Submitted: 19 Aug. 2017.

Wright's Strict Finitism

Master's Thesis

Takahiro Yamada

Utrecht University Student Number: 5580846

Supervisor: dr. R. Iemhoff Second reader: prof. dr. D. Cohnitz Third reader: dr. W. Veldman

Cover page corrected on 25 Aug. 2017; Typographycal and grammatical errors corrected on 31 Aug. 2017.

Contents

1. Introduction: Strict finitism in the semantic realism deba	te	1
1.1 The semantic realism debate as a discussion of worldvi	iews	2
1.1.1 The ontic realism debate and the semantic realism d	lebate	3
1.1.2 Semantic realism \ldots \ldots \ldots \ldots \ldots \ldots		4
1.1.3 Semantic anti-realism		5
1.1.4 Example		8
1.2 The anti-realist attacks and strict finitism		9
1.2.1 Two kinds of anti-realist attack \ldots \ldots \ldots \ldots		9
1.2.2 Strict finitism \ldots \ldots \ldots \ldots \ldots \ldots \ldots		11
1.2.3 The structure of this thesis		14
2. The negative programme: The failure of anti-realism		16
2.1 The key notions of general anti-realism are interlocked		16
2.1.1 The usual approach fails \ldots \ldots \ldots \ldots		17
2.1.2 The notational approach fails		18
2.1.3 The Dedekindian approach fails		20
2.1.3.1 Infinity may be made finite		20
2.1.3.2 An improvement \ldots		21
2.1.3.3 Finitude is made 'not finite' \ldots		21
2.1.4 The Dedekindian approach is harmful \ldots .		22
2.1.4.1 Surveyability predicates		22
2.1.4.2 Finitude is made infinite		23
2.1.5 The conclusion so far \ldots \ldots \ldots \ldots \ldots \ldots		25
2.2 No loophole for arithmetical anti-realism		26
2.2.1 The anti-realist's argument for the bivalence		28
2.2.1.1 The 'implementation' interpretation		28
2.2.1.2 Two statements (C) and (D) \ldots		29
2.2.1.3 The transition from (C) to (D) $\ldots \ldots \ldots$		31
2.2.2 The strict finitist's reply		33
3. The positive programme (1): The basis of strict finitism		35
3.1 The basic concepts \ldots \ldots \ldots \ldots \ldots \ldots		35
3.1.1 Verifiability in practice		35
3.1.2 The features of a surveyability predicate		36

3.1.2.1 The tolerance \ldots	36
3.1.2.2 The non-vagueness	39
3.1.2.3 The actual weak decidability	40
3.1.3 The notation \ldots	44
3.1.4 Indefiniteness: The strict finitist numbers	45
3.2 Against inconsistency	47
3.2.1 Wang's paradox and its presentation	47
3.2.2 Wright's solution	48
3.2.3 A support to Wright's solution	50
4. The positive programme (2): The formal theories of strict finitism	54
4.1 The language	54
4.2 Semantics	55
4.2.1 Model	55
4.2.2 Verification-conditions	57
4.2.2.1 The lemma \ldots	59
4.3 Logic	59
4.3.1 The partial justification of intuitionistic logic	60
4.3.1.1 Modus Ponens and (iii)	61
4.3.1.2 The refutation of (9)	63
4.3.1.3 Other semantic results	64
4.4 Mathematics	66
4.4.1 Arithmetic	66
4.4.1.1 Wright's first observation	67
4.4.1.2 Wright's second observation	67
4.4.1.3 Wright's third observation	68
4.4.2 Intuitionistic mathematics	69
4.4.2.1 The idea	70
4.4.2.2 The formalisation \ldots	71
5. Summary and the topics left untouched	74
5.1 Summary \ldots	74
5.2 The topics left untouched	76
References	78

1. Introduction: Strict finitism in the semantic realism debate

This master's thesis is devoted to an investigation and a support of the constructivist standpoint called 'strict finitism', especially its version which Crispin Wright argued for in his paper 'Strict Finitism' [27]. This position stands in the tradition of the semantic realism debate which the British philosopher Michael Dummett spearheaded. Strict finitism represents the severest anti-realist view. While the vast amount of Dummett's work forms a thorough study against realism in favor of anti-realism, the strict finitist regards anti-realism as insufficient and opposes it.

