A first step in the AI implementation of pattern recognition as a solution to the access problem for non-eliminative structuralism

Utrecht University

Marieke Gelderblom (5717035)
Supervisor: dr. Johannes Korbmacher
Second reader: dr. Colin Caret

Bachelor thesis (7,5 ECTS)
Bachelor Artificial Intelligence
Faculty of Humanities
January 30, 2020

Contents

1 Introduction

This thesis is about pattern recognition as a solution to the access problem for non-eliminative structuralism. We propose a first attempt for implementing this using artificial intelligence, which will provide new perspectives on the suitability of this solution.

The brief description of the topic of this thesis above might include some unfamiliar terms to those not too acquainted with philosophy of mathematics. Despite this, we can describe the underlying topic of this thesis, without using technical terms, in a single question: 'What is the nature of mathematical objects?' This is an *ontological* question: a philosophical question about the status of being and existence. It is the central question of this thesis and does not have a straightforward answer. There are many different kinds of objects that we call 'mathematical'. A small selection of such objects include numbers (integers, real numbers...), points, functions, sets, relations, permutations, graphs (including nodes and edges), lines, vectors, groups, rings... Although the list does not end here, it does show the wide range of objects that fall into this category of being 'mathematical'.

Now we have seen some examples of mathematical objects, we can return to an examination of their ontology. Different philosophers of mathematics have taken different positions. In chapter 2 these positions are discussed, with a focus on a position that is called *structuralism*. Structuralists maintain that the ontological status of mathematical objects is not determined by their internal nature, but by the way they relate to each other. These relations form what we call *structures* (hence the name 'structuralism'). Finally, chapter 2 concentrates on one particular structuralist position that we will assume throughout this thesis. This position is called *non-eliminative structuralism*.

We will see in chapter 3 that the non-eliminative structuralism that we assume has to provide answers to some *epistemological* questions: it has to provide a theory of how mathematical knowledge can be obtained. We believe that at least some of us have justified knowledge of mathematical objects – but how can this knowledge be obtained? The corresponding problem is called the *access problem*. We examine a possible solution, based on a process using pattern recognition. After a detailed analysis of this process we argue why criticism on this solution (as provided by MacBride) does not hold.

Chapter 4 builds on pattern recognition as a method of obtaining knowledge of mathematical objects. Now that we have examined how this can be done by a human subject, we want to describe a similar process for an artificial learner. We use techniques of artificial intelligence to propose a first step for the implementation of the process. This includes discussing the possibilities for a formal representation of structures, a suitable reformulation of the intended task and an analysis of the result. The formalization that is required to do so will allow us to consider new perspectives on the suitability of pattern matching as a solution to the access problem.

We conclude our analysis in chapter 5. This chapter includes an overview of the implications of our research and gives some suggestions for further research.

This thesis thus elaborates on a debate in the literature, but looks at the topic from a new angle by looking at an AI implementation. By doing so it provides a contribution to the ongoing discussion.

Although a better understanding of the ontological status of mathematical objects is on itself definitely relevant for the field of artificial intelligence, this is not the main connection between the current research and its background in AI. The main purpose of this thesis is to explore the first steps of an AI implementation of pattern recognition – and by doing so, we will use the techniques of artificial intelligence as a tool for the philosophical argumentation. Formalizing a process in this way will point out flaws (and/or strengths) of the argumentation, as it does not allow any ambiguity.

2 Mathematical structuralism

In this thesis, we discuss the objects that mathematicians are interested in. Many objects seem to qualify as being (part of) the subject matter of mathematics. Sets, functions, relations but also lines and graphs take an important place. But how can we describe this diverse range of entities? What are they, and how can we have knowledge about them? In other words, what can we subsequently say about the ontological and epistemological status of mathematical objects?

The answers to these questions will play an important role in the rest of this thesis. We will take this chapter to elaborate on the position that we will take in the rest of this thesis, called *non-eliminative structuralism*. However, before we can get to this, we will have to set out the involved problems in some more detail.

2.1 Three positions: realism, intuitionism and nominalism

Keep in mind the examples of mathematical objects that we just named: numbers, sets, functions, relations, lines, graphs et cetera. A first observation is that these objects are quite different from everyday objects. Compare the number 3 to a pen. I can say that I have the pen in my hands at this moment, but such a thing will never be applicable to the number 3. This is because mathematical objects (like the number 3) are not located in space and time, as pens obviously are. We say that mathematical objects are *abstract*, with which we will mean that they are not spatially or temporally localized. Furthermore abstract entities are not able to exert any causal powers [23].

In the discussion regarding the ontological status of these abstract mathematical objects, we can determine three schools of thought [34, p. 25]. Firstly, one can take a *realist* position, as for example for example done by Frege [7], Gödel [11] and Russell [36]. Proponents of mathematical realism believe that mathematical entities are as real as physical objects and exist independent from the mathematician. A second view is called *intuitionism*. Intuitionists believe that mathematical entities exist, but that they are dependent of the mathematicians mind and/or his language, conventions, et cetera. Examples of intuitionists are Brouwer [3], Heyting [16] and Dummett [5]. Finally, there are *nominalists*, like Hilbert [17], who deny the existence of mathematical entities at all. Their aim is to reformulate mathematics in such a way that it can do without assuming the existence of these entities.[1]

We can categorize these three positions by looking at their claims about the *existence* and *independence* of mathematical objects. Realism asserts both these claims. Intuitionism, then, asserts existence but not independence. Finally, nominalism denies both. Therefore the questions of existence and the independence of this existence provide relevant criteria to divide the three positions. In this thesis we will focus on a realist approach. But by adhering to the belief that mathematical objects exist independently, we do have to specify what these objects are. This will be the topic of the next sections.[2]

2.2 Realism and sets

Recall that we named some obvious examples of mathematical entities in the beginning of this chapter: numbers, sets, functions, lines, graphs. Instead of studying each of these different objects separately, it was discovered that these could all be reduced to (or modelled in) set theory [34, p. 265]. This enables us to interpret, for example, the arithmetic language via set theory. In this set theoretic reduction, the same theorems can be proven as in the arithmetic language itself [34, p. 42]. This encouraged mathematicians to reduce any kind of mathematical entity to sets. Talking of sets, numbers, points and so on does not provide any advantage when talking of only sets suffices [34, p. 265]. Therefore, many philosophers of science hold that the foundations of mathematics can be found in set theory [18]; this position is for example adhered by Maddy [26]. We have thus found one possible answer to the question what mathematical objects are: sets! One realist position can now be formulated as the position that holds that sets exist and do so independently.[3]

We will take a closer look at the implications of such reductions for the natural numbers. There are multiple systems in set theory we can construct such that they exemplify the natural number system. The two most well-known are the finite *Zermelo ordinals* and the finite *Von Neumann ordinals*. We will briefly discuss these reductions.

Following Ernst Zermelo, we take $\emptyset$ for 0 and define for each number n its successor to be $\{n\}$. In this way, we obtain $\emptyset$ for 0, $\{\emptyset\}$ for 1, $\{\{\emptyset\}\}$ for 2, $\{\{\{\emptyset\}\}\}$ for 3, et cetera. In other words, we could say that we start with $\emptyset$ for 0 and define the successor function $s : n \to \{n\}$. The natural numbers are now recursively defined as the smallest set containing $\emptyset$ that is also closed under the

[1]We often make a distinction between realist and *anti-realist* positions. Intuitionism and nominalism are the two most important examples of such anti-realist positions.

[2]There are certain epistemological questions related to the realist view as well, some of which we will get back to later in chapter 2.

[3]We should note that not all realists agree on this set-theoretic reduction and that there are also other realist positions. Criticism is for example articulated by Nelson Goodman [13].

given successor function [29].

John von Neumann proposed a different construction. In this construction, we also take $\emptyset$ for 0, but define for each number n its successor as the set of all numbers less than n. This way, we obtain $\emptyset$ for 0, $\{\emptyset\}$ for 1, $\{\emptyset, \{\emptyset\}\}$ for 2, $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ for 3, et cetera. A different description would be to define $\emptyset$ for 0 but now have the successor function $s : n \longrightarrow n \cup \{n\}$. Again, the natural numbers are recursively defined as the smallest set containing $\emptyset$ that is closed under this successor function [29].

