds.) (2006).

[15] Baillie, R. and Bollerslev, T. (1989). The message in daily exchange rates: A conditional variance tale. *Journal of Business and Economic Statistics*, **7**(3):297–305.

[16] Baillie, R., Bollerslev, T., and Mikkelsen, H. (1996). Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, **74** (1):3–30.

[17] Balzer, W., Moulines, C. U., and Sneed, J. D. (1987). *An Architectonic for Science: The Structuralist Program*. Reidel, Dordrecht.

[18] Bekker, I. (1831–1870). *Aristoteles Opera*. G. Reimerum, Berlin.

[19] Bell, J. L. (2015). The axiom of choice. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Summer 2015 edition. URL `https://plato.stanford.edu/archives/sum2015/entries/axiom-choice/`.

[20] Benacerraf, P. (1965). What numbers could not be. *The Philosophical Review*, **74**(1):47–73.

[21] Benacerraf, P. (1973). Mathematical truth. *The Journal of Philosophy*, **70**(19): 661–79. Reprinted in Benacerraf and Putnam (1983), pages 403–20.

[22] Benacerraf, P. and Putnam, H. (Eds.) (1983). *Philosophy of Mathematics: Selected Readings*. Cambridge University Press, Cambridge, 2nd edition. Reprinted in 1985, first published by Prentice-Hall, New Jersey, 1964.

[23] Bennett, C. (1990). How to define complexity in physics, and why. In Zurek, W. (Ed.), *Complexity, Entropy and the Physics of Information*, pages 137–48. Westview, Boulder.

[24] Bernays, P. (1935). Sur le platonisme dans les mathématiques. *L'enseignement mathématiques*, **34**:52–69. English translation by C. D. Parsons in Benacerraf and Putnam (1983), pages 258–71.

[25] Bernstein, E. and Vazirani, U. (1997). Quantum complexity theory. *SIAM Journal on Computing*, **26**(5):1411–73.

[26] Best, J. (2013). IBM Watson: The inside story of how the Jeopardy-winning supercomputer was born, and what it wants to do next. *TechRepublic*. URL `http://www.techrepublic.com/article/ibm-watson-the-inside-story-of-how-the-jeopardy-winning-supercomputer-was-born-and-what-it-wants-to-do-next` (retrieved: 6 March 2018).

[27] Biagini, F., Hu, Y., Øksendal, B., and Zhang, T. (2008). *Stochastic Calculus for Fractional Brownian Motion and Applications*. Springer, London.

[28] Bird, A. and Tobin, E. (2017). Natural kinds. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Spring 2017 edition. URL `https://plato.stanford.edu/archives/spr2017/entries/natural-kinds/`.

[29] Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, **81**(3):637–54.

[30] Blanchette, P. (2014). The frege-hilbert controversy. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Spring 2014 edition. URL `https://plato.stanford.edu/archives/spr2014/entries/frege-hilbert/`.

[31] Boesch, B. (2015). Scientific representation. In Fieser, J. and Dowden, B. (Eds.), *The Internet Encyclopedia of Philosophy*. URL `http://www.iep.utm.edu/sci-repr/`.

[32] Bogen, J. (2017). Theory and observation in science. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Summer 2017 edition. URL `https://plato.stanford.edu/archives/sum2017/entries/science-theory-observation/`.

[33] Bogen, J. and Woodward, J. F. (1988). Saving the phenomena. *The Philosophical Review*, **97**(3):303–52.

[34] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**(3):307–27.

[35] Bollerslev, T. (1987). A conditional heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, **69**(3): 542–7.

[36] Bollerslev, T. (2008). Glossary to ARCH (GARCH). *CREATES Research Paper*, **2008-49**.

[37] Bostrom, N. (2003). Are we living in a computer simulation? *The Philosophical Quarterly*, **53**(211):243–55.

[38] Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, Oxford.

[39] Bouchaud, J.-P. (2008). Economics needs a scientific revolution. *Nature*, **455** (7217):1181.

[40] Bouchaud, J.-P. and Potters, M. (2003). *Theory of Financial Risk and Derivative Pricing*. Cambridge University Press, Cambridge.

[41] Bourbaki, N. (1939–). *Éléments de Mathématique*. Hermann/Springer, Paris/Heidelberg. English translation as *Elements of Mathematics*, Springer, Heidelberg.

[42] Bourbaki, N. (1968). *Theories of Sets*. Elements of Mathematics. Addison-Wesley, Reading.

[43] Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B*, **26**(2):211–52.

[44] Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day, Oakland. Revised Edition. Reprinted in 4^{th}, repeatedly revised edition with G. Reinsel, Wiley, 2008.

[45] Brandom, R. (1994). *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard University Press, Cambridge. 2^{nd} Revised edition in 1998.

[46] Brandt, A. (1986). The stochastic equation $y_{n+1} = a_n + b_n$ with stationary coefficients. *Advances in Applied Probability*, **18**(1):211–20.

[47] Brentano, F. (1874). *Psychologie vom empirischen Standpunkt*. Duncker & Humblot, Leipzig. English translation in Brentano (1973).

[48] Brentano, F. (1973). *Psychology from an Empirical Standpoint*. L. L. McAlister (Ed.). Routledge, London. Translation by A. C. Rancurello, D. B. Terrell, and L. L. McAlister. 2^{nd} edition with introduction by P. Simons, in 1995.

[49] Bridges, D. and Palmgren, E. (2013). Constructive mathematics. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Winter 2016

edition. URL https://plato.stanford.edu/archives/win2016/entries/mathematics-constructive/.

[50] Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer. 2nd edition, first edition in 1987.

[51] Brown, J. (1994). *Smoke and Mirrors: How Science Reflects Reality*. Routledge, London.

[52] Bueno, O. and French, S. (2011). How theories represent. *The British Journal for the Philosophy of Science*, **62**(4):857–94.

[53] Burge, T. (1992). Frege on knowing the third realm. *Mind*, **101**(404):633–50.

[54] Cajori, F. (1960). *Sir Isaac Newton's Mathematical Principles of Natural Philosophy and His System of the World*. University of California Press, Berkeley (CA).

[55] Campbell, M., Hoane Jr., A. J., and Hsu, F.-H. (2002). Deep blue. *Artificial Intelligence*, **134**(1):57–83.

[56] Cantor, G. (1895). Beiträge zur Begründung der transfiniten Mengenlehre. *Mathematische Annalen*, **46**(4):481–512. English translation by P. Jourdain in Cantor (1915).

[57] Cantor, G. (1897). Beiträge zur Begründung der transfiniten Mengenlehre (Zweiter Artikel). *Mathematische Annalen*, **49**(2):207–46. English translation in Cantor (1915).

[58] Cantor, G. (1915). *Contributions to the Founding of the Theory of Transfinite Numbers*. Dover Publications, Berlin. English translation by P. Jourdain. Reprinted in 1955.

[59] Carnap, R. (1932). Über Protokollsätze. *Erkenntis*, **3**(1):215–28.

[60] Carnap, R. (1958). Beobachtungssprache und Theoretische Sprache. *Dialectica*, **12**(3-4):236–48.

[61] Carnap, R. (1966). *Philosophical Foundations of Physics: An Introduction to the Philosophy of Science*. M. Gardner (Ed.). Basic Books, New York.

[62] Carus, P. (1892). What does Anschauung mean? *The Monist*, **2**(4):527–32.

[63] Cassirer, E. (1936). Determinismus und Indeterminismus in der Modernen Physik: Historische und systematische Studien zum Kausalproblem. *Göteborgs Högskolas Årsskrift*, **42**(3). Reprinted 1937, Wettergren & Kerber, Göteborg. English translation as *Determinism and Indeterminism in Modern Physics* by O. T. Benfey, 1956, Yale University Press, New Haven.

[64] Chaitin, G. (1992). Information-theoretic incompleteness. *Applied Mathematics and Computation*, **52**(1):83–101.

[65] Chakravartty, A. (2011). Scientific realism. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Winter 2016 edition. URL `https://plato.stanford.edu/archives/win2016/entries/scientific-realism/`.