Wright's paper 'Strict Finitism' was published in 1982 as a response to the issues brought forward in Dummett's paper 'Wang's Paradox' published in 1975^{*1}. Wang's paradox is the paradox which Dummett supposes strict finitism incurs, and he rejects strict finitism, as the result of the consideration on it. Responding to this, Wright, in 'Strict Finitism', offers meticulous analyses and shows us how to conceptualise strict finitism as a standpoint in the semantic realism debate, how to avoid the paradox reported by Dummett, and what the strict finitist formal theories may look like. I consider Wright's analyses as deep enough that we could develop strict finitism as a new mathematical view in the future, if we can trace and defend his thoughts well. This is what I attempt in this thesis: I aim to provide philosophically firm grounds for strict finitism and a starting set of formal theories of it, by examining and modifying Wright's 'Strict Finitism'.

The semantic realism debate has a close relationship with mathematical views: while realism, anti-realism and strict finitism are views in the semantic realism debate, realism has platonism, and anti-realism corresponds to intuitionism. The origin of Dummett's philosophical antirealism is the very first influential form of mathematical constructivism in the modern age, 'intuitionism', created by the Dutch mathematician L.E.J. Brouwer, and Dummett called the classical mathematical view, which intuitionism opposes, 'platonism'. In my understanding, these mathematical views are picked up because they fit the theoretical representations of the corresponding philosophical views: the standpoints in the semantic realism debate can be theorised as semantic theories, and semantic theories specify the mathematical views that fit themselves. Thus strict finitism, as a standpoint in the debate, must have a mathematical view that fits it. However, such a mathematics is not well-known. As far as I could reach by

*1 Wright's 'Strict Finitism' first appeared in 1982 as a journal article (*Synthese*, Vol.51, No.2, pp.203-82), and was reprinted in [18]. We refer to another reprint, collected in Wright's anthology, as [27].

Dummett's 'Wang's Paradox' was first read orally in 1970 at the University of New York and then published as a journal article in 1975 (*Synthese*, Vol.30, No.3/4, pp.301-24, see [9, p.460]). This is why [9] displays 'Wang's Paradox' as a work in 1970 (see [9, p.248]). We refer to this as [7].

the time I write this thesis, Wright's 'Strict Finitism' is the only attempt. This is one of the reasons I consider this paper deserves a close investigation: it contains an appendix where he displays his strict finitist mathematics [27, pp.167-75]. In this thesis, I will try to even slightly advance the strict finitist mathematics by developing Wright's 'outline' of the formal systems.

In this essay, I will behave as though the anti-realist arguments that reject realism are successful enough. The main focus will be on why the strict finitist thinks anti-realism is untenable and what the formal representation of strict finitism will be like. Strict finitism in this essay starts at the point where the anti-realist has made it clear that realism founders. What is interesting is that, as I will describe shortly, anti-realism is attacked by the strict finitist with the same form of arguments which the anti-realist uses to attack realism. In other words, the strict finitist insists that, if one admits that the anti-realist arguments are correct and we must reject realism, then we must reject anti-realism also and slide her position to strict finitism. Certainly, this is merely a conditional statement to support strict finitism since it has the form 'if the anti-realist arguments against realism are sound, then anti-realism is mistaken and strict finitism is correct', and I assume the antecedent without discussion. But I hope one may be content with this basic stance of my thesis, because the study in the semantic realism debate has already had a heap of literature, and what this thesis tries is to investigate the implications of the strategies of anti-realism, the leading standpoint of the debate.

In the rest of this introductory section, I will display the background and the conceptual framework for the discussion of this thesis. I will first explain the metaphysical discussion between realism and anti-realism over the reality (determinacy) of the world. The interest in the determination of the world starts as a talk between intuitive worldviews, but it turns into a theoretical debate thanks to Dummett's keen philosophical insights. After displaying what this theoretical realism debate is like, I will explain the arguments used by the anti-realist to point out the conceptual inadequacy of realism. The strict finitist learns them as the model of how to attack a point of view in this debate. I will have strict finitism enter the picture as a rival position against anti-realism.

1.1 The semantic realism debate as a discussion of worldviews

We first look at the general settings of the semantic realism debate. In this subsection, I will start with realism and anti-realism as intuitively conceived worldviews, then introduce the Dummettian framework of the debate, and explain what the realism and the anti-realist as theoretically contrived standpoints aim to establish.