It may be clear that both the Zermelo and the Von Neumann ordinals describe the natural numbers. We can also see that these descriptions are different from each other, as different claims are true for both constructions. An example is the statement that $1 \in 3$. This is true for the Zermelo ordinals but not for the Von Neumann ordinals. Therefore we can conclude that there is more that one way to describe the natural numbers with a set. There are many set-theoretic reductions that can be used to describe the natural numbers. In fact, there are infinitely many sets that describe the natural numbers.[4]

2.3 Multiple set-theoretic descriptions

Having multiple sets to describe one single object poses a problem for the realists committed to set-theoretic reduction. If the number 3 exists and does so independently of the mathematician, then it must be that either $1 \in 3$ or $1 \notin 3$. It is impossible for the two statements to be both true at the same time. Therefore the number 3 cannot be both represented by the Zermelo ordinals and the Von Neumann ordinals. It is not possible to be both $\{\{\{\emptyset\}\}\}$ and $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$. We must choose the one or the other. Benacerraf describes this problem clearly:

> If numbers are sets, then they must be particular sets, for each set is some particular set. But if the number 3 is really one set rather than another, it must be possible to give some cogent reason for thinking so; for the position that this is an unknowable truth is hardly tenable. But there seems to be little to choose among the accounts. [1, p. 62]

We have no grounds to prefer one representation over the other. Benacerraf convincingly argues that the number 3 is neither of these representations. In fact, no set-theoretic or other representation will *be* the number 3. The realist has the problem of describing what mathematical objects are, as he adheres to the belief that these objects exist independently. One possible answer that seemed promising was to answer this via set theoretic foundations – but it appears that this might not

[4] One simple way to see that there are infinitely many sets describing the natural numbers would be to change the Zermelo ordinals slightly: replace the successor function by $f : x \longrightarrow \{\{x\}\}$. Obviously we could also replace it by $f : x \longrightarrow \{\{\{x\}\}\}$, or by $f : x \longrightarrow \{\{\{\{x\}\}\}\}$, et cetera. Alternatively we could choose not to start with taking $\emptyset$ for 0, but for 1, or for 2, or for 3, or... In this way we have found infinitely many ways to describe the natural numbers.

work after all. We have not found a satisfying answer for describing mathematical objects and this remains a question that needs answering.

This gives us two options to continue. We could either (1) abandon this question altogether and thus give up our realist position, or (2) remain a realist but discard the belief that all mathematical objects are sets. Option (1) is unattractive, as the reasons to adhere to realism in the first place have not changed. Therefore it must be that we go for option (2) and decide that mathematical objects are not sets. This means that we have come back to our original question of what mathematical objects are.

According to Benacerraf, trying to find the correct set theoretic description of mathematical object indeed does not make sense. There is nothing that sets the number 3 apart from the other natural numbers, except that is is preceded by the numbers 2, 1 and possibly 0, and followed by the numbers 4, 5, 6, et cetera [1, p. 70]. There is no object that *is* the number 3, as it is nothing more than 'the object in the third place'. Any object could play the role of the number 3 [1, p. 70]. This brings us to a view called *structuralism*.

2.4 Structuralism

Structuralists believe that the main objective of mathematics is to study the structures of mathematical objects that are abstracted from the individual objects that make up those structures [30, p. 341]. For example, structuralists hold that arithmetic studies the structure of the natural numbers, Euclidean geometry the Euclidean-space structure et cetera. These structures are exemplified through systems of objects [33, p. 9]. We already discussed a first structuralist, Benacerraf [1], and others include Shapiro [33], Resnik [31] and Hellman [15].

With a *system*, we mean a collection of objects that contains certain relations among them [34, p. 259]. A *structure* is the abstract form of such a system, that highlights the interrelationships among the objects, and ignores any features that do not affect how they relate to other objects in the system [34, p. 259].[5]

An example of such a system is the set $\{0, 1, 2, 3, 4, \ldots\}$, exemplifying the natural number structure. However, the systems $\{1, 2, 3, 4, 5, \ldots\}$ and $\{0, 2, 4, 6, 8, \ldots\}$ both exemplify the same natural number structure as well. And so does an infinite sequence of dashes (for example the sequence $|, ||, |||, ||||, \ldots$) too. All these systems have a first element and a certain successor function; therefore

[5]This has certain similarities to the type-token dichotomy. A structure is much like a *type*, whereas a system is more like a *token*. A token is a concrete particular, such as each of the six 3's in this list: '$3, 3, 3, 3, 3, 3$'. These numbers exists on my computer screen and are composed of pixels. When I print this thesis, it will instead have six different tokens. The type of the number '3', however, will stay the same. It is the abstract and unique idea that creates the shape that all 3's share [34, p. 262]. In much the same way, a *system* of the natural numbers is a concrete particular that exemplifies the *structure* of the natural numbers - and a different system can do this as well.

they are instances of the abstract form of the natural numbers.

Therefore, mathematics is about the internal relations within a structure, according to the structuralist. Reck and Price give three *intuitive theses* to summarize structuralism: '(1) that mathematics is primarily concerned with "the investigation of structures"; (2) that this involves an "abstraction from the nature of individual objects"; or even, (3) that mathematical objects "have no more to them than can be expressed in terms of the basic relations of the structure" ' [30, p. 341].

An important notion for structuralists is that of an *isomorphism*. An isomorphism is a bijective mapping between systems that preserves the structure [21, p. 3]. Therefore, two systems that are isomorphic to each other are of the same structure. We can say that a structure is *unique up to isomorphism*. When we construct a natural number system, we can in fact construct infinitely many of these system – all being isomorphic to each other [30, p. 343].

2.5 Different views: (non-)eliminative structuralism

The description of structuralism that we have given in the previous section captures the intuitive ideas of structuralism, but remains vague and imprecise about the details. In fact, there are many different approaches to structuralism. In this thesis we will assume a position known as *non-eliminative structuralism*. But before we describe this exact position, we will first briefly discuss what *eliminative structuralism* entails.

What divides these positions is similar to the difference between mathematical realism (whose proponents claim that abstract entities exist independently) and anti-realism (whose proponents claim that they do not). At stake in the current issue is the the existence of the structures as abstract objects [29]. In eliminative forms of structuralism the structure itself is eliminated, resulting in what is also called *structuralism without structures*. (As to be expected, the alternative name for non-eliminative structuralism is *structuralism with structures*.)

There are two main forms of eliminative structuralism: we have the *relativist* and the *universalist* position. Relativist structuralists like Gleason [10] defend that an arithmetical statement is relative to the interpretation that is chosen. A structure is nothing but the systems that exemplify it. Therefore there does not exist anything like the structure of natural numbers apart from the (set-theoretic) system(s) that represent it [21, p. 9]. Universalist structuralists as Hellman [15] also believe that the natural number structure does not exist apart from its set-theoretic representations, but for different reasons. They claim that we should understand an arithmetical statement about the natural numbers as a quantified statement concerning all set-theoretic representations of these numbers [21, p. 10]. For this, modal logic turns out to be useful and this position is therefore also sometimes described as *modalized universalist structuralism* or *modal structuralism* [30, p. 359].

In other words, a statement is true if it is true for all *logically possible* systems [34, p. 274].

Both these positions hold that although systems exist, structures do not. Non-eliminative structuralists plead for a different view: they propose that structures exist and that they do so independently.[6] Structures are seen to stand apart from the community of mathematicians and scientists and exist independently of their minds, languages, et cetera [35, p. 130]. We can speak of structures *sui generis*: meaning that structures are seen as objects that cannot be reduced to any kind of set theoretical construction, but are entities in their own right [21, p. 12]. Non-eliminative structuralism is therefore the position that takes patterns to exist independent of any systems that exemplify it [34, p. 263]. Apart from the existence of structures themselves, non-eliminative structuralism also commits to the existence of the places within a structure [25, p. 156]. This makes it a realist doctrine. Non-eliminative structuralism is sometimes also called *ante rem structuralism*.[7]

The realist needs to provide an answer to the question what mathematical objects are. Non-eliminative structuralism is attractive, as it provides such an answer: mathematical objects are structures that can be exemplified through different systems. Additionally, structuralism is not susceptible to Benacerraf's objection. We can now use both the Zermelo and the Von Neumann ordinals as systems to describe the natural numbers, but we are not obliged to say that either *is* the natural numbers.

Let us finally take one example. What is the number 3? The non-eliminative structuralist answer is that it is nothing but a certain place in the natural number structure. Any number is only a place in the pattern that is shared between all models of arithmetic - whether this model is described in a set-theoretical hierarchy or a different one. We have thus found the ontological position that we will take. However, this does not mean that our study of mathematical objects is finished. The next chapter will introduce a particular epistemological question that needs resolving as well.

[6]This indeed reminds us of the realist position that we described in section 1.1.

[7]The term *ante rem structuralism* is especially used for denoting Shapiro's version of structuralism as described in for example [33, 34].