[66] Chou, R. (1988). Volatility persistence and stock valuations: Some empirical evidence using GARCH. *Journal of Applied Econometrics*, **3**(4):279–94.

[67] Church, A. (1932). A set of postulates for the foundation of logic. *Annals of Mathematics*, **33**(2):346–66.

[68] Church, A. (1936). An unsolvable problem of elementary number theory. *American Journal of Mathematics*, **58**(2):345–63. Reprinted in Davis (1965), pages 88–107.

[69] Church, A. (1940). A formulation of the simple theory of types. *Journal of Symbolic Logic*, **5**(2):56–68.

[70] Churchland, P. M. (1985). The ontological status of observables: In praise of the superempirical virtues. In Hooker, C. A. and Churchland, P. M. (Eds.), *Images of Science: Essays on Realism and Empiricism with Replies from Bas C. Van Frassen*, pages 35–47. University of Chicago Press.

[71] Collins, H. (1985/1992). *Changing Order*. The University of Chicago Press, Chicago. 2nd ed.

[72] Colyvan, M. (2001). *The Indispensability of Mathematics*. Oxford: Oxford University Press, Oxford.

[73] Cootner, P. (1964). *The Random Character of Stock Market Prices*. MIT Press, Cambridge.

[74] Copeland, B. J. (2002). The church-turing thesis. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Summer 2015 edition. URL `https://plato.stanford.edu/archives/sum2015/entries/church-turing/`.

[75] Corry, L. (1992). Nicolas Bourbaki and the concept of mathematical structure. *Synthese*, **92**(3):315–48.

[76] Creath, R. (2011). Logical empiricism. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Spring 2014 edition. URL `http://plato.stanford.edu/archives/spr2014/entries/logical-empiricism/`.

[77] Crovella, M. E. and Bestavros, A. (1997). Self-similarity in world wide web traffic: Evidence and possible causes. *IEEE/ACM Transactions on Networking*, **5**(6):835–46.

[78] Crupi, V. (2015). Confirmation. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Fall 2016 edition. URL `http://plato.stanford.edu/archives/fall2016/entries/confirmation/`.

[79] Cruz-Inigo, A. E., Ladizinski, B., and Sethi, A. (2011). Albinism in Africa: Stigma, slaughter and awareness campaigns. *Dermatologic Clinics*, **29**(1):79–87.

[80] Davidson, D. (1980). *Essays on Actions and Events*. Oxford University Press, New York. 2nd edition in 2002.

[81] Davidson, D. (2004). *Problems of Rationality*. Oxford University Press, New York.

[82] Davis, M. (Ed.) (1965). *The Undecidable: Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions*. Raven, Hewlett (NY). Reprint in 2004 by Dover, Mineaola (NY).

[83] Davis, M. (2000). *The Universal Computer: The Road from Leibniz to Turing*. W. W. Norton, New York.

[84] Davis, M. and Etheridge, A. (Eds.) (2006). *Louis Bachelier's Theory of Speculation: The Origins of Modern Finance*. Princeton University Press, Princeton.

[85] de Beauregard Robinson, G. (1959). *Foundations of Geometry*. University of Toronto Press, 4th revised edition.

[86] Dedekind, R. (1888). *Was sind und was sollen die Zahlen?* Friedrich Vieweg und Sohn, Braunschweig. Several reprints. English translation by W. W. Beman in *Essays on the Theory of Numbers* (1901), Open Court, Chicago, and revised by H. Pogorzelski as *What Are Numbers and What Should They Be?* (1995), Research Institute for Mathematics, Maine.

[87] DeepMind Technologies Limited. (2016). AlphaGo | DeepMind. URL `https://deepmind.com/research/alphago` (retrieved: 6 March 2018).

[88] Deiser, O. (2002). *Einführung in die Mengenlehre*. Springer, Berlin.

[89] Deiser, O. (2010). *Grundbegriffe der wissenschaftlichen Mathematik*. Springer, Berlin.

[90] Dennett, D. (1987). *The Intentional Stance*. MIT Press, Chicago.

[91] Dennett, D. (1991). Real patterns. *The Journal of Philosophy*, **88**(1):27–51.

[92] Descartes, R. (1641). *Meditationes de prima philosophia, in qua Dei existentia et animae immortalitas demonstratur*. Michel Soly, Paris. Most recent English translation by M. Moriarty as *Meditations on First Philosophy: With Selections from the Objections and Replies*, 2008, Oxford University Press, Oxford.

[93] Doukhan, P., Oppenheim, G., and Taqqu, M. S. (2003). *Theory and Applications of Long-Range Dependence*. Birkhäuser, Boston.

[94] Dummett, M. (1991). *Frege: Philosophy of Mathematics*. Harvard University Press, Cambridge.

[95] Ebbinghaus, H.-D., Flum, J., and Thomas, W. (1994). *Mathematical Logic*. Springer, New York. 2nd edition in 1996.

[96] Eberlein, E., Keller, U., and Prause, K. (1998). New insights into smile, mispricing, and value at risk: The hyperbolic model. *Journal of Business*, **71**(3):371–405.

[97] Eilenberg, S. and Mac Lane, S. (1945). General theory of natural equivalences. *Transactions of the American Mathematical Society*, **58**:231–94.

[98] Elgin, C. Z. (2010). Telling instances. In Frigg and Hunter (2010), pages 1–18.

[99] Ellerman, D. P. (1988). Category theory and concrete universals. *Erkenntnis*, **28** (3):409–29.

[100] Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**(4):987–1008.

[101] Engle, R. F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Review*, **5**(1):1–50.

[102] Everett, H. (1957). "Relative state" formulation of quantum mechanics. *Review of Modern Physics*, **29**(3):454–62.

[103] Ewald, W. B. (Ed.) (1996). *From Kant to Hilbert. A Source Book in the Foundations of Mathematics*, volume 2. Oxford University Press, Oxford.

[104] Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal*, **51**(1):55–9.

[105] Fama, E. F. (1965). The behavior of stock-market prices. *Journal of Business*, **38** (1):34–105.

[106] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, **25**(2):383–417.

[107] Feyerabend, P. (1962). Explanation, reduction, and empiricism. In Feigl, H. and Maxwell, G. (Eds.), *Minnesota Studies in the Philosophy of Science*, volume III, pages 28–97. University of Minnesota Press, Minneapolis.

[108] Feyerabend, P. (1975). *Against Method: Outline of an Anarchist Theory of Knowledge*. New Left Books, London. Earlier version in *Minnesota Studies in Philosophy of Science*, **4**:17–130, 1970. Later editions with significant revisions in 1988 and 1993.

[109] Field, H. (1980). *Science without Numbers: A Defense of Nominalism*. Princeton University Press, Princeton.

[110] Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science*, **66**(Proceedings Supplement):362–78.

[111] Fraenkel, A. (1922). Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre. *Mathematische Annalen*, **86**(3):230–7.

[112] Franses, P. H. (1995). IGARCH and variances change in the US long-run interest rate. *Applied Economic Letters*, **2**(4):113–4.

[113] Franses, P. H. and van Dijk, D. (1996). Forecasting stock market volatility using (non-linear) Garch models. *Journal of Forecasting*, **15**(3):229–35.

[114] Frege, G. (1879). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Louis Nebert, Halle a. S. English translation by S. Bauer-Mengelberg in van Heijenoort (1967), pages 1–82.

[115] Frege, G. (1884). *Die Grundlagen der Arithmetik: Eine logisch-mathematische Untersuchung über den Begriff der Zahl*. Wilhelm Koebner, Breslau. English translation by J. Austin in Frege (1959).

[116] Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, **100**:25–50. English translation in Frege (1980).

[117] Frege, G. (1893/1903). *Grundgesetze der Arithmetik*, volume I/II. Verlag Hermann Pohle, Jena. Partial english translation of volume I by M. Furth in Frege (1965).

[118] Frege, G. (1918–1919). Der Gedanke: Eine Logische Untersuchung. *Beiträge zur Philosophie des deutschen Idealismus*, **1**(1):58–77. English translation by A. Quinton and M. Quinton in Frege (1968).