1.1.1 The ontic realism debate and the semantic realism debate

To begin with, I want to describe what the usual kind of discussion between realism and anti-realism is like. Realism and anti-realism are worldviews, in the sense that they are about the 'reality' of the world. Given that we encounter many realist views and anti-realist views in philosophy, it seems to be important to distinguish two kinds of worldview: the worldviews about specific aspects of the world are the first kind, and the general worldviews are the second^{*2}. On this conception, a realist view about, say, mathematics is a realist view regarding the mathematical aspect of world, and an anti-realist view about morality is an anti-realist view regarding the moral aspect of the world. We take it that one's view about the whole world is the combination of the views about the aspects of the world. Let us say, again usually, a realism (and an anti-realism) about an aspect, X, of the world takes the form of 'realism (anti-realism) about objects of certain kind, say Y, and hence about aspect X'; realism about objects of kind Y is defined as the view that such objects exist independently of whether we know that they exist; and anti-realism about Y is defined to be the negation of realism about Y, i.e., either that they do not exist, or that they exist but only dependently on our recognitional capacities. General realism (and general anti-realism), then, is defined to be the standpoint which adopts realism (anti-realism) for each aspect of the world. For instance, a realist view about mathematics contains the view that sets exist independently of us, and hence that the mathematical aspect of the world is real; and an anti-realist view about morality includes *either* the view that moral facts do not exist, or the view that moral facts exist only because, e.g., we have socially constructed them. It is worth noting that, on this conception, the 'reality' of the world is conceived in terms of what exists in the world (e.g., mathematical objects) and how they exist (i.e. whether independently or dependently). Brouwer's intuitionistic mathematics is the typical example of an anti-realist view of the mathematical aspect of the world – and as I said earlier, is the origin of Dummett's philosophical inspiration: we can observe traits of Dummett's standpoint already residing in Brouwer's thoughts. Intuitionistic mathematicians think that mathematical objects exist just because they are mentally constructed, and hence only the legitimate states of affairs are those which they can confirm by means of proof. $*^3$

^{*2} I learned this view on realism and anti-realism from Dummett's writings, especially [4, p.145], [6, p.226] and [10, p.230] among others.

^{*3} By the way, what will happen if there are, e.g., two kinds of mathematical object that are relevant to the reality of the mathematical aspect of the world, and someone takes realism about the first kind and anti-realism about the second? One of the cases which happen in this case is that she is self-contradictory. To deny this to be the case, a metaphysician might argue that there is one principal kind of mathematical object, and that a worldviews about any other kinds must be consistent with the worldview about the principal one.

The realism and the anti-realism described above are *ontic*, and the grounds of them are, ultimately, mere intuitions – this is Dummett's diagnosis [5, p.202], [9, p.xxv], [11, pp.9-12, pp.14-5]. This is why the conflict between them never seems to be resolved, and will never be resolved. Dummett's keen insight consists in the observation that this metaphysical conflict can (only) be solved by comparing the *semantic* grounds of those worldviews. The semantic realism debate urges the realist and the anti-realist to spell out their own conceptions of meaning and convince the other with them: this debate is where the better theory of meaning is the better worldview. We call the theory of meaning that suits the realist, 'semantic realism', and that which suits the anti-realist, 'semantic anti-realism'. It is true that 'ontic' metaphysicians raise objections to Dummett's view. Many would say that realism and antirealism *are* ontic standpoints and never semantic: it is indeed an ontic conception of the reality to think that a view about an aspect of the world is primarily a view about objects in the world^{*4}. But, in the Dummettian semantic debate, the aspects of the world are individuated through the syntactical features of the statements. The realist (and anti-realist) in this debate discusses the meaning and the linguistic behaviour of numerical expressions instead of numbers, of moral or evaluative expressions instead of moral facts, and of sentences in the past tense and in the future tense instead of time.