3 The access problem

In the previous chapter we discussed multiple positions in the philosophy of mathematics. Up to now we have focused on the ontological aspects of these positions. In this chapter we move on and take a look at some of the epistemological questions involved, as we generally believe that at least some of us know some things about these mathematical entities.[1] Surely some mathematicians are also *justified* in believing that they have knowledge about these objects. But how is such justified knowledge of mathematical objects possible?

For non-realist positions this question has an easy answer.[2] However, for all realist positions, including non-eliminative structuralism, this question needs more attention. If abstract mathematical entities exist independent of the mathematician, it is not at all clear how we know things about them. We need to find out how we get access to these abstract objects. In this chapter we study the corresponding epistemological problem called the *access problem*.

In our investigation we are not so much interested in this problem for any realist position, but will base our investigation on assuming a non-eliminative structuralist position. We have seen that non-eliminative structuralism holds that structures exist independently of any systems that exemplify them. As mathematics is taken to be the study of these structures, our mathematical knowledge is nothing but the knowledge we have of these structures [34, p. 276]. The access problem for non-eliminative structuralism does therefore only call for a theory of how we acquire knowledge of structures. One suggestion is that this is done via a process called *pattern recognition*, which we will examine in more detail later.

3.1 The access problem

The access problem originates from the article *Mathematical Truth* by Benacerraf [2]. It is known by multiple names and has also been described as the 'access problem', the 'Benacerraf problem' and the 'reliability challenge' [4].

[1]As for our present purposes we will set skepticism aside.

[2]This is for example the case for the nominalist, who claims that mathematical objects do not exist at all. Certainly knowledge of non-existing objects is not possible, so this epistemological question does not need much attention.

The problem applies to the realist position, that take mathematical objects to exist independently. We will elaborate on this topic in the next section (section 3.2). In the previous chapter we have concluded that these mathematical objects differ from physical, everyday objects because they are *abstract*, so *nonspatially and nontemporally localized and causally inefficacious*. Mathematicians attempt to study these abstract and infinite mathematical entities. However, mathematicians themselves are finite beings (as all humans), with senses only suited to obtain information about their concrete surroundings [24, p. 309]. These (independently existing) mathematical objects are thus located in a realm that mathematicians cannot travel to and are not able to receive signals from [35, p. 131]. How is knowledge of mathematical objects possible, then?

This is what we call the *access problem*. It appears to be a mystery how the mathematician can obtain knowledge of the mathematical domain. In order to do this it seems we have to find a way to cross from the concrete realm to the abstract domain of mathematics [24, p. 309]. Moving from the concrete to the abstract appears to be impossible [35, p. 131].

3.2 Problematic positions

As we mentioned in the previous section, the access problem applies not to all positions in philosophy of mathematics. The realist needs to provide an answer to this problem, as it needs to explain how mathematicians are able to form reliable beliefs about the mathematical objects they claim to exist [35, p. 131]. Nominalists do not have to do so (as it is impossible to have knowledge of objects that do not exist), neither do intuitionists (mathematical objects are clearly accessible to the 'knower' when they depend on him).

In this thesis, we assume a non-eliminative structuralist position and accordingly mathematical knowledge is taken to be knowledge of structures. Take for example arithmetic: having arithmetical knowledge now implies must have knowledge of the relevant arithmetical structures [35, p. 134]. Because the non-eliminative structuralist takes structures to exist *sui generis*, it has to explain how reliable beliefs about them are possible. The later sections of this chapter will be devoted to an attempt to achieve this.

However, the solution that we will present depends on our assumption of non-eliminative structuralism. For a different position we might need a different solution. We outlined possible other structuralist positions in section 2.5, discussing relativist and universalist structuralism. Of these two eliminative structuralist position, we will firstly examine the former.

As we specified in section 2.5, the relativist structuralist believes that there exist no structures apart from the systems that exemplify it. Relativist structuralism is therefore not a realist position (contrary to non-eliminative structuralism): it is a structuralist position that denies the existence of abstract structures. As mathematical knowledge is taken to be knowledge of systems and thus of concrete entities, it follows that there is no access problem. According to the relativist structuralist, there is no need to access abstract entities in order to obtain mathematical knowledge.[3]

For the universalist structuralist this is a different story, and the access problem does turn out problematic for this position. By claiming that mathematical knowledge is knowledge of all *logically possible* systems (see section 2.5), this position demands a justification of how reliable beliefs about these systems can be formed [34, p. 273]. The problem for the universalist structuralist turns out to be related to the problem of the non-eliminative structuralist. Because the non-eliminative structuralist says that we can obtain knowledge of abstract structures, it has to provide an account of how this is possible. The universalist structuralist, on the other hand, has to show which systems are possible and how we can obtain knowledge of what holds in these different possible systems [34, p. 276]. So for non-eliminative structuralism, knowledge can be obtained by demonstrating the existence of some structure, while for universalist structuralism this can be done by demonstrating the existence of a corresponding logically possible system. This links the two positions. Shapiro points out that there are 'there are direct, straightforward translations from the realist's language to the language of the modal structuralist and vice versa'[4] and provides these in [33, p. 229]. Using this translation we can make slight modifications to make the solution presented in the following paragraph (as a solution for the access problem for non-eliminative structuralism) also suitable for universalist structuralism [34, p. 276].

3.3 Solutions to the access problem

The access problem can be interpreted in various ways. One possible way is articulated by MacBride [25]. He believes that the epistemology that is provided as a solution to the access problem must describe the process in which mathematical knowledge is obtained in purely non-mathematical terms. From this it must be clear that the 'knowers' indeed end up with mathematical knowledge. This view is quite extreme, as it is unclear whether a non-mathematical theory could be more secure than mathematics itself [35, p. 133].

A view on the other side of the spectrum is proposed by philosophical naturalists as for example Maddy [27]. These claim that there is no access problem at all, because according to the naturalist, we can only criticize mathematics in mathematical terms [35, p. 133]. They believe that it is impossible to account for mathematical practice in non-mathematical terms and such an epistemology will therefore yield nothing [35, p. 133]. No such epistemological account can make us more secure than we are already, so this position thus assumes that we obviously have

[3]Although the relativist structuralist escapes the access problem, this position does face different problems; see also [30, p. 348-354].

[4]Recall that universalist structuralism using modal logic is also called modal structuralism, as described in section 2.5.

mathematical knowledge.

Both these approaches do not seem satisfying, as the first demands something impossible and the second assumes what is questioned. We seek for a solution to the access problem that is in between these two views. We desire an epistemology that is neither unreachable nor obvious. One of the strategies that has been proposed as a process that could result in justified mathematical knowledge is pattern recognition.

3.4 Pattern recognition

One of the proposed solutions to the access problem is that of *pattern recognition* and main proponents are Shapiro [33] and Resnik [31]. This paragraph is mostly based on the version of pattern recognition as proposed by Shapiro in [33, 34]. The core idea is that a subject comes into contact with multiple systems of concrete objects that are organised differently, and from this she extracts the corresponding pattern (or: the structure) [35, p. 136]. Pattern recognition in this sense is only possible for small finite structures and not for infinite structures, as it is impossible to oversee an infinite structure. So how does the subject learn infinite structures, then? The global idea is the following. The subject begins to mentally arrange the first few structures using pattern recognition. Then she realizes that they themselves contain a pattern. The pattern that these patterns contain, can be projected to continue the sequence. This leads to knowledge of larger finite structures and eventually infinite structures [35, p. 136]. At this stage, it is recognized how new structures can be formed from old ones. However, this limits the subject to learning new structures that are at most countably infinite. Knowledge of even larger mathematical structures can be obtained through a final stage where the subject learns to provide axiomatic characterizations of structures, as for example the Peano axioms are for arithmetic [34, p. 285].

This last step includes obtaining the ability to describe structures with a size larger than countably infinite. This mechanism relies on implicit definitions and goes further than the process of pattern recognition (a detailed account can be found in [34, p. 283-289] and [33, p. 129-136]). For our purposes it will suffice to speak only of structures that are at most countably infinite in size, our prime example being the natural number structure. We will therefore not describe the techniques behind implicit definition nor go into the details of this last stage. Instead we focus on just the process of pattern recognition, as the process up to establishing the natural number structure.

Two steps can be identified in this process of pattern recognition. The first is the learning of finite structures, via the recognition of a pattern as exemplified in different systems. The second step is the recognition of a higher-order structure in the formerly identified patterns itself. This crucial step requires a shift between what is seen as object and what is seen as structure [34, p. 280].