[119] Frege, G. (1959). *The Foundations of Arithmetic: A Logico-Mathematical Enquiry Into the Concept of Number*. Blackwell, Oxford. Translated by J. Austin.

[120] Frege, G. (1965). *Basic Laws of Arithmetic*. University of California Press, Berkeley. English translation by M. Furth.

[121] Frege, G. (1980). On sense and reference. In Geach, P. and Black, M. (Eds.), *Translations from the Philosophical Writings of Gottlob Frege*. Blackwell, Oxford, 3rd edition.

[122] Freudenthal, H. (Ed.) (1961). *The Concept and the Role of the Model in Mathematics and Natural and Social Sciences*. Reidel, Dordrecht.

[123] Friedman, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy*, **71**(1):5–19.

[124] Frigg, R. (2010a). Fiction and scientific representation. In Frigg and Hunter (2010), pages 97–138.

[125] Frigg, R. (2010b). Models and fiction. *Synthese*, **172**(2):251–68.

[126] Frigg, R. and Hartmann, S. (2017). Models in science. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Spring 2017 edition. URL `https://plato.stanford.edu/archives/win2016/entries/scientific-representation/`.

[127] Frigg, R. and Hunter, M. C. (2010). *Beyond Mimesis and Convention: Representation in Art and Science*. Springer, Berlin and New York.

[128] Frigg, R. and Nguyen, J. (2016). Scientific representation. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Winter 2016 edition. URL `https://plato.stanford.edu/archives/win2016/entries/scientific-representation/`.

[129] Gabriele Contessa. (2007). Scientific representation, interpretation, and surrogative reasoning. *Philosophy of Science*, **74**(1):48–68.

[130] Geman, H. (2002). Pure jump Lévy processes for asset price modelling. *Journal of Banking & Finance*, **26**(7):1297–316.

[131] Geman, S. (2016). Ulf Grenander, 1923-2016. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **179**(4):1122–4.

[132] Giere, R. N. (1988). *Explaining Science: A Cognitive Approach*. Chicago University Press, Chicago.

[133] Giere, R. N. (2004). How models are used to represent reality. *Philosophy of Science*, **71**(5):742–52.

[134] Giere, R. N. (2010). An agent-based conception of models and scientific representation. *Synthese*, **172**(2):269–81.

[135] Glymour, B. (2000). Data and phenomena: A distinction reconsidered. *Erkenntnis*, **52**(1):29–37.

[136] Godfrey-Smith, P. (2006). The strategy of model-based science. *Biology and Philosophy*, **21**(5):725–40.

[137] Goldman, A. (1967). A causal theory of knowing. *The Journal of Philosophy*, **64** (12):357–72.

[138] Goldman, A. (1976). Discrimination and perceptual knowledge. *The Journal of Philosophy*, **73**(20):771–91.

[139] Goldstein, S. (2001). Boltzmann's approach to statistical mechanics. In Bricmont, J., Ghirardi, G., Dürr, D., Petruccione, F., Galavotti, M., and Zanghi, N. (Eds.), *Chance in Physics*, volume 574 of *Lecture Notes in Physics*, pages 39–54. Springer, Berlin.

[140] Goodman, N. (1976). *Languages of Art*. Hackett, Indianapolis and Cambridge, 2^{nd} edition.

[141] Goodman, N. and Quine, W. V. O. (1947). Steps towards a constructive nominalism. *Journal of Symbolic Logic*, **12**(4):97–122.

[142] Grau-Carles, P. (2000). Empirical evidence of long-range correlations in stock returns. *Physica A: Statistical Mechanics and its Applications*, **287**(3):396–404.

[143] Grenander, U. (1963). *Probabilities on Algebraic Structures*. Wiley, New York. Latest reprint in 2008, Dover.

[144] Grenander, U. (1970). A unified approach to pattern analysis. *Advances in Computers*, **10**(1):175–216.

[145] Grenander, U. (1993). *Pattern Theory: A Mathematical Study of Regular Structures*. Clarendon Press, Oxford.

[146] Grenander, U. and Miller, M. (2007). *Pattern Theory: From Representation to Inference*. Oxford University Press, Oxford.

[147] Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, **38**(1):173–198. English translation in van Heijenoort (1967), pages 596–616.

[148] Gödel, K. (1934). On undecidable propositions of formal mathematical systems. Lecture notes taken by S. Kleene and J. B. Rosser at the Institute for Advanced Study, Princeton (NJ). Reprinted in Davis (1965), pages 39–74.

[149] Gödel, K. (1944). Russell's mathematical logic. In Schilpp, P. A. (Ed.), *The Philosophy of Bertrand Russell*, The Library of Living Philosophers, pages 123–53. Northwestern University, Evanston and Chicago. Reprinted in Benacerraf and Putnam (1983), pages 447–69.

[150] Hacking, I. (1985). Do we see through a microscope? In Hooker, C. and Churchland, P. (Eds.), *Images of Science: Essays on Realism and Empiricism with Replies from Bas C. Van Frassen*, pages 132–52. University of Chicago Press, Chicago.

[151] Hacking, I. (1992). The self-vindication of the laboratory sciences. In Pickering, A. (Ed.), *Science as Practice and Culture*, pages 29–64. University of Chicago Press, Chicago.

[152] Hanson, N. R. (1958). *Patterns of Discovery*. Cambridge University Press, Cambridge.

[153] Haugeland, J. (1993). Pattern and being. In Dahlbom, B. (Ed.), *Dennett and His Critics*, pages 53–69. Blackwell, Oxford. Reprint in *Having Thought: Essays in the Metaphyics of Mind*, 1998, pages 267-90. Harvard University Press, Cambride (Massachusetts).

[154] Heath, T. L. (1908). *The Thirteen Books of Euclid's Elements*, volume 1, 2 and 3. Cambridge University Press, Cambridge.

[155] Heidegger, M. (1927). *Sein und Zeit*. Niemeyer, Tübingen. Reprinted in 19th edition, 2006. English translation in Heidegger (1962).

[156] Heidegger, M. (1962). *Being and Time*. Basil Blackwell, Oxford. Translated by J. Macquarrie and E. Robinson. Reprinted by Harper Perennial, 2008.

[157] Hellman, G. (1989). *Mathematics without Numbers*. Oxford University Press, Oxford.

[158] Hellman, G. (2003). Does category theory provide a framework for mathematical structuralism? *Philosophia Mathematica*, **11**(2):129–57.

[159] Hempel, C. G. (1943). A purely syntactical definition of confirmation. *Journal of Symbolic Logic*, **8**(4):122–43.

[160] Hempel, C. G. (1945a). Studies in the logic of confirmation (I.). *Mind*, **54**(213): 1–26. Reprinted in Hempel (1965), pages 3–47.

[161] Hempel, C. G. (1945b). Studies in the logic of confirmation (II.). *Mind*, **54**(214): 97–121. Reprinted in Hempel (1965), pages 3–47.

[162] Hempel, C. G. (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. Free Press, New York.

[163] Hempel, C. G. and Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, **15**(2):135–75. Reprinted in Hempel (1965), pages 245–90.

[164] Herbrand, J. (1932). Sur la non-contradiction de l'arithmetique. *Journal für die reine und angewandte Mathematik*, **166**:1–8.

[165] Hilbert, D. (1926). Über das Unendliche. *Mathematische Annalen*, **95**(1):161–90. English translation by E. Putnam and G. J. Massey in Benacerraf and Putnam (1983), pages 183–201.

[166] Hilbert, D. and Bernays, P. (1934/1939). *Grundlagen der Mathematik*, volume I/II. Springer, Berlin. Second revised edition by same publisher in 1968/1970.

English translation of second edition by the *Hilbert Bernays Project*, part A of I in 2011, College Publications, London, and the rest upcoming.

[167] Hopcroft, J. E., Motwani, R., and Ullman, J. D. (2001). *Introduction to Automata Theory, Languages and Computability*. Addison-Wesley, Boston (MA), 2nd edition.