1.1.2 Semantic realism

The definitional character of the Dummettian realism ('semantic realism') about an aspect of the world can be boiled down to the acceptance of the 'principle of bivalence'. This is the principle that says that every statement either holds or does not hold. It is the semantic realist's custom to express this by 'every statement is either true or not true' using the notion of truth, and to call being not true 'false'. The definitional character of the Dummettian antirealism ('semantic anti-realism'), on the other hand, is the rejection of the bivalence [3, p.14] [8, p.56, p.62, p.64, p.66] [10, p.230] [11, p.9]. The realist accepts the principle of bivalence because she assumes that a statement, which reflects a state of affairs concerning the aspect of the world at issue, is true just in case the state of affairs holds, and is false just in case it fails, and that any state of affairs either holds or fails, independently of our recognition. Here we can see how the realist relates the determination of the world to the notion of truth. Roughly speaking, the reasoning behind it is as follows: the world is the collection of facts; and a fact is a state of affairs that holds good; a statement stands or falls together with the state of affairs

^{*4} There in fact are philosophers who clearly states that the problem of the reality is not a matter of semantics or language, but of ontology [2, pp.3-4], [20, p.161]. I will leave it inconclusive whether a theory about meaning or an ontic theory grasps the world better, although my opinion is that this dilemma is ill-posed and we need both kinds of theory about the world, plus one. I think that the most effective form of a worldview would be the triple of a semantics, an ontology and an epistemology [22].

which it stands for; therefore, to say that (an aspect of) the world is determined independently of the human recognition is to say that the truth-value of each statement (which belongs to the class of statements that stands for the aspect of the world) is independently determined.

Semantic realism thus uses the notion of truth. One could think that truth is chosen in order to have the principle of bivalence obtain most straightforwardly: the semantic realist is the one who endorses the bivalence, and she does so by equating that a statement holds with that it is true. She makes use of the nature of the notion of truth that every statement, even if there is no clue available about the truth-values, is either true or false. This brings about a crucial character to the realist account of meaning, when it is combined with a certain feature of the theories of meaning in this framework. The theories of meaning which appear in the semantic realism debate should treat meaning in the form of the 'understanding of meaning': a theory of meaning is a theory of how meaning is understood [6, pp.216-7]. Thus, the realist conception of understanding of meaning is the truth-conditional conception, according to which one is considered to know the meaning of a statement, P, just in case she knows the condition under which P is true. A realist theory of meaning is equipped with a 'theory of reference', and this theory specifies truth-conditions, in other words, it determines for each statement whether truth applies to it or not^{*5}. A realist theory of reference assigns the references to the terms and the predicates of the language at issue. Here, the realist takes a set-theoretical way of thinking: the reference of a term is an object; and the reference of, e.g., a one-place predicate is a set of objects (or a function from objects to truth-values). [8, p.84]

Referring to this conception, we say that truth is employed as the central notion to specify what the meaning of a statement is. The crucial character of the realist theory of meaning is that, because truth is evidence-transcendent, it results in the evidence-transcendent kind of understanding. In the ordinary life, we may find many pieces of evidence for and against statements, but we can also easily come up with statements which we cannot find any evidence neither for nor against. But one of the implications of the law of bivalence is that even such statements are true or false. The realist conception of meaning is thus that we understand this kind of truth-condition.

1.1.3 Semantic anti-realism

One can think of mathematical statements as the example of the statements which we have no evidence for nor against. But in order to sharpen our analyses, we use the notion of decidability (and that of undecidability) instead of the notion of evidence. We postpone the detailed discussion about decidability and, for the time being, let us be content with

^{*5} A theory of meaning in this framework is said to possess its 'theory of reference', its 'theory of sense', and its 'theory of force' [8, p.40, p.84]. But I will not discuss the latter two theories.

the following rather simple characterisation: a statement is decidable just in case *either* we have a proof of it or a proof of the negation of it (i.e., a 'disproof'), *or* we have a decision procedure for it which brings us within a finite number of steps to the realisation of either that it holds or that it fails. We call the statements which are not decidable, 'undecidable' [21, p.329]. Dummett points out that there are at least three sentential operations which induce undecidability: (1) the use of quantification over an infinite totality (or a totality beyond the human capacity): (2) the use of the subjunctive conditional; and (3) the reference to the space-time which is not accessible to us [8, p.46; p.60] [12, p.69]. Mathematics is the area in which one regularly uses the quantification over an infinite totality, and is thus prone to the phenomenon of undecidability.