Let us take a closer look at this first step. How does repeated exposure to a system lead to knowledge of its structure? This process is similar to learning the alphabet [34, p. 276]. Suppose that the letter 'T' is pointed out to a subject, combined with the corresponding sound. The subject might at first think that this sound belongs to the specific token of the letter T, but as more and more instances of this letter are pointed out she will realize that it actually belongs to the specific shape of the letter - the crossing of a vertical and a (smaller) horizontal line. Even later she will discover that there are even more shapes that should be identified with this sound (capital T's, lower case t's...). Then she will realize that the sound applies to the type whose tokens include all capital and lower case tokens of T's. That is, that 'being a T' is just a place in a pattern: what all T's have in common, is their role in the alphabet and different strings [34, p. 277].

In a similar way, the 4-pattern is exemplified by different collections of four objects. By coming into contact with different collections (systems) of four objects, the subject will see what pattern is shared between these systems [34, p. 277-278]. We will call this pattern the structure $\$_4$, and denote any system exemplifying this structure as S_4. Following this process, small cardinal structures as the the 1-pattern $\$_1$, 2-pattern $\$_2$ et cetera are learned.

We can imagine that our subject now has a mental arrangement of the 1-pattern, the 2-pattern, the 3-pattern and the 4-pattern. That is, it knows the following sequence of structures:

$$\$_1,\ \$_2,\ \$_3,\ \$_4.$$

We have come to the second step in our process of pattern recognition. Reflecting on the sequence above, the subject realizes that this sequence of structures provides a new structure itself [34, p. 279]. These four structures together, seen as system, exemplify the 4-pattern themselves! The sequence above, consisting of structures, is thus on itself an instance of $\$_4$.

This step invokes a faculty of projection [35, p. 136]. After having realized that these finite structures together can be seen as a system exemplifying a *new* structure, the subject realizes that each element of this sequence seems to have a unique element succeeding it. So, as the elements are in fact structures, it realizes that every structure has a unique structure succeeding it.

The subject thus learns to extend any structure. In this way, the subject can gain knowledge of very large structures that it has never encountered as concrete systems [35, p. 136]. Continuing this process will eventually result in the pattern of an infinite structure $\$_\infty$ [34, p. 280].

Our subject can now create a mental arrangement of a sequence of structures that also contains infinite structures. It is now a sequence that does not end:

$$\$_1,\ \$_2,\ \$_3,\ \$_4,\ \dots .$$

This process provides a possible way of learning structures that are at most countably infinite in size, such as the natural numbers. These can be represented by the sequence that we have just given, or by many different sequences [34, p. 281]. We could therefore equally say that the mental arrangement of the infinite sequence of structures above looks like

$$a,\ aa,\ aaa,\ aaaa,\ \ldots,$$

because the structure of the natural numbers is independent of the elements constituting it. The crucial point is learning what is meant with the ellipses in the end (. . .), that allow the subject to coherently discuss infinite structures [34, p. 281].

Now that we have outlined all steps in the process of pattern recognition, where does that bring us? We assumed a non-eliminative structuralist position and therefore needed to justify the possibility of knowledge of structures *sui generis*. The process of pattern recognition provides this justification. It makes it plausible that mathematical knowledge is nothing but knowledge of structures *sui generis*, as it describes a process that allows us to form reliable beliefs about abstract mathematical structures. This makes pattern recognition a solution to the access problem.

3.5 (A response to) MacBride's criticism

Although we have now found that pattern recognition is a solution for the access problem, its suitability is not undisputed. Fraser MacBride presents his criticism in the article 'Can *ante rem* structuralism solve the access problem?' [25].[5] This article is particularly aimed at the ideas of Shapiro as presented in [33] and [34], which also formed the basis for the previous paragraph. Shapiro reacts to MacBride's criticism in his paper 'Epistemology of Mathematics: What are the questions? What count as answers?' [35]. In this paragraph we will give a short overview of the debate and argue for the suitability of pattern recognition as a solution to the access problem.

According to MacBride the problems do not lie in the first step of pattern recognition. He writes that 'for present purposes I shall grant that we have a capacity for abstracting simple linguistic and numerical patterns from the systems that exhibit them' [25, p. 158]. MacBride is prepared to consent with Shapiro, so this is not the problematic stage. However, we should note that he also believes that 'an illuminating philosophical description of the processes involved remains to be given' [25, p. 158-159].

A second point on which MacBride agrees with Shapiro is in acknowledging that 'we *do*

[5] We should recall from section 2.5 that non-eliminative structuralism is also denoted as *ante rem* structuralism, especially when speaking about Shapiro's position.

project from the structure of the systems of objects with which we are acquainted, and thereby come to grasp structures the instantiating systems of which we cannot literally survey' [25, p. 159]. However, MacBride argues that this descriptive claim alone does not suffice. The problem is the following:

> [T]here are indefinitely many ways of projecting the series of numerals – indefinitely many rules for constructing numerals which are consistent with the evidence supplied by an initial finite sequence arrived at by abstraction, rules which fail to generate systems of numerals isomorphic to the natural numbers. [25, p. 161]

Even if the subject knows (1) how to continue some particular cardinality structures to create a successor distinct from it, then where does she learn that (2) *every* structure has a unique next-longest extension? So, suppose that the subject has a mental arrangement of a system of structures in the sense that we discussed in section 3.4. We will represent this system as $\$_1$, $\$_2$, $\$_3$, $\$_4$, ... Although the subject might know that in this sequence $\$_4$ succeeds $\$_3$ (and have similar knowledge of other finite structures), MacBride challenges that the subject would know that $\$_{n+1}$ succeeds $\$_n$ for all n.

According to MacBride, we can only *deduct* claim (2) from claim (1) by appealing to truths that are just as general in character as (2) [25, p. 160].[6] MacBride argues that 'without this knowledge [the subject] can have no assurance that the pattern of cardinality patterns extends not only *just* beyond but also *far* beyond the cardinality structures of which [she] has encountered instances' [25, p. 160].

Shapiro, in his reaction to MacBride, accepts the difference between describing the source of a given faculty and justifying the reliability of that faculty [35, p. 139]. However, he opposes to MacBride's idea that we should *deduct* (2) from (1). Instead, he points out the gradual borders between premises, postulates, working hypotheses, established beliefs and known propositions. What once is regarded only a hypothesis could later become an established belief. We should see claim (2) in this light. A subject, noticing that small finite cardinal structures have successors, might hypothesize that this applies to all structures and subsequently assert (2) as a hypothesis worth exploring. Later, when this hypothesis proved successful for various purposes, she might see that (2) has taken up a more central role – and might possibly even acquire the status of an axiom [35, p. 140]. Exploring the consequences of hypothesising (2) will never enable us to *prove* it, but it will allow us to recognize its role in our intellectual enterprise [35, p. 142].

[6] MacBride does concede (for the sake of the argument) that there are at least some people who in fact *do* move from (1) to (2). However, he holds that this still does not explain why it is rational that some people do so; it merely presupposes so [25, p. 160].

With this reaction, I believe Shapiro provides a fitting response against MacBride's criticism. Although it is indeed the case (as MacBride points out correctly) that the subject's interpretation decides what structure is abstracted, I will argue that this is not problematic in any way.

We should keep in mind that the purpose of the stage of 'projection' is to show that there exist (cardinality) structures that go beyond the ones we encountered as concrete systems. The natural number structure is one such example. Suppose that a subject, having access to small finite structures, adopts a 'non-standard' way of continuing an initial sequence. This does not meant that our goal can not be achieved: we should only admit that there exists a certain structure that can be accessed in this way this subject does. But now suppose that a different subject, that again has access to small finite structures, does interpret the rule of finding a successor in the usual way. Would this subject then have a special problem? Even if the subject's interpretation were based on a coincidence, the goal of showing the existence of structures beyond the ones we encountered as concrete systems would still be achieved. In fact, this indeed shows the existence of the natural number structure.[7]

Although this argument on itself might invite new criticism, that is not a discussion that we want to have here. The purpose of this section was to provide some insight into the debate on the suitability of pattern recognition as a solution to the access problem, as we presented in section 3.4. We have seen relevant criticism, but nonetheless believe that the arguments in favour of this solution are convincing enough to make pattern recognition an acceptable assumption for the upcoming chapter. There we will present a first attempt at formalizing pattern recognition – and therefore respond to MacBride's request for 'an illuminating philosophical description of the processes involved' [25, p. 158-159].

[7]In the debate, MacBride and Shapiro also discuss an additional argument against the suitability of this solution to the access problem. At stake here is the mechanism for moving from knowledge of countably infinite structures to knowledge of even larger structures. As we decide not to elaborate on this in this thesis, we consider it not relevant for our purposes and will not cover it here.