[168] Horwich, P. (1982). *Probability and Evidence*. Cambridge University Press, Cambridge.

[169] Hsieh, D. A. (1989). Modeling heteroscedasticity in daily foreign exchange rates. *Journal of Business and Economic Statistics*, **7**(3):307–17.

[170] Hsu, F.-H. (2002). *Behind Deep Blue: Building the Computer that Defeated the World Chess Champion*. Princeton University Press.

[171] Hsu, J. (2017). Meet the new AI challenging human poker pros - IEEE Spectrum. URL `http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/meet-the-new-ai-challenging-human-poker-pros` (retrieved: 6 March 2018).

[172] Hughes, R. I. G. (1997). Models and representation. *Philosophy of Science*, **64**: 325–36.

[173] Hurst, H. E. (1951). Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, **116**(1):400–10.

[174] Husserl, E. (1913). *Logische Untersuchungen. Theil 1 & 2*. Niemeyer, Halle, 2nd edition. English translation by J. Findlay in Husserl (2001).

[175] Husserl, E. (2001). *Logical Investigations. Vols. One and Two*. D. Moran (Ed.). Routledge, London. Translation by J. Findlay.

[176] Huxley, T. H. (1874). On the hypothesis that animals are automata, and its history. *The Fortnightly Review (New Series)*, **16**:555–80. Reprinted in *Method and Results: Essays by Thomas H. Huxley*, pages 199–250, 1898, D. Appleton and Company, New York.

[177] IBM. (2001). IBM Research | Deep Blue | Overview. URL `https://www.research.ibm.com/deepblue/meet/html/d.3.2.html` (retrieved: 6 March 2018).

[178] IBM. (2009). IBM - DeepQA Project: FAQs. URL `https://www.research.ibm.com/deepqa/faq.shtml` (retrieved: 6 March 2018).

[179] Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, **32**(127): 127–36.

[180] Jackson, F. (1986). What Mary didn't know. *Journal of Philosophy*, **83**(5):291–5.

[181] Kanamori, A. (2011). ZFC. In *Encyclopedia of Mathematics*. Springer. URL http://www.encyclopediaofmath.org/index.php?title=ZFC&oldid=19298.

[182] Kant, I. (1787). *Critik der reinen Vernunft*. Johann Friedrich Hartknoch, Riga, 2nd, extended edition. Reprinted in 1998, Meiner, Hamburg. English translation by P. Guyer and A. Wood in Kant (1999).

[183] Kant, I. (1999). *Critique of Pure Reason*. Cambridge University Press, Cambridge. Translation by P. Guyer and A. Wood.

[184] Kesten, H. (1973). Random difference equations and renewal theory of products of random matrices. *Acta Mathematica*, **131**(1):207–48.

[185] Keynes, J. M. (1936). *General Theory of Employment, Interest and Money*. Palgrave Macmillan.

[186] Kleene, S. C. (1935). A theory of positive integers in formal logic. Part I and II. *American Journal of Mathematics*, **57**(1-2):153–73 and 219–44.

[187] Kleene, S. C. (1936a). General recursive functions of natural numbers. *Mathematische Annalen*, **112**:727–42. Reprinted in Davis (1965), pages 236–53.

[188] Kleene, S. C. (1936b). λ-definability and recursiveness. *Duke Mathematical Journal*, **2**(2):340–53.

[189] Klemeš, V. (1974). The Hurst phenomenon: a puzzle? *Water Resources Research*, **10**(4):275–88.

[190] Klemke, E. D. (Ed.) (1968). *Essays on Frege*. University of Illinois Press, Chicago.

[191] Klenke, A. (2008). *Probability Theory: A Comprehensive Course*. Springer, Berlin. 2nd edition in 2014.

[192] Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*, volume 2 of *Ergebnisse der Mathematik und Ihrer Grenzgebiete*. Springer, Berlin. English translation in Kolmogorov (1950).

[193] Kolmogorov, A. N. (1950). *Foundations of the Theory of Probability*. Chelsea Publishing Company, New York. Translation by N. Morrison.

[194] Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission*, **1**(1):1–7.

[195] Konrad Zuse Internet Archive. (2017). Z1. URL http://zuse.zib.de/z1 (retrieved: 6 March 2018).

[196] Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago.

[197] Kuratowski, C. (1921). Sur la notion de l'ordere dans la Théorie des Ensembles. *Fundamenta Mathematica*, **2**:161–71.

[198] Kurzweil, R. (2005). *The Singularity Is Near: When Humans Transcend Biology*. Viking Penguin, New York.

[199] Ladyman, J. (2014). Structural realism. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Winter 2016 edition. URL `http://plato.stanford.edu/archives/win2016/entries/structural-realism/`.

[200] Ladyman, J. and Ross, D. (2007). *Every Thing Must Go*. Oxford University Press, Oxford.

[201] Ladyman, J. and Ross, D. (2013). The world in the data. In Ross, D., Ladyman, J., and Kincaid, H. (Eds.), *Scientific Metaphysics*, chapter 6, pages 108–50. Oxford University Press, Oxford.

[202] Lakatos, I. (1978). *The Methodology of Scientific Research Programmes: Philosophical Papers.* Worrall, J. and Currie, G. (Eds.), volume I and II. Cambridge University Press, Cambridge.

[203] Lake, B. M., Salakhutdinov, R., and Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, **350**(6266): 1332–8.

[204] Lambek, J. (1968). Deductive systems and categories: I. Syntactic calculus and residuated categories. *Mathematical Systems Theory*, **2**(4):287–318.

[205] Lambek, J. (1969). Deductive systems and categories II. standard constructions and closed categories. In *Category Theory, Homology Theory and their Applications I*, volume 99 of *Lecture Notes in Mathematics*, pages 76–122. Springer, Berlin.

[206] Lambek, J. (1972). Deductive systems and categories III. cartesian closed categories, intuitionistic propositional calculus, and combinatory logic. In *Toposes, Algebraic Geometry and Logic*, volume 274 of *Lecture Notes in Mathematics*, pages 57–82. Springer, Berlin.

[207] Landry, E. (2011). How to be a structuralist all the way down. *Synthese*, **179**(3): 435–54.

[208] Landry, E. and Marquis, J.-P. (2005). Categories in context: Historical, foundational and philosophical. *Philosophica Mathematica*, **13**(1):1–43.

[209] Lawvere, F. W. (1963). *Functorial Semantics of Algebraic Theories and Some Algebraic Problems in the context of Functorial Semantics of Algebraic Theories*. PhD thesis, Columbia University. Partly published in *Reports of the Midwest Category*

Seminar II, 1968, 41–61 and republished in *Reprints in Theory and Applications of Categories*, **5**:1–121, 2004.

[210] Lawvere, F. W. (1964). An elementary theory of the category of sets. *Proceedings of the National Academy of Sciences U.S.A*, **52**(6):1506–11.

[211] Lebowitz, J. (1993). Boltzmann's entropy and time's arrow. *Physics Today*, **46** (9):32–8.

[212] Lebowitz, J. (1994). Is Boltzmann entropy time's arrow's archer?: Replies to reader's responses to Lebowitz (1993). *Physics Today*, **47**(11):113–6.

[213] Lebowitz, J. (1995). Microscopic reversibility and macroscopic behavior: Physical explanations and mathematical derivations. In Brey, J., Marro, J., Rubí, J., and San Miguel, M. (Eds.), *25 Years of Non-Equilibrium Statistical Mechanics, Proceedings of the XIII Sitges Conference, Barcelona, Spain, 1994*, volume 445 of *Lecture Notes in Physics*, pages 1–20. Springer, Berlin.

[214] Leibniz, G. W. (1714). *La Monadologie*. Unpublished. First published as German translation by H. Köhler in 1720. First original French edition by J. E. Erdmann in 1840, Berlin.

[215] Leland, W. E., Taqqu, M. S., Willinger, W., and Wilson, D. V. (1994). On the self-similar nature of ethernet traffic. *IEEE/ACM Transactions on Networking*, **2** (1):1–15.