The anti-realist believes that the central notion of the theory of meaning must be epistemic: a theory of meaning must appreciate the epistemic distinction between the verifiable statements and the non-verifiable statements; and it must only allow the verifiable statements to be meaningful (significant). The notion the anti-realist chooses to employ is the notion of verifiability in principle, instead of that of truth: for the anti-realist, a statement holds good just in case it is verified with a finite method [8, pp.45-6, pp.63-4, passim]. A statement is verifiable in principle particularly (but not necessarily) when we possess a decision procedure for it. The decision procedure for a statement, which takes finitely many steps at most, is a method to know whether the statement is to be affirmed or not, as the result of the implementation of the prescribed algorithm. To make a contrast with strict finitism, which we will see shortly, the anti-realist's basic stance is that finite methods are harmless no matter how long they may take to bring about the results, and we can see this idea in Brouwer's work already [1, p.510]. The notion of 'verifiability in principle' is itself a focus of analysis in this thesis (2.1). But for the time being, we can understand this notion as follows: if the result of the decision procedure is going to be affirmative, we say the statement is 'verifiable in principle', and when not going to be affirmative, we say 'falsifiable in principle' – even if the result cannot actually be given to us. The anti-realist uses these notions and equates a statement's holding good with its being verified, and not holding good with being falsified: it is an epistemic matter whether a statement holds or not. The anti-realist calls the condition under which a statement holds the 'verification-condition', just as the realist calls it the 'truth-condition'. On this conception, the bivalence fails for any domain of statements which contains at least one undecidable statement, and mathematics is such a domain.

An anti-realist theory of meaning possesses its theory of reference just as a realist theory of meaning does, and the theory of reference determines what statements are verifiable in principle. An anti-realist theory of reference assigns to a term an object^{*6}, and to a predicate

^{*6} One may expect that the anti-realist uses the notion of objects differently than the realist does. In fact,

'an effective means of recognizing, for any object, a conclusive demonstration that the predicate applies to that object' [8, p.84]. This is Dummett's phrase, but I think that we can interpret this (since it must not be the case that any predicate applies to any object) as 'a means of deciding, for any given object, whether the predicate applies to the object, which only has finitely many steps at most', i.e. a decision procedure. This reading is consistent with the fact that in the realist case, a function from objects to truth-values is assigned to a predicate as its reference. Notice, by the way, that a proof and a disproof for a statement can be taken as a kind of decision procedure for the statement, which have only one step: proofs and disproofs may have complex structures, but as long as we recognise an entity as a proof or a disproof for a statement, it lets us immediately know that the statement is correct or incorrect.

The choice of the central notion also affects what it is to know the meaning of a statement: for the anti-realist it is to know the verification-condition. To know the meaning of a statement is to know when the statement is said to be verifiable in principle. I think that it is basically an epistemic activity to spell out in what cases a given statement is considered to be verifiable in principle. In scientific cases, for example, one may require 'stronger' evidence for a statement than the others may require. But it is ultimately a distinctive activity from establishing antirealism as a philosophical standpoint to clarify the several opinions about the verificationconditions and to justify one of them: of course the anti-realist herself can engage in such an activity, but she also can simply adopt the epistemologist's opinion. Also, notice that the anti-realist theory of meaning needs to have a theory of reference, and to have such a theory is to complete the epistemic task to spell out the collection of the verification-conditions: the anti-realist theory of meaning inevitably contains such an epistemic part. In this thesis, by the way, I will be concerned mainly with the mathematical cases and say that we know the meaning of a mathematical statement just in case we recognise, for anything mathematical and constructible in principle, whether it is a proof of the statement or not. But when it comes to 'what a proof of a given statement is', I will simply accept what the (Brouwerian) mathematicians would think to be it.

It is important to notice that to possess a decision procedure for a statement is sufficient but not necessary to know the meaning of it. This is because, when we have a decision procedure for a statement, we know in what cases the statement is said to be verified: it is when the final result of the procedure is going to be affirmative – even if we cannot in practice see the result. Even when we have no decision procedure for a given statement, we can be said to know the meaning of it if we can recognise a proof as it is when it is presented to us [8, p.70].

it is very rare to see, in the semantic realism debate literature, the notion of objects discussed explicitly. My opinion is that the anti-realist can adopt the same notion of objects as the realist's, because what matters to the realist and the anti-realist is not the ontic question 'What exists?', but the semantic question 'Which is the better conception of meaning?'.

This is why the mathematicians can try to prove a statement for which they have no proof yet. Their knowledge of the meaning of the statement lies in their ability to discriminate the proofs of the statement from those which are not. The mathematicians can recognise what they construct as a proof because they know the meaning of the statement.