4 A first step in the AI implementation of pattern recognition

In the previous chapters we have discussed non-eliminative structuralism and its ontological and epistemological views regarding mathematical objects. Assuming non-eliminative structuralism and the belief that at least some humans have some mathematical knowledge, it must follow that at least some humans have some knowledge of structures. But knowledge of abstract structures is not unproblematic, as we witnessed when we spoke of the access problem and a corresponding solution based on the process of pattern recognition. In this chapter we assume that pattern recognition is indeed a working solution that enables us to gain knowledge of abstract structures.

If pattern recognition is a process that enables us, humans, to learn about the abstract mathematical objects that we call structures, we can wonder if this same mechanism is applicable to an artificial learner (an 'AI') as well. Mimicking human learning in an agent using artificial intelligence will allow us to discuss whether it is possible for such an AI to learn to recognize mathematical structures. In this chapter we will explore what a first step in this learning process for an AI would look like and what criteria it should adhere. The formalization that this implementation requires provides us insights about pattern recognition in human learning as well.

The current available explanations of the processes involved in pattern recognition can be made more rigorous. We saw in section 3.5 that MacBride argues that 'an illuminating philosophical description of the processes involved [in pattern matching] remains to be given' [25, p. 158-159]. By exploring the possibilities of an AI implementation of the process, we respond to his request: the model explicating our intuitions does not allow any vagueness, so we will have to make all steps formally precise.

Our ultimate aim, then, is to provide the AI with knowledge of infinite structures, employing a method similar to pattern recognition as described for humans in section 3.4. Just as in this section, we will focus on the prime example of the natural numbers and their corresponding (infinite) structure. This is a complicated operation and no earlier attempt can be found in the literature, so we will not be able to discuss the complete process here. As described in the previous chapter, pattern recognition comprises of several sub-processes. In this chapter we will confine ourselves to the necessary first step that needs to be worked out for any attempt to be successful.

4.1 The goal: approaching human pattern matching

It is important to begin with noting that it is an obvious requirement that the AI-based approach towards pattern matching resembles the human-based approach, as we want to closely compare the two and analyze their respective differences in outcome. The reason we want to implement pattern recognition for an AI is because this will help us to point out the strengths and weaknesses of pattern recognition as a solution to the access problem. The formalization that an AI implementation requires points out all deficiencies and does not allow any vagueness. Furthermore a successful implementation will result in a model that can be used to test hypotheses regarding the particular epistemic position that was intended; an example of such a test will be given in section 4.6.

Keeping the process similar to the human-based approach means for example that the duality of the process as described in paragraph 3.4 should be preserved. The artificial learner should begin by learning to recognize finite structures, just as the human learner does. A second step then includes recognizing that these structures together can be seen as a system that itself can be expanded to become the natural number structure.

Any process where an AI learns to recognize the natural number structure therefore begins with learning to recognize finite structures. This essential step cannot be omitted in any successful implementation and therefore it will be the subject of the present chapter. We will call this sub-process *finite-structure recognition*.

4.2 AI learning

Learning processes for AI are united under the term ‘machine learning’, which includes a broad range of different techniques. Examples of such algorithms are linear or logistic regression, naive Bayes and Gradient Boosting algorithms [12, p. 3]. We will not go into the details of one specific technique, but instead use high-level descriptions. A machine learning algorithm is an algorithm that is not explicitly programmed for its task, but instead uses input data to train itself. Extracting patterns from this raw data allows the AI to learn [12, p. 2]. To specify what we mean with ‘learning’ when speaking about an AI, we will use the following definition:

> A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E. [12, p. 97]

This definition uses three variables that should be specified: T, E and P. We will now investigate some particular details of these variables with regard to our purposes.

4.2.1 Task T

The task T intuitively has something to do with recognizing finite structures. But how can we formalize this? We can see our task as a *binary classification task*: meaning the AI is asked which of 2 categories the input belongs to [12, p. 98]. For us, this would mean that the AI has to decide whether a certain input system S is of some structure \$ or not. The task T thus consists of providing a yes/no answer to the question: 'Does system S instantiate structure \$?'. We can formulate this more precisely using predicate logic. Let $I(S, \$)$ be true if system S instantiates structure \$. The task T is then to determine whether $I(S, \$)$ is true or false.

Although we will limit to finite-structure recognition here, this formulation of our task enables us to speak about our ultimate goal (which is learning the AI to recognize the natural number structure) as well. Let us denote the 1-pattern as $\$_1$, the 2-pattern as $\$_2$ et cetera. In section 3.4 we spoke of an infinite system consisting of structures, denoted as $\$_1,\ \$_2,\ \$_3,\ \$_4, \ldots$. Following the process of pattern recognition as discussed in that section, we say that a subject has learned the natural number structure when she recognizes that this system (extended infinitely) exemplifies the natural number structure. Simulating this in the task for our AI, we can say that the AI has learned the natural number structure when it evaluates $I([\$_1,\ \$_2,\ \$_3,\ \$_4, \ldots], \$_{\mathbb{N}})$ correctly as being true.

4.2.2 Performance measure P

We now move on to describing an appropriate performance measure P. At first sight, *accuracy* might seem a good choice [12, p. 102]. Accuracy is the proportion of correct answers, calculated as the number of correct answers divided by the number of wrong answers:

$$\text{Accuracy} = \frac{\text{correct answers}}{\text{total answers}}.$$

However, a correct answer can be either a true positive ('TP') or a true negative ('TN') and similarly, wrong answers are either false positives ('FP') or false negatives ('FN'). We should take into account that the accuracy does not provide any information about which prevail in our case [12, p. 102]. Although it shows the proportion of mistakes that the algorithm makes, it does not indicate of what kind these mistakes are. We can see this by specifying the definition of accuracy in the following way:

$$\text{Accuracy} = \frac{\text{correct answers}}{\text{total answers}} = \frac{TP + TN}{TP + FP + TN + FN}.$$

This makes it important also to report on the *precision* and *recall*. Precision gives the proportion of correct positive identifications: a model that produces no false positives has a precision of 1. It is calculated by:

$$\text{Precision} = \frac{TP}{TP + FP}.$$

Recall, on the other hand, provides the proportion of all actual positives that are also identified: a model without false negatives has a recall of 1. This is calculated by:

$$\text{Recall} = \frac{TP}{TP + FN}.$$

Together, these three performance measures can accurately describe the performance of the task. An AI that is accurately trained will ideally have a high accuracy, consisting of a high precision as well as a high recall.

4.2.3 Experience E

Lastly we specify the experience E fitting for our task T. The performance of the AI in determining whether $I(S, \$)$ is true or false for any system S and structure \$, as measured by P, should increase with experience E. This means that the number of false positives and false negatives should decrease (or the number of true positives and true negatives should increase). This is possible by training the AI via supervised learning: providing the AI input data and its desired output. The AI can train itself on this data and subsequently determine if it calculated the correct output.

What should this training data look like? A triple seems suitable, as we know that the experience should consist of three parts. The first is the system S that has to be evaluated. The second is the structure \$ that the system might or might not exemplify. The third is the correct yes/no answer A to the question ‘Does system S instantiate structure \$?’.[1] This makes E the triple $(S, \$, A)$.[2]

Defining experience E as the triple $(S, \$, A)$ gives us some specification, but not enough. We should also go into the details of how the different components of E are represented themselves. For A, this might not be too difficult: as it is the truth value of $I(S, \$)$, the obvious choice for A is to be a Boolean. For S and \$, however, it is more difficult to choose an appropriate representation. We will use the following paragraphs to discuss this more elaborately.

[1]Or: the truth value of the predicate $I(S, \$)$.

[2]The experience E might also be a data set consisting of a collection of these triples instead of just a sole instance, but without loss of generality we can speak of E as if it is just a single instance.

4.3 Representing a system S

There are many systems that can exemplify a particular structure. Let us say, as we did in section 3.4, that the 4-pattern $\$_4$ is the structure common to all collections of four objects. A group of four physical objects exemplifies this structure; but so do the letters in the word 'fork', the sound of a doorbell ringing four times, a picture of 4 banana's, all prime numbers under 10 and the set $\{2,3,5,7\}$.

The most basic representation of systems in our AI implementation might be a representation using sets: then the sets $\{2,3,5,7\}$ and $\{99,98,97,96\}$ instantiate the 4-pattern, but $\{1,3,5\}$ or $\{50,100,150,200,250\}$ does not. However, training on just these representations might lead an AI to believe (mistakenly) that the set $\{a,b,c,d\}$ does not exemplify $\$_4$ either. As opposed to the earlier sets, this is a set consisting of letters and not numbers and the AI has not learned that $\$_4$ can also be instantiated by some other set than a set consisting of numbers. It is therefore important that the AI should become aware that there is more to 'exemplifying $\$_4$' than there is to 'being a set containing four elements which are all numbers'.