[216] Leonelli, S. (2009). On the locality of data and claims about phenomena. *Philosophy of Science*, **76**(5):737–49.

[217] Leonelli, S. (2015). What counts as scientific data? a relational framework. *Philosophy of Science*, **82**(5):810–21.

[218] Lewis, C. I. (1929). *Mind and the World Order: Outline of a Theory of Knowledge*. Charles Scribner's Sons, New York. Reprinted in Dover, New York, 1991.

[219] Li, M. and Vitányi, P. (2008). *An Introduction to Kolmogorov Complexity and Its Applications*. Springer, New York (NY). 3rd edition.

[220] Lichtl, A. and Jones, S. (17–20 March 2015). GPUs to Mars: Full scale simulation of SpaceX's Mars rocket engine. URL `http://on-demand.gputechconf.com/gtc/2015/video/S5398.html` (retrieved: 19 January 2018).

[221] Linnebo, Ø. and Pettigrew, R. (2011). Category theory as an autonomous foundation. *Philosophia Mathematica*, **19**(3):227–54.

[222] Lobato, I. N. and Savin, N. E. (1997). Real and spurious long memory in stock market data. *Journal of Business and Economic Statistics*, **16**(3):261–8.

[223] Lutz, S. (2014). What's right with a syntactic approach to theories and models? *Erkenntnis*, **79**(8):1475–92.

[224] Lux, T. (1996). Long-term stochastic dependence in financial prices: Evidence from the german stock market. *Applied Economic Letters*, **3**(11):701–6.

[225] Mac Lane, S. (1971). *Categories for the Working Mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer, New York. Second edition in 1998.

[226] Machamer, P (Ed.). (2011). Phenomena, data and theories: a special issue of *Synthese*. *Synthese*, **182**(1).

[227] Maddy, P. (1990). *Realism in Mathematics*. Oxford University Press, Oxford.

[228] Mahner, M. (2007). Demarcating science from non-science. In Kuipers, T. (Ed.), *Handbook of the Philosophy of Science: General Philosophy of Science – Focal Issues*, pages 515–75. Elsevier, Amsterdam.

[229] Malkiel, B. G. (1979). The capital formation problem in the United States. *The Journal of Finance*, **34**(2):291–306.

[230] Mandelbrot, B. B. (1963). The variation of certain speculative prices. *Journal of Business*, **36**(4):394–419.

[231] Mandelbrot, B. B. (1969). Long-run linearity, locally Gaussian processes, H-spectra and infinite variances. *International Economic Review*, **10**(1):82–113.

[232] Mandelbrot, B. B. (1971). When can price be arbitraged efficiently? a limit to the validity of random walk and martingal models. *Review of Economics and Statistics*, **53**(3):225–36.

[233] Mandelbrot, B. B. (1983). *The Fractal Geometry of Nature*. Freeman and Co., New York. Revised and enlarged translation from French *Les Objets fractals: Forme, hasard et dimension*, 1975, Flammarion, Paris.

[234] Mandelbrot, B. B. (2004). A theory of roughness. In Brockman, J. (Ed.), *Edge*. URL `http://www.edge.org/3rd_culture/mandelbrot04/mandelbrot04_index.html` (retrieved: 6 March 2018).

[235] Mandelbrot, B. B. and Wallis, J. R. (1968). Noah, Joseph, and operational hydrology. *Water Resources Research*, **4**(5):909–18.

[236] Marcus, G., Rossi, F., and Velos, M. (Eds.). (2016). Beyond the Turing Test. *AI Magazine*, **37**(1).

[237] Marquis, J.-P. (2000). Three kinds of universals in mathematics? In Woods, J. and Brown, B. (Eds.), *Logical Consequence: Rival Approaches and New Studies*

in Exact Philosophy: Logic, Mathematics and Science, volume II, pages 191–212. Hermes, Oxford.

[238] Marquis, J.-P. (2006). Categories, sets and the nature of mathematical entities. In van Benthem, J., Heinzmann, G., Rebuschi, M., and Visser, H. (Eds.), *The Age of Alternative Logics. Assessing philosophy of logic and mathematics today*, Logic, Epistemology, and the Unity of Science, pages 181–92. Springer, Dordrecht.

[239] Marquis, J.-P. (2013). Mathematical forms and forms of mathematics: Leaving the shores of extensional mathematics. *Synthese*, **190**(12):2141–64.

[240] Marquis, J.-P. (2014). Category theory. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Winter 2015 edition. URL `http://plato.stanford.edu/archives/win2015/entries/category-theory/`.

[241] Martin-Löf, P. (1970). *Notes on Constructive Mathematics*. Almqvist & Wiksell, Stockholm.

[242] Martin-Löf, P. (1973). An intuitionistic theory of types: Predicative part. In Rose, H. and Shepherdson, J. (Eds.), *Proceedings of the logic colloquium, Bristol, July 1973*, volume 80 of *Studies in logic and the foundations of mathematics*, pages 73–118. North-Holland, Amsterdam. Later review in *Journal of Symbolic Logic*, **49**(1):311–3, 1984.

[243] Massimi, M. (2008). Why there are no ready-made phenomena: What philosophers of science should learn from Kant. *Royal Institute of Philosophy Supplement*, **83** (63):1–35.

[244] Massimi, M. (2011). From data to phenomena: a Kantian stance. *Synthese*, **182** (1):101–16.

[245] McAllister, J. W. (1997). Phenomena and patterns in data sets. *Erkenntnis*, **47** (2):217–28.

[246] McAllister, J. W. (2011). What do patterns in empirical data tell us about the structure of the world? *Synthese*, **182**(1):73–87.

[247] McCurdy, T. H. and Morgan, I. G. (1988). Testing the martingale hypothesis in Deutsche Mark futures with models specifying the form of heteroskedasticity. *Journal of Applied Econometrics*, **3**(3):187–202.

[248] McDowell, J. (1994). *Mind and World*. Harvard University Press, Cambridge. Reprinted in 1996 and 2003.

[249] McLarty, C. (1993). Numbers can be just what they have to. *Noûs*, **27**(4):487–98.

[250] McLarty, C. (2011). Recent debate over categorical foundations. In Sommaruga, G. (Ed.), *Foundational Theories of Classical and Constructive Mathematics*, pages 145–54. Springer, New York.

[251] McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.

[252] Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, **4**(1):141–83.

[253] Merton, R. K. (1973). *The Sociology of Knowledge*. University of Chicago Press, Chicago.

[254] Minsky, M. L. (1967). *Computation: Finite and Infinite Machines*. Prentice-Hall, Engelwood Cliffs (NJ).

[255] Moore, A. W. (1990). *The Infinite*. Routledge, London. 2nd edition in 2001.

[256] Moran, D. (2000). *Introduction to Phenomenology*. Routledge, New York.

[257] Mukhopadhyay, N. (2006). A conversation with Ulf Grenander. *Statistical Science*, **21**(3):404–26.

[258] Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. S. A. Nelson, New York.

[259] Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, **83**(4):435–50.

[260] Neurath, O. (1932). Protokollsätze. *Erkenntis*, **3**(1):204–14.

[261] Nobel Media AB 2014, *nobelprize.org*. (1997). The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 1997, URL `http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1997/` (retrieved: 8 November 2016).

[262] Nobel Media AB 2014, *nobelprize.org*. (2003). The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2003, URL `http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2003/` (retrieved: 8 November 2016).

[263] NVIDIA. (2017). About NVIDIA brochure. URL `http://images.nvidia.com/content/pdf/about-nvidia/nvidia-2017-updated.pdf` (retrieved: 4 August 2017).

[264] Oetting, W. S., Fryer, J. P., and King, R. A. (1998). Mutations of the human tyrosinase gene associated with tyrosinase related oculocutaneous albinism (OCA1). *Human Mutation*, **12**(6):433–4. Mutations in brief no. 204. Online.

[265] Papineau, D. (Ed.) (1996). *The Philosophy of Science*. Oxford Readings in Philosophy. Oxford University Press, Oxford.