The anti-realist insists that when we have a decision procedure for a statement, the principle of bivalence holds for it [6, pp.243-5] [8, p.70] [13, pp.348-9]. For example, consider the predicate 'x is a prime number' and let us write it as Pr(x). The anti-realist can say that we surely know the meaning of Pr(n) for every n (hence so do we the meaning of Pr), because we have a decision procedure for Pr(x) for any x: the sieve of Eratosthenes is one of the decision procedures. The anti-realist insists that, further, Pr(n) either holds or fails to hold, no matter how large x is. One can in principle attain the result of the sieve of Eratosthenes and can in principle know whether the Pr(x) is correct or incorrect [8, p.73]. We will carefully examine this point later (mainly in 2.2).

Anti-realism as the philosophical project which Dummett prompted aims to establish a general worldview. The Dummettian anti-realist starts attacking realism and defending her position in the area of mathematics, and by doing so tries to find a plausible general account of meaning [3, p.17] [6, pp.226-7] [8, pp.70-1]. We will concentrate on the mathematical case and look at how the anti-realist attacks realism, and whether the anti-realist can defend from the strict finitist attack. The motivation of this strategy is the fact that mathematics is the area for which Dummett most frequently argued for anti-realism, for which the largest amount of studies have been devoted, and whose statements are easiest to deal with, since they are expressible in a formal language.

1.1.4 Example

Before moving to the next topic, let us here introduce an example of a statement in mathematics with respect to which the realist and the anti-realist disagree. We define several notions at first. We call the decimal part of π , d. We can in principle calculate any digit of d. Let us regard d as a sequence and write the *i*-th number of it as d(i). Notice that $\pi = 3 + \sum_{n=0}^{\infty} d(n) \cdot 10^{-n-1}$ and d(0) = 1, d(1) = 4, d(2) = 1, etc.^{*7} Define, for any $n \in \mathbb{N}$, $n < k_{99}$ to mean that $\neg \exists j \leq n \forall i < 99 [d(j + i) = 9]$ (in words, there is no sequence of consecutive ninety-nine 9's which starts by the *n*-th member of d); $k_{99} \leq n$ to mean that $\exists j \leq n \forall i < 99 [d(j + i) = 9]$; and $n = k_{99}$ to mean that n is the least j such that

^{*7} On the realist view, d is an already determinate infinite sequence of natural numbers, and according to intuitionism, d is a sequence of natural numbers to which each digit is added as we calculate them. Note that this intuitionistic conception of an infinite sequence is not something which directly can be implied by semantic anti-realism, but something which belongs to the tradition of the Brouwerian intuitionistic mathematics. We will revisited the intuitionist's way of dealing with mathematical objects later.

 $\forall i < 99 \left[d(j+i) = 9 \right].$

The opinions of the realist and the anti-realist differ regarding $\forall n \in \mathbb{N} [n < k_{99}]$: whereas the realist regards this statement as either true or false, it is not acceptable for the anti-realist that this either holds or fails. Notice, first, that while the realist naturally takes any well-defined predicate as significant, the intuitionist also regards both d(x) = 9 and $x < k_{99}$ significant. This is because, for any natural number n, no matter how large the given n is, the intuitionist can in principle decide whether d(n) = 9 or not, and whether $n < k_{99}$ or not, by looking at the final results of the calculation. However, we now by proof know neither that it holds, nor that it fails. So, it should be when we have a decision procedure that this statement is decidable. But a decision procedure for this statement would tell us *either* that for each natural number $n, n < k_{99}$ holds, or that there is an n such that $k_{99} \leq n$, as the result of the implementation of it. We do not have such a procedure. Even though for each n it is decidable whether $n < k_{99}$ or not, the universally quantified statement is not (now) decidable. Therefore, it is unacceptable for the anti-realist to conclude that $\forall n \in \mathbb{N} [n < k_{99}]$ either holds or fails^{*8}.

1.2 The anti-realist attacks and strict finitism

How strict finitism emerges and attacks anti-realism is ironic, because it is the anti-realist's strategies against realism that call upon the strict finitist ideas and give rise to an attack to anti-realism from the strict finitist. Strict finitism figures in the debate as a natural standpoint when one considers how the anti-realist attacks realism. In this section, I will describe what the anti-realist's strategies are, how strict finitism can be formed out of the very same strategies, and what strict finitism is like. How successful the strict finitist attack is, however, should be postponed to the next section (2.).