However, not all the systems exemplifying the 4-pattern that we named in the first paragraph of this section are equally suitable for a representation in AI learning. Luckily, it is also not strictly necessary to be able to represent all possible different sort of systems: we should just present *enough* different sort of systems that the AI learns that instantiating $\$_4$ is not limited to a specific input style. We can justify this idea by looking at the human based process of pattern matching, where the (human) learner does not come into contact with all possible systems that instantiate the 4-pattern either. It might also be very well possible that this learner has never encountered a certain specific system (she might for example never have seen a four coloured object), but she could still be able to form ideas about what it means to exemplify $\$_4$. Therefore, the AI should just encounter *enough* different sorts of systems to be able to create an idea of what it is to exemplify a structure.

Now we will describe some possible representations of systems. We already named the use of sets. But an AI could also be able to analyse an image containing a certain number of colors or dots, a sound tape of a doorbell or a picture of some objects. Furthermore it could look at strings, the number of times a key on the keyboard is pressed, et cetera. There are endless possibilities to expand this idea.

What we should note is that all these systems are direct representations, in the sense that the representation is a concrete instance. This works for finite systems; but not for infinite systems. This is an important observation, as this is one of the reasons why the process of pattern matching was introduced in the first place! Shapiro notes that we speak about large structures as the 12.444-

pattern $\$_{12.444}$ just as easily as we speak about the 4-pattern $\$_4$, even though no one has ever observed a system exemplifying this structure [34, p. 279]. He writes that the reason that we can speak of these large structures coherently is because pattern recognition enables us to. Therefore an AI implementation will not have to be able to represent extremely large or infinite systems.

4.4 Representing a structure $\$$: unlabeled graphs

Whereas in the case of systems we can restrain to looking at finite systems, this is not possible for structures. We know many infinite structures and for our purposes we will need to be able to talk coherently about them. Take our usual example of the natural number structure, which is an important infinite structure. We should therefore take into account that our representation of a structure $ should also be appropriate for representing infinite structures. Furthermore we are assuming a non-eliminative structuralist position throughout this chapter, taking structures to exist independently. We should find an account that does justice to this view as well. To find such a suitable representation we will explore the possibilities provided by Leitgeb's unlabeled graph theory, as described in [21, 22].

In his articles, Leitgeb provides a case-study of non-eliminative structuralism about unlabeled graphs. He gives an axiomatic account of unlabeled graphs as structures *sui generis*. In the rest of this chapter we will assume that the ideas that Leitgeb presents in his papers provide a proper way to talk about structures – in other words, we assume that it is appropriate for our purposes.

What unlabeled graphs are, can intuitively be understood by looking at figure 4.1. This figure shows all possible graphs with 4 points.

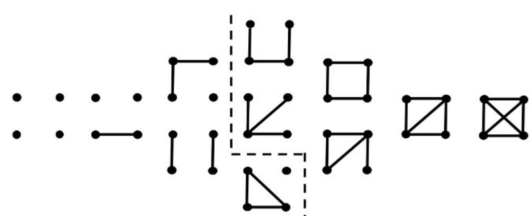

Figure 4.1: The unlabeled graphs with 4 points [21, p. 15].

Unlabeled graphs are useful because they resemble structures in a sense that two nodes within one or different graphs are only said to differ for as much that they have a different *position* or *role* in the graph. When these nodes become separated from their respective graphs, they look exactly the same [21, p. 16].

There are multiple ways to speak about unlabeled graphs coherently and one (popular) possibility is to do this using set theory.[3] However, Leitgeb takes a different approach and instead introduces an axiomatic account called *Unlabeled Graph Theory* (UGT). This axiomatic approach is closer to the informal idea of unlabeled graphs and allows us to understand graphs as primitive graph theoretic structures *sui generis* [21, p. 19]. We will only discuss a small part of this theory here, but more details can be found in Leitgeb's articles [21, 22].

The language to formulate the axioms of UGT is second-order logic and will mainly consist of three predicates:

$$\begin{aligned}
Graph(G) &= \text{the value of the variable } G \text{ is an unlabeled graph,}\\
Vertex(v,G) &= \text{the value of variable } v \text{ is a vertex (or: node) that is a component of the value}\\
&\quad\ \text{of the variable } G,\\
Connected(v,w,G) &= \text{there is an edge between the values of the variables } v \text{ and } w \text{ within the graph}\\
&\quad\ \text{that is the value of the variable } G \text{ [21, p. 20].}
\end{aligned}$$

Furthermore Leitgeb provides axioms for adding isolated vertices and adding edges, allowing to build new graphs from existing ones. We can prove that all finite unlabeled graphs exist, and, very relevant for our purposes, that there exists an unlabeled infinite graph $G\infty$. In order to do this, let $Isomorphism(f,G,G-v_0)$ be a metalinguistic abbreviation for there being an isomorphism f between G and what remains from G once v_0 is ignored; a formal definition can be found in [21, p. 25]. Then $G\infty$ is postulated by the axiom:

$$\begin{aligned}
&\mathbf{E}\infty:\\
&\quad \exists G\ \exists v_0,v_1 : Vertex(v_0,G) \land Vertex(v_1,G) \land Connected(v_0,v_1,G) \land\\
&\qquad \forall w(Connected(w,v_0,G) \to w = v_1) \land\\
&\qquad \exists f(Isomorphism(f,G,G-v_0) \land f(v_0) = v_1).
\end{aligned}$$

The graph $G\infty$ must look as depicted in figure 4.2.

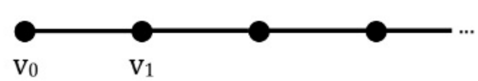

Figure 4.2: The graph $G\infty$ [21, p. 33].

[3]This follows the idea that all parts of mathematics can be reduced to set theory as presented in section 2.2.

This graph can be distinguished from any finite graph as it is a proper subgraph of itself: there exists an isomorphism between $G\infty$ and the class of nodes from $G\infty$ that have more than one neighbour.[4]

UGT as an axiomatic account of unlabeled graphs as structures *sui generis* provides us with the tools suitable for our purposes. We were in need of a way to represent a structures that could be finite or infinite: UGT provides this possibility. The axiomatic account allows us to describe a structure \$ using second-order predicate logic. Take any n-pattern as example. Now $\$_1$ can be described by a single node. By extending this with one extra node and adding an edge between those, we can represent $\$_2$. Then $\$_3$ is in turn created by adding a new node again, and connecting those last two nodes. This way, the structure that is the n-pattern $\$_n$ is created from the $(n-1)$-pattern $\$_{n-1}$. Continuing this infinitely brings us to the complete structure of the natural numbers, $\$_{\mathbb{N}}$, which is represented by $G\infty$.

4.5 Does learning $I(S, \$)$ provide knowledge?

With all representations in place, we can return to the bigger picture of finite-structure recognition. We assumed non-eliminative structuralism and described criteria for this necessary first step. Out task was to provide a yes/no answer to the question 'Does system S instantiate structure \$?' by determining the truth value of the predicate $I(S, \$)$. Performance (measured by accuracy, precision and recall) is to increase by training on triples $(S, \$, A)$. Here, the system S can be represented by various data types as long as this is done by providing concrete instances. We assumed Leitgeb's UGT and used its axiomatic description of unlabeled graphs to represent the structure \$.

In this way the AI can learn to recognize finite structures: it will be able to say whether any given system exemplifies any given structure. It does this by determining the truth value of $I(S, \$)$ for systems and structures that it has not encountered before.

We set out to describe the first step of the process of pattern matching and aimed to provide an AI with knowledge of finite structures. Suppose that we succeed in implementing what is described in this chapter and are able to train the AI. How reasonable is it then that we have achieved our goal? That is, can we say now that the AI indeed has *knowledge* of finite structures?