[266] Parsons, C. (1990). The structuralist view of mathematical objects. *Synthese*, **84** (3):303–46.

[267] Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, **2**(11):559–72.

[268] Peterson, J. B. (2017). *The Psychological Significance of the Biblical Stories*. Internet lecture series, YouTube.

[269] Pindyck, R. S. (1984). Risk, inflation, and the stock market. *American Economic Review*, **74**(3):335–51.

[270] Poincaré, H. (1902). *La Science et l'Hypothèse*. Bibliothèque de philosophie scientifique. Flammarion, Paris. English translation as *Science and Hypothesis* by W.J.G. [only initials indicated], The Walter Scott Publishing Co., London, 1905 and (reprint) Dover, New York, 1952.

[271] Poincaré, H. (1905). *La Valeur de la Science*. Bibliothèque de philosophie scientifique. Flammarion, Paris. English translation as *The Value of Science* by G. B. Halsted, The Science Press, New York, 1907, and (reprint) Dover, New York, 1958.

[272] Popper, K. (1935). *Logik der Forschung*. Springer, Wien. English translation in Popper (1959).

[273] Popper, K. (1954). Degree of confirmation. *The British Journal for Philosophy of Science*, **5**(18):143–9.

[274] Popper, K. (1957). A second note on degree of confirmation. *The British Journal for Philosophy of Science*, **7**(28):350–3.

[275] Popper, K. (1958). A third note on degree of corroboration or confirmation. *The British Journal for Philosophy of Science*, **8**(32):294–302.

[276] Popper, K. (1959). *The Logic of Scientific Discovery*. Hutchinson, London.

[277] Popper, K. (1962). *Conjectures and refutations. The growth of scientific knowledge*. Basic Books, New York.

[278] Preus, A. (2007). *Historical Dictionary of Ancient Greek Philosophy*. The Scarecrow Press, Lanham, Maryland.

[279] Putnam, H. (1967). The nature of mental states. In Capitan, W. H. and Merrill, D. D. (Eds.), *Art, Mind, and Religion*, pages 37–48. Pittsburgh University

Press, Pittsburg. Originally published as “Psychological Predicates”. Reprinted in Chalmers, D. J. (Ed.), *Philosophy of Mind: Classical and Contemporary Readings*, 2002, pages 73–79, Pittsburgh University Press.

[280] Putnam, H. (1971). *Philosophy of Logic*. Harper & Row, New York.

[281] Putnam, H. (1975). *Mathematics, Matter and Method*. Cambridge University Press, Cambridge.

[282] Quine, W. V. O. (1948). On What There Is. *Review of Metaphysics*, **2**(5):21–38. Reprinted in Quine (1961).

[283] Quine, W. V. O. (1961). *From a Logical Point of View*. Harvard University Press. 2nd revised edition with several reprints, inter alia, 1980. 1st edition in 1953.

[284] Quine, W. V. O. (1969). Epistemology naturalized. In *Ontological Relativity and Other Essays*, pages 69–90. Columbia University Press, New York.

[285] Quine, W. V. O. (1981). *Theories and Things*. Harvard University Press.

[286] Raatikainen, P. (2015). Gödel’s incompleteness theorems. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Spring 2015 edition. URL `https://plato.stanford.edu/archives/spr2015/entries/goedel-incompleteness/`.

[287] Reck, E. and Price, P. (2000). Structures and structuralism in contemporary philosophy of mathematics. *Synthese*, **125**(3):341–83.

[288] Regal, B. (2009). *Pseudoscience: A Critical Encyclopedia*. Greenwood, Santa Barbara (CA).

[289] Resnik, M. (1997). *Mathematics as a Science of Patterns*. Oxford University Press.

[290] Robb, D. and Heil, J. (2013). Mental causation. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Spring 2014 edition. URL `https://plato.stanford.edu/archives/spr2014/entries/mental-causation/`.

[291] Rosen, G. (2012). Abstract objects. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Fall 2014 edition. URL `http://plato.stanford.edu/archives/fall2014/entries/abstract-objects/`.

[292] Ross, D., Ladyman, J., and Collier, J. (2007). Rainforest realism and the unity of science. In Ladyman and Ross (2007), pages 190–257.

[293] Russell, B. (1903). *The Principles of Mathematics*. Cambridge University Press, Cambridge.

[294] Russell, B. (1908). Mathematical logic as based on the theory of types. *American Journal of Mathematics*, **30**(3):222–62.

[295] Russell, B. (1927). *The Analysis of Matter*. Kegan Paul, London. Several reprints, inter alia, 2014, Martino, Mansfield Centre.

[296] Russell, S. and Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Prentice Hall, Englewood Cliffs (NJ), 3rd edition. According to the authors "the leading textbook in Artificial Intelligence".

[297] Schlick, M. (1934). Über das Fundament der Erkenntnis. *Erkenntnis*, **4**(1):79–99.

[298] Schrinner, A. (2008). Volkswagenaktien: Schein und Haben, Handelsblatt, October 28th, URL `http://www.handelsblatt.com/politik/deutschland/aktienkurs-volkswagenaktien-schein-und-haben/3043788.html` (retrieved: 6 March 2018).

[299] Schupbach, J. N. (2011). Comparing probabilistic measures of explanatory power. *Philosophy of Science*, **78**(5):813–29.

[300] Schupbach, J. N. and Sprenger, J. (2011). The logic of explanatory power. *Philosophy of Science*, **78**(1):105–27.

[301] Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, **3**(3):417–24.

[302] Searle, J. R. (2014). What your computer can't know. *The New York Review of Books*, **61**(15).

[303] Sellars, W. (1956). Empiricism and the philosophy of mind. In Feigl, H. (Ed.), *Minnesota Studies in the Philosophy of Science*, volume 1 (*Foundations of Science & the Concepts of Psychology and Psychoanalysis*), pages 253–329. University of Minnesota Press. Republished by Harvard University Press, Cambridge, in 1956 and 1997 with an Introduction by R. Rorty and a Study Guide by R. Brandom.

[304] Shannon, C. E. (1948a). A mathematical theory of communication, part I. *Bell System Technical Journal*, **27**(3):379–23.

[305] Shannon, C. E. (1948b). A mathematical theory of communication, part II. *Bell System Technical Journal*, **27**(4):623–56.

[306] Shapiro, S. (1997). *Philosophy of Mathematics: Structure and Ontology*. Oxford University Press, New York.

[307] Shapiro, S. (2005). Categories, structures and frege-hilbert controversy: the status of metamathematics. *Philosophia Mathematica*, **13**(1):61–77.

[308] Shapiro, S. (2007). *The Oxford Handbook of Philosophy of Mathematics and Logic*. Oxford University Press.

[309] Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, **71**(3):421–36.

[310] Shiller, R. J. (1981). The use of volatility measures in assessing market efficiency. *Journal of Finance*, **36**(2):291–304.

[311] Shoemaker, S. (2001). Realization and mental causation. In Gillett, C. and Loewer, B. M. (Eds.), *Physicalism and Its Discontents*, pages 74–98. Cambridge University Press, Cambridge. Reprinted in Shoemaker (2003) *Identity, Cause, and Mind (Expanded Edition)*, Clarendon Press, Oxford, pages 427–51.

[312] Shope, R. K. (1983). *The Analysis of Knowing: A Decade of Research*. Princeton University Press, Princeton.

[313] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., and Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, **529**:484–9.

[314] Skolem, T. (1923). Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre. In *Matematikerkongressen i Helsingfors den 4-7 Juli, 1922. Den femte Skandinaviska Matematikerkongressen*, pages 217–32. Sana Helsingfors. English translation by S. Bauer-Mengelberg in van Heijenoort1967 (1967), pages 290–301.

[315] Smith, D. W. (2008). Phenomenology. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Fall 2008 edition. URL `http://plato.stanford.edu/archives/fall2008/entries/phenomenology/`.

[316] Sneed, J. D. (1971). *The Logical Structure of Mathematical Physics*. Reidel, Dordrecht.