Although interesting and important, we cannot elaborate on this topic too much here. Instead we will only briefly discuss the topic to indicate that it is not completely unreasonable to assume that this AI implementation yields knowledge. For this, we will use an account that describes knowledge as *Justified True Belief* (JTB); although most philosophers agree that this view is not a

[4] A more formal definition of 'proper subgraph' can be found in [21, p. 31].

sustainable position, we can see it as a useful starting point [19].[5]

For now, however, we will assume that this account of knowledge as JTB is useful – at least for a first examination of the subject. Furthermore we assume that we can ascribe knowledge to an AI in a similar way that we ascribe it to a human subject. Our analysis will consist of three parts, as we have to evaluate all three aspects of JTB. We will start with the latter, *belief*: although it is debatable whether we can ascribe belief to an AI, we will here interpret the fact that the AI gave this particular output as if it is indeed a belief. Whether this is a *true* belief or not can be easily checked – but assuming that the training functioned properly (meaning that the training set was large enough and this resulted in a high score on all performance measures), we will assume that the AI performs good enough to be usually right in its judgements. We finally look at the *justification*. The AI was trained to deal with similar situations, so again assuming that its training functioned properly we should be able to say that the AI is justified in its belief. All in all, this shows that it is not completely unreasonable to believe that the process described in this chapter would be able to provide an AI knowledge of finite structures.

4.6 A test for non-eliminative structuralism: $I(\$, \$)$

Up to now, have seen many instances of systems exemplifying a certain structure. But as we have assumed a non-eliminative structuralist position, there is one case particularly interesting. Recall from section 2.5 that non-eliminative structuralists believe that structures exist and do so independently. This makes structures also objects that can be looked at in the same way that we usually look at systems. Therefore structures themselves can also exemplify a structure.

Take the natural numbers and look at them from our non-eliminative structuralist point of view: these numbers themselves exemplify the natural number structure [34, p. 269]. There might not even be a better example to illustrate the natural number structure than these numbers themselves. The natural number structure, seen as a system, thus exemplifies the natural number structure. The places of the natural number structure as a system are organised as follows: the first element takes the role of being the initial element, the second element takes the two-role, the third the three-role, et cetera. This results in the natural number structure.

In much the same way, *all* structures, seen as systems of places, exemplify their own structure. At least, as long as we endorse *non-eliminative* structuralism. Assuming this position makes it sensible to ask the question 'Does structure \$ (seen as system) instantiate its own structure \$?'. The non-eliminative structuralist then answers this question positively.[6]

[5]Examples of attacks on this 'JTB' analysis can be found in Hazlett [14] (attacking the truth-condition) and Fantl McGrath [6] and Radford [28] (both attacking the belief condition). Further criticism is articulated via *Gettier scenarios*, that provide examples of cases where justified true belief does not result in knowledge [9].

[6]Shapiro presents two different perspectives to look at structures. The first is the *places-are-offices* perspective, the

This observation makes it interesting to look at the truth value of $I(\$, \$)$ in our AI implementation. According to our non-eliminative position this proposition would have to be true. If we agree that the process described in the previous chapter provides the AI with knowledge of finite structures, then the truth value of $I(\$, \$)$ presents interesting insight into the nature of this knowledge. If we do not train the AI on any triple of the form $(\$, \$, A)$ (with A either true or false), this gives us a useful test in determining whether our AI indeed follows our non-eliminative approach. The test therefore allows us to check if we are working in the right (non-eliminative structuralist) direction before implementing the next steps in the process of pattern recognition.

second the *places-are-objects* perspective. He uses the switch between these perspectives to explain why for a non-eliminative structuralist a structure, seen as a system of places, always exemplifies its own structure [34, p. 268-270].

5 Conclusion

In this thesis we encountered multiple positions regarding the ontological status of mathematical objects, beginning with realism, intuitionism and nominalism. After discussing a (prevalent) set theoretic interpretation of realism and its problematic implications, we argued that the correct approach would be structuralist. Structuralists hold that mathematical objects are not defined by their internal nature but by the relations they bear to each other. These relations form abstract structures which are exemplified by one or multiple systems. In particular we focused on a non-eliminative version of structuralism, which is a position maintaining that structures exist independently. Its proponents furthermore believe that all mathematical knowledge is actually knowledge of structures.

We assumed a non-eliminative structuralist position, thus agreeing that all mathematical knowledge is knowledge of structures *sui generis*. As we generally believe that there are at least some humans having some mathematical knowledge, we need to explain how this knowledge can be obtained – because it is not obvious how one could gain access to these abstract structures. This problem is called the access problem. In particular, we examined a solution based on pattern recognition and argued why (despite MacBride's criticism) this solution is useful for our purposes.

Finally, we investigated the suitability of pattern recognition as a solution to the access problem by looking at an AI implementation of this process. This also serves as a response to MacBride's request for a more illuminating account of the processes involved. Limiting ourselves to a first step of pattern recognition (namely finite-structure recognition), we discussed some criteria that any successful implementation should adhere. We formalized the process of learning to recognize finite structures as determining the truth value of the predicate $I(S, \$)$ for any (finite) system S and any structure \$. Furthermore we discussed appropriate performance measures (accuracy, precision and recall) and training the AI on experience $(S, \$, A)$. Working with systems S, we argued, should be done using finite, direct representations that include enough variety. To represent a structure \$ we adopted Leitgeb's axiomatic description of unlabeled graph theory, enabling us to describe both finite and infinite structures \$. We argued (using the simplistic account that sees knowledge as JTB) that it is not unreasonable to believe that learning $I(S, \$)$ can yield in knowledge. Finally, we presented a test that indicates whether the AI indeed follows the non-eliminative structuralist position we assumed throughout: the predicate $I(\$, \$)$ should be be assigned 'true'.

5.1 Implications of this research

By working out the details of an AI implementation of finite-structure recognition, we have provided a first attempt to formalize pattern recognition as a solution to the access problem. Only practice can show if such an algorithm would indeed be successful, but our analysis has brought this development one step closer. We discussed how our task should be formulated and sketched the restrictions necessary to obtain suitable representations. That this proposal at least has the potential to yield knowledge of finite structures is shown by an analysis of knowledge as JTB, although this analysis should be worked out in more detail.

Our proposal is particularly useful as it provides a test that indicates whether the AI is indeed gaining knowledge of structures using an non-eliminative approach. This check is possible in an early stage, after only implementing finite-structure recognition, the first step of pattern recognition. It is valuable to have the possibility to experimentally verify the correctness of this approach at this moment in the implementation because this makes it possible to make adjustments before continuing with the implementation of the rest of the process.

This proposal is therefore a useful first step in creating an AI implementation for pattern recognition. If we will eventually succeed in creating a successful algorithm for all stages of pattern recognition, we are able to use this algorithm as an argument in favour of pattern recognition as a suitable solution for the access problem. A formal implementation does not allow any vagueness – therefore, creating such an implementation will point out all deficiencies in the arguing and a successful account will automatically be thorough. This can yield in a relevant argument in the debate.

5.2 Further research

What we discussed in this thesis may invite to further research. There are many open questions and possibilities to continue in this direction. In this section we will do various suggestions.

5.2.1 Further steps in implementing pattern recognition

We discussed only the first step of implementing pattern recognition: finite-structure recognition. There is still work to be done in describing the additional steps. Keeping to the human-based process of pattern recognition, we can lay out the phases that follow. Let us suppose that the AI is able to successfully interpret structures as a system. In other words, a test similar to the one we introduced in section 4.6 – which was evaluating $I(\$, \$)$ – is passed. Ultimately, the task for the AI would be to correctly evaluate $I([\$_1,\ \$_2,\ \$_3,\ \$_4, \ldots], \$_{\mathbb{N}})$, where $[\$_1,\ \$_2,\ \$_3,\ \$_4, \ldots]$ is an infinite system consisting of structures (the 1-pattern, the 2-pattern, the 3-pattern et cetera). To do this, the AI needs to be able to recognize and construct this infinite system of structures.

How can we extend the pattern that is recognized in these structures-as-system? One possibility might be to formulate a learning task based on extending structures. We split this into multiple sub-processes. A first step would be for the AI to learn to extend some system S^n, exemplifying the n-pattern, to some system S^{n+s}, exemplifying the $(n + s)$-pattern (for small $n, s \in \mathbb{N}$). For example, this can include learning to extend the set $\{2, 4, 6, 8, 10\}$, which should be recognized as instantiating the 5-pattern, to a set instantiating the 6-pattern. The correct answer should be $\{2, 4, 6, 8, 10, 12\}$.

The next step could now be to learn this for structures instead of systems. This means that the AI would learn to extend the n-pattern itself to the $(n + s)$-pattern. If this succeeds it is definitely a step in the right direction, but it will (by definition) still be limited to finite systems. To achieve our goal of gaining knowledge of the natural number structure (via $I([\$_1,\ \$_2,\ \$_3,\ \$_4, \ldots], \$_{\mathbb{N}})$), the AI still has to learn to expand this to an infinite system as well. The topic of representing infinite systems is especially interesting, and we treat it in the upcoming section 5.2.2.