[317] Sober, E. (2015). *Ockham's Razor: A User's Manual*. Cambridge University Press.

[318] Sornette, D. (2006). *Critical Phenomena in Natural Sciences*. Springer, Berlin.

[319] Spiegelberg, H. (1994). *The Phenomenological Movement: A Historical Introduction*. Springer, Dordrecht, 3rd revised edition.

[320] Spritz, R. A., Oh, J., Fukai, K., Holmes, S. A., Ho, L., Chitayat, D., France, T. D., Musarella, M. A., Orlow, S. J., Schnur, R. E., Weleber, R. G., and Levin, A. V. (1997). Novel mutations of the tyrosinase (TYR) gene in type I oculocutaneous albinism (OCA1). *Human Mutation*, **10**(2):171–4.

[321] Stanford Vision Lab. (2016). ImageNet. URL `http://image-net.org` (retrieved: 6 March 2018).

[322] Ströing, P. (2012). Das Multifraktalmodell für Anlagerenditen zum Risikomanagement von Aktienpositionen. Diploma thesis, Ludwig-Maximilians-Universität München.

[323] Ströing, P. (2018). Data, evidence and explanatory power. *Philosophy of Science*, **85**(3):422–41.

[324] Suppes, P. (1960). A comparison of the meaning and uses of models in mathematics and the empirical sciences. *Synthese*, **12**(3):287–301. Reprinted in Freudenthal (1961), pages 163–77, and in Suppes (1969), pages 10–23.

[325] Suppes, P. (1960). *Axiomatic Set Theory*. Van Nostrand, New York. Slightly revised edition published by Dover, New York, 1972.

[326] Suppes, P. (1962). Models of data. In Nagel, E., Suppes, P., and Tarski, A. (Eds.), *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress*, pages 252–61. Stanford University Press, Stanford. Reprinted in Suppes (1969), pages 24–35.

[327] Suppes, P. (1969). *Studies in the Methodology and Foundations of Science: Selected Papers from 1951 to 1969*, volume 22 of *Synthese Library*. Springer, Dordrecht.

[328] Suppes, P. (2002). *Representation and Invariance of Scientific Structures*. CSLI Publications.

[329] Suárez, M. (1999). Theories, models, and representations. In Magnani, L., Nersessian, N. J., and Thagard, P. (Eds.), *Model-Based Reasoning in Scientific Discovery*, pages 75–83. Springer, New York.

[330] Suárez, M. (2003). Scientific representation: Against similarity and isomorphism. *International Studies in the Philosophy of Science*, **17**(3):225–44.

[331] Suárez, M. (2004). An inferential conception of scientific representation. *Philosophy of Science*, **71**(5):767–79.

[332] Suárez, M. (2015). Deflationary representation, inference, and practice. *Studies in History and Philosophy of Science Part A*, **49**(Supplement C):36–47.

[333] Swart, E. R. (1980). The philosophical implications of the four-color problem. *The American Mathematical Monthly*, **87**(9):697–707.

[334] Taqqu, M. (2001). Bachelier and his times: A conversation with Bernard Bru. *Finance and Stochastics*, **4**(1):3–32.

[335] Tarski, A. (1953). *Undecidable theories*. North-Holland, Amsterdam.

[336] Tarski, A. (1994). *Introduction to Logic and the Methodology of the Deductive Sciences*, volume 24 of *Oxford Logic Guides*. Oxford University Press, 4th edition. Revised edition of O. Helmer's translation from the author's Polish (1936) and German (1937) version, 1941.

[337] Teller, P. (2001). Whither constructive empiricism? *Philosophical Studies*, **106** (1-2):123–50.

[338] Thom, R. (1971). "Modern" mathematics: an educational and philosophical error? *American Scientist*, **59**(6):695–99.

[339] Tierney, M. (1972). Sheaf theory and the continuum hypothesis. In Lawvere, F. W. (Ed.), *Toposes, Algebraic Geometry and Logic*, volume 274 of *Lecture Notes in Mathematics*, pages 13–42. Springer, Berlin.

[340] Tintner, G. (1940). *The Variate Difference Method*. Principia Press, Bloomington, Indiana.

[341] Tomita, Y., Takeda, A., Okinaga, S., Tagami, H., and Shibahara, S. (1989). Human oculocutaneous albinism caused by single base insertion in the tyrosinase gene. *Biochemical and Biophysical Research Communications*, **164**(3):990–6.

[342] Tsai, Y., Sawaya, M. R., Cannon, C. G., Cai, F., Williams, E. B., Heinhorst, S., Kerfeld, C. A., and Yeates, T. O. (2007). Structural analysis of CsoS1A and the protein shell of the *Halothiobacillus neapolitanus carboxysome*. *PLoS Biology*, **5** (6):e144.

[343] Turing, A. M. (1937). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, **s2-42** (1):230–65. Correction *ibid.*, (1938) **s2-43**(1):544–6. Reprinted in Davis (1965), pages 115–53.

[344] Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, **59**(236): 433–60.

[345] Tymoczko, T. (1979). The four-color problem and its philosophical significance. *The Journal of Philosophy*, **76**(2):57–83.

[346] van Atten, M. (2004). *On Brouwer*. Wadsworth, Belmont (CA).

[347] van Fraassen, B. (1985). Empiricism in the philosophy of science. In Hooker, C. A. and Churchland, P. M. (Eds.), *Images of Science: Essays on Realism and Empiricism with Replies from Bas C. Van Frassen*, pages 245–308. University of Chicago Press, Chicago.

[348] van Fraassen, B. C. (1980). *The Scientific Image*. Oxford University Press, Oxford.

[349] van Fraassen, B. C. (2001). Constructive empiricism now. *Philosophical Studies*, **106**(1-2):151–170.

[350] van Fraassen, B. C. (2008). *Scientific Representation: Paradoxes of Perspective*. Oxford University Press, Oxford.

[351] van Heijenoort, J. (1967). *From Frege to Gödel: A Source Book in Mathematical Logic, 1979–1931*. Harvard University Press, Cambridge. Reprinted in 2002 by same publisher.

[352] Vitányi, P. and Li, M. (2000). Minimum description length induction, Bayesianism, and Kolmogorov complexity. *IEEE Transactions on Information Theory*, **46**(2): 446–64.

[353] Vollmer, G. (1992). Wozu Pseudowissenschaften gut sind. Argumente aus Wissenschaftstheorie und Wissenschaftspraxis. *Universitas*, **47**(Feb.):155–68. Reprinted in Vollmer, G., *Wissenschaftstheorie im Einsatz* (1993) Hirzel-Verlag, Stuttgart, pages 11–34.

[354] von Neumann, J. (1923). Zur Einführung der transfiniten Zahlen. *Acta Scientiarum Mathematicarum (Szeged)*, **1**(4):199–208. English translation by S. Bauer-Mengelberg in van Heijenoort (1967), pages 346–54.

[355] von Neumann, J. (1925). Eine Axiomatisierung der Mengenlehre. *Journal für die reine und angewandte Mathematik*, **154**:219–40. English translation by S. Bauer-Mengelberg and D. Føllesdal in van Heijenoort (1967), pages 393–413.

[356] Waldermann, A. (2009). Fusion statt Übernahme: Wiedeking scheitert mit Angriff auf Volkswagen, SPIEGEL ONLINE, May 6th, URL `http://www.spiegel.de/wirtschaft/fusion-statt-uebernahme-wiedeking-scheitert-mit-angriff-auf-volkswagen-a-623308.html` (retrieved: 6 March 2018).

[357] Wallace, D. (2003). Everett and structure. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, **34**(1): 87–106.

[358] Wang, Y., Guo, X., Li, W., and Lian, S. (2009). Four novel mutations of TYR gene in chinese OCA1 patients. *Journal of Dermatological Science*, **53**:80–1.

[359] Watkins, E. (2008). Kant and the myth of the given. *Inquiry*, **51**(5):512–31.

[360] Wegner, D. M. (2002). *The Illusion of Conscious Will*. MIT Press, Cambridge (MA).

[361] Wegner, D. M. (2004). Précis of *The Illusion of Conscious Will*. *Behavioral and Brain Sciences*, **27**(5):649–59.

[362] Weiner, J. (1995). Realism *bei* Frege: Reply to Burge. *Synthese*, **102**(3):363–82.

[363] Weiß, M. G. (2009). Die Auflösung der menschlichen Natur. In Weiß, M. G. (Ed.), *Bios und Zoë: Die menschliche Natur im Zeitalter ihrer technischen Reproduzierbarkeit*, pages 34–54. Suhrkamp, Frankfurt a. M.

[364] Wiener, N. (1914). A simplification of the logic of relations. *Proceedings of Cambridge Philosophical Society*, **17**:387–90. Reprinted in van Heijenoort (1967), pages 224–7.

[365] Wikipedia. (2017). Beobachtungssatz, URL `https://de.wikipedia.org/wiki/Beobachtungssatz` (retrieved: 12 September 2017).

[366] Wikipedia. (2018). Comparison of orbital rocket engines, URL `https://en.wikipedia.org/wiki/Comparison_of_orbital_rocket_engines` (retrieved: 19 January 2018).

[367] Willinger, W., Paxson, V., Riedi, R. H., and Taqqu, M. S. (2003). Long-range dependence and data network traffic. In Doukhan *et. al.* (2003), pages 373–407.

[368] Winther, R. G. (2016). The structure of scientific theories. In Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Winter 2016 edition. URL `https://plato.stanford.edu/archives/win2016/entries/structure-scientific-theories/`.

[369] Wittgenstein, L. (1953). *Philosophical Investigations.* Anscombe, G. E. M. and Rhees, R. (Eds.). Blackwell, Oxford. Posthumously first published from a German manuscript from 1936–1946 translated by Anscombe, original text in (370). Several reprints and revised editions, inter alia, 2001, Blackwell, Hoboken.

[370] Wittgenstein, L. (2001). *Philosophische Untersuchungen. Kritisch-genetische Edition.* J. Schulte (Ed). Suhrkamp, Frankfurt.

[371] Wolfram, S. (2015). Wolfram Language Artificial Intelligence: The Image Identification Project—Stephen Wolfram Blog. URL `http://blog.stephenwolfram.com/2015/05/wolfram-language-artificial-intelligence-the-image-identification-project` (retrieved: 6 March 2018).

[372] Wolfram Research. (2015a). The Wolfram Language Image Identification Project. URL `https://www.imageidentify.com` (retrieved: 6 March 2018).

[373] Wolfram Research. (2015b). How the Wolfram Language Image Identification Project Works. URL `https://www.imageidentify.com/about/how-it-works` (retrieved: 6 March 2018).

[374] Woodward, J. F. (1989). Data and phenomena. *Synthese*, **79**(3):393–472.

[375] Woodward, J. F. (2000). Data, phenomena, and reliability. *Philosophy of Science*, **67**(Supplement):163–79. *Proceedings of the 1998 Biennial Meetings of the Philosophy of Science Association. Part II: Symposia Papers (Sep., 2000).*

[376] Woodward, J. F. (2011). Data and phenomena: A restatement and defense. *Synthese*, **182**(1):165–79.

[377] Worrall, J. (1989). Structural realism: The best of both worlds? *Dialectica*, **43** (1-2):99–124. Reprinted in Papineau (1996), pages 139–65.

[378] Zermelo, E. (1904). Beweis, dass jede Menge wohlgeordnet werden kann: Aus einem an Herrn Hilbert gerichteten Briefe. *Mathematische Annalen*, **59**(4):514–6. English translation by S. Bauer-Mengelberg in van Heijenoort (1967), pages 139–41.

[379] Zermelo, E. (1908). Neuer Beweis für die Möglichkeit einer Wohlordnung. *Mathematische Annalen*, **65**(1):107–28. English translation by S. Bauer-Mengelberg in van Heijenoort (1967), pages 183–98.

[380] Zermelo, E. (1908). Untersuchungen über die Grundlagen der Mengenlehre I. *Mathematische Annalen*, **65**(2):261–81. English translation by S. Bauer-Mengelberg in van Heijenoort (1967), pages 199–215.

[381] Zermelo, E. (1930). Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre. *Fundamenta Mathematicae*, **16**:29–47. English translation by M. Hallett in Ewald (1996), pages 1208–33.

Subject Index

The *Subject Index* lists places at which the respective subjects are either introduced or at which the subject plays an important role for the philosophical reasoning. The following philosophical subjects are not listed, because they play an important throughout the majority of the thesis: *phenomenon*, *relation*.

Person Index

The *Person Index* lists places at which I mention or refer to quotes or ideas from the named individuals to a relevant extent. The following individuals are not listed, because I refer to their works throughout the majority of the thesis: *James Bogen*, *James F. Woodward*.

Mathematical Notation

The mathematical notation used in the text is listed below. These symbols origin from different branches of mathematics, such as set theory, analysis, theory of probability and stochastic processes, and statistics. I list only symbols that may be ambiguous and are not explicitly introduced in the text elsewhere.

$\{a,b\}$	set with elements a and b
$\mathbb{N}$	set of natural numbers $\{1,2,3,...\}$
$\mathbb{N}_0$	natural numbers unified with $\{0\}$
$\mathbb{N}_{\leq a}, \mathbb{N}_{\geq a}$	natural numbers less/greater than or equal to $a \in \mathbb{R}$
$\mathbb{Z}$	set of integers $\{..., -2, -1, 0, 1, ...\}$
$\mathbb{Q}$	set of rational numbers, *i.e.* $\frac{z_1}{z_2}$ with $z_1, z_2 \in \mathbb{Z}$
$\mathbb{R}, \mathbb{R}_{\geq 0}, \mathbb{R}_{>0}$	set of real -/non-negative real -/positive real numbers
$\emptyset$	empty set $\{\}$
2^M	power set (*i.e.* set of all subsets) of a set M
$\lvert M \rvert$	cardinal number of a set M
$M \setminus N$,	set difference
$M \subset N, M \supset N$	set M is a strict subset/strict superset to set N
$M \subseteq N, M \supseteq N$	set M is subset or identical/superset or identical to set N
$]a,b[$, $[a,b]$, $]a,b]$, $[a,b[$	$a, b, \in \mathbb{R}$: open/closed/half-open interval on $\mathbb{R}$
$(e_1, ..., e_n)$	ordered list of $n \in \mathbb{N}$ elements from $(e_i)_{i \in \{1,...,n\}}$
A^n	ordered list of $n \in \mathbb{N}$ elements from set A
$M \times N$	Cartesian product of sets, *i.e.* $\{(m,n) \mid m \in M, n \in N\}$
$\vec{v}$	variable for vector/ordered list (transposition indistinct)
$(X_i)_{i \in I}$	family $\{X_i : i \in I\}$ of mathematical objects with an index set I
$\lvert x \rvert$	absolute value of an $x \in \mathbb{R}$
$X \sim \mu$	random variable X has a distribution μ in some unambiguous probability space
$\mathcal{L}(X)$	distribution of random variable X in some unambiguous probability space
$\mathcal{N}(\mu, \sigma^2)$	normal distribution in the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with expected value μ and variance σ^2

$\mathrm{var}(X)$	variance of a random variable X
$\mathrm{cov}(X, Y)$	covariance of two random variables X, Y, **or** estimated empirical covariance of two time series $(X_t)_{t\in\{1,2,\ldots,N\}}, (Y_t)_{t\in\{1,2,\ldots,N\}}, N \in \mathbb{N}$
$\mathrm{corr}(X, Y)$	correlation of two random variables X, Y, **or** estimated empirical correlation of two time series $(X_t)_{t\in\{1,2,\ldots,N\}}, (Y_t)_{t\in\{1,2,\ldots,N\}}, N \in \mathbb{N}$
$\mathbb{E}^{\mathbb{P}}[X]$	expected value of a random variable X under probability measure $\mathbb{P}$
$\mathbb{E}[X]$	expected value of a random variable X, if applied probability measure is unambiguous