This suggestion is just one possible way of continuing the AI implementation of pattern recognition, working in the spirit of what is proposed in the previous chapters. Is is only presented here to show that there are indeed approaches that could provide a possible way of finishing what we set out to do. However, this does not at all mean that this suggestion is the correct approach – just that there is at least one.

5.2.2 Representing infinite systems

We return to the problems involved in representing infinite systems, as this also provides an interesting topic for further research. As mentioned in section 4.3, we do not need to be able to represent infinite systems for the purposes intended in chapter 3, as infinite systems are not encountered in the physical world. That is convenient, as direct representations of infinite systems is not possible, because we can impossibly provide a concrete example of an infinite instance (such as an infinite set). But as we mentioned in section 5.2.1 we will need the possibility to represent infinite systems in order to be able to evaluate $I([\$_1,\ \$_2,\ \$_3,\ \$_4, \ldots], \$_{\mathbb{N}})$. Luckily, formal methods can provide us with new tools. We *can* describe an infinite object using recursion: we should just specify a starting element and the recursive rule for creating a next element.

However, this might not be as simple as it appears. Recursion over the natural numbers, for example, does not suffice as a suitable method. By simply defining a first element and the rule to create the next element, the system inherently exemplifies the natural number structure. It is defined just to have a first element, a second, a third, a fourth, et cetera. This makes the question of whether or not this system instantiates the natural number structure trivial. This strategy therefore does not work.

Should we abandon recursion, then? Not yet, because there are more versions of recursion than this one. Think of the recursive definition of all well-formed formulas $\mathcal{L}$ in first-order logic, defined as:

(i) $\mathcal{P} \subseteq \mathcal{L}$ (where $\mathcal{P}$ is the set of propositional letters in the vocabulary of $\mathcal{L}$);

(ii) If $\phi \in \mathcal{L}$, then $\neg\phi \in \mathcal{L}$;

(iii) If ϕ, $\psi \in \mathcal{L}$, then $(\phi \wedge \psi)$, $(\phi \vee \psi)$, $(\phi \rightarrow \psi)$, $(\phi \leftrightarrow \psi) \in \mathcal{L}$;

(iv) Only that which can be generated by the clauses (i)-(iii) in a finite number of steps is a formula in $\mathcal{L}$ [8, p. 35].

This definition consists of multiple recursive rules. The kind of recursion we discussed in the previous paragraph was similar to that of (*ii*), inherently resulting in a natural number structure. However, looking at (*iii*), we see that this does not have to be the case: even when confining only to $(\phi \wedge \psi)$, there are multiple possible resulting structures.

We have seen that some recursively defined objects inherently exemplify the natural number structure, but that this is not the case for all versions of recursion. This formal method for describing infinite systems therefore provides an interesting new view that is worth exploring. All in all, the usefulness of recursion might be a promising topic for further research.

5.2.3 Other topics

Finally, we will present some topics that might be interesting for further research but that we will not discuss thoroughly. The first is obviously the creation of the actual code for an implementation as we proposed. This is the ultimate goal and an important step in determining the success of our proposal. If this proposal were to be made into a concrete algorithm, the suitability of using a particular machine learning algorithm needs to be examined (think of neural networks, Bayesian learning et cetera). Studying which of these is the most appealing is an interesting topic for further research as well.

It might also be interesting to look at the ideas presented in this thesis but instead use the integers or real numbers as a prime example (instead of the natural numbers). The same applies to adopting a different structuralist view than the non-eliminative position that we assumed. Furthermore, the analysis of knowledge as justified true belief could be worked out in more detail, as this analysis ultimately determines whether we have achieved our goal of obtaining knowledge of abstract structures. Finally, there is also a lot of further research that could be done on the theoretical background of mathematical structuralism, such as investigating the use of directed (instead of undirected) graphs for describing structures.

Bibliography

[1] Benacerraf, Paul: *What Numbers Could not Be*. The Philosophical Review, 74(1):47–73, 1965.

[2] Benacerraf, Paul: *Mathematical Truth*. The Journal of Philosophy, 70(19):661–679, 1973.

[3] Brouwer, L. E. J: *Intuitionism and formalism*.

[4] Clarke-Doane, Justin: *What Is the Benacerraf Problem?*, volume 28, pages 17–43. January 2016, ISBN 978-3-319-45978-3.

[5] Dummett, Michael: *The philosophical basis of intuitionistic logic*. In *Studies in Logic and the Foundations of Mathematics*, volume 80, pages 5–40. 1975.

[6] Fantl, Jeremy and Matthew McGrath: *Knowledge in an Uncertain World*. Oxford University Press, 2009.

[7] Frege, Gottlob: *Die Grundlagen der Arithmetik: eine logisch mathematische Untersuchung über den Begriff der Zahl*. w. Koebner, 1884.

[8] Gamut, L.T.F.: *Logica, Taal en Betekenis: Inleiding in de logica*. Uitgeverij Het Spectrum, 1982.

[9] Gettier, Edmund: *Is justified true belief knowledge?* Analytica, 1:123–126, 1963.

[10] Gleason, A. M.: *Fundamentals of Abstract Analysis*. Bartlett, Boston, 1991.

[11] Gödel, Kurt: *What is Cantor's continuum problem?* The American Mathematical Monthly, 54(9):515–525, 1947.

[12] Goodfellow, Ian, Yoshua Bengio, and Aaron Courville: *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.

[13] Goodman, Nelson: *A world of individuals*. In Benacerraf, P. and H. Putnam (editors): *Philosophy of Mathematics: Selected Readings*. Prentice-Hall, 1956.

[14] Hazlett, Allan: *The myth of factive verbs*. Philosophy and Phenomenological Research, 80(3):497–522, 2010.

[15] Hellman, Geoffrey: *Mathematics without numbers: Towards a modal-structural interpretation*. Clarendon Press, 1989.

[16] Heyting, Arend: *Die intuitionistische grundlegung der mathematik*. Erkenntnis, 2:106–115, 1931.

[17] Hilbert, David: *Foundations of geometry*. Open Court, La Salle, IL, 1971.

[18] Horsten, Leon: *Philosophy of Mathematics*. In Zalta, Edward N. (editor): *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2019 edition, 2019.

[19] Ichikawa, Jonathan Jenkins and Matthias Steup: *The analysis of knowledge*. In Zalta, Edward N. (editor): *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2018 edition, 2018.

[20] Korbmacher, Johannes and Georg Schiemer: *What are Structural Properties?* Philosophia Mathematica, 26(3):295–323, 2018.

[21] Leitgeb, Hannes: *On Non-Eliminative Structuralism. Unlabeled Graphs as a Case Study (Part A)*. Philosophia Mathematica, Forthcoming, January 2020.

[22] Leitgeb, Hannes: *On Non-Eliminative Structuralism. Unlabeled Graphs as a Case Study (Part B)*. Philosophia Mathematica, (Forthcoming), January 2020.

[23] Linnebo, Øystein: *Platonism in the philosophy of mathematics*. In Zalta, Edward N. (editor): *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2018 edition, 2018.

[24] MacBride, Fraser: *Can Structuralism solve the Access Problem?* Analysis, 64(4):309–317, 2004.

[25] MacBride, Fraser: *Can ante rem Structuralism solve the Access Problem?* Philosophical Quarterly, 58(230):155–164, 2008.

[26] Maddy, Penelope: *Realism in mathematics*. Clarendon Press, 1990.

[27] Maddy, Penelope: *Naturalism in mathematics*. 1997.

[28] Radford, Colin: *Knowledge—by examples*. Analysis, 27(1):1–11, 1966.

[29] Reck, Erich and Georg Schiemer: *Structuralism in the Philosophy of Mathematics*. In Zalta, Edward N. (editor): *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2019 edition, 2019.

[30] Reck, Erich H. and Michael P. Price: *Structures and Structuralism in Contemporary Philosophy of Mathematics*. Synthese, 125(3):341–383, 2000.

[31] Resnik, Michael D: *Mathematics as a Science of Patterns*. Oxford University Press, 1997.

[32] Rosen, Gideon: *Abstract objects*. In Zalta, Edward N. (editor): *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2018 edition, 2018.

[33] Shapiro, Stewart: *Philosophy of Mathematics:Structure and Ontology*. Oxford University Press, 1997.

[34] Shapiro, Stewart: *Thinking about mathematics*. Oxford University Press, 2000.

[35] Shapiro, Stewart: *Epistemology of Mathematics: What are the questions? What count as answers?* The Philosophical Quarterly, 61(242):130–150, 2011.

[36] Whitehead, Alfred North and Bertrand Russell: *Principia Mathematica*. Cambridge: Cambridge University Press, 1910, 1912, 1913.