1.2.1 Two kinds of anti-realist attack

Today's participants in the semantic realism debate recognise two related but distinctive kinds of anti-realist attack on realism, thanks to Wright's illumination [29, esp. pp.13-23]. Wright calls the first the 'acquisition argument', and the second the 'manifestation argument'. Both arguments impose, as he prefers to regard, specific kinds of 'challenge' on realism such that if the realist fails in answering to them, it means that the realist understanding of meaning is inadequate.

The acquisition argument demands an explanation of how one could acquire the realist

 $^{^{*8}}$ I learned this example from [24].

Certainly, it might happen in the future that mathematicians come up with a method to grasp relevant general features of π so that they can judge whether $\forall n \in \mathbb{N} [n < k_{99}]$ holds or not. But we do now not know for sure that this invention in the future really happens. To say that we know this will happen would be tantamount to say now sthat $\forall n \in \mathbb{N} [n < k_{99}]$ now either holds or does not.

understanding of meaning. As we saw, the realist understanding is evidence-transcendent and covers non-verifiable statements. Some of statements obtain without being known to obtain by us. It seems to be impossible to learn with case-based instructions the truth-conditions of such statements, since we may not be able to know that they are true, when they are true. For instance, there are statements containing the quantification over an infinite totality, such as $\forall n \in \mathbb{N} [n < k_{99}]$, for which there is no decision procedure. In the case of a finite statement, e.g. $\forall n < 9 [n < k_{99}]$, one can easily learn under which circumstances the statement holds: it suffices to check that $n < k_{99}$ holds for the first nine n's. But by what means could one learn what it is that $\forall n \in \mathbb{N} [n < k_{99}]$ holds good?

I cannot examine the history of the correspondence between the pros and the cons over this kind of argument in this thesis, but it would be worthwhile to summarise the realist's answers which Wright is aware of, when he introduced the term 'the acquisition argument' ([29]). He sketches two possible realist responses. (1) The realist may say that her understanding of statements can be acquired by the 'idealization' of the understanding of the anti-realistically uncontroversial statements [29, p.15]. On my reconstruction, this idea utilises the fact that the intuitionist does not reject the significance of statements such as $\forall n < 9 [n < k_{99}], \forall n < 999 [n < k_{99}], etc.: in other words, the intuitionist admits that the verification-conditions of the quantified statements can be learned as long as the ranges of the quantifications are finite. The realist thinks that the truth-condition of <math>\forall n \in \mathbb{N} [n < k_{99}]$ can be attained as the 'limit' (or an expansion in a sense) of the verification-conditions. Wright, as an anti-realist, is against this line of thought: the realist understanding is not like any 'limit' version of the anti-realist understanding [29, pp.14-5].

(2) The realist may say, on my interpretation of Wright's description, that the realist understanding can be acquired via the compositionality of language. The condition under which a statement holds, and hence our understanding of it, seems to be piecemeal: such a condition is composed of the ways the constituents of the statement contribute to the condition. I think that there are two fundamental suppositions behind this: (1) every semantic atom has its own way of contributing to the condition under which the statement containing the semantic atom holds; (2) the way a semantic atom contributes is uniform, regardless of whatever other semantic atoms it is combined with. By 'semantic atoms', I refer to the individual notions out of which we compose a statement. The realist would say that we know the truth-condition of $\forall n [n \in \mathbb{N} \to n < k_{99}]$ for example, because we know how universal quantification ($\forall n$), belonging to the set of natural numbers ($n \in \mathbb{N}$), conditional (\rightarrow), etc. affect the condition under which the entire statement holds. We have learned each of their effects before encountering this statement, and each of them is not anti-realistically controversial. Since they do not change, we can understand $\forall n [n \in \mathbb{N} \to n < k_{99}]$, a mere composition of what we have already learned. Wright admits that this line of answer is plausible.^{*9}

Wright, then, leads us to the other kind of argument, the manifestation argument. This argument urges the realist to show that she really has the realist understanding: to the antirealist, the realist understanding appears to lack a practical sort of evidence. The challenge needs the later Wittgenstein's so-called 'use theory' of meaning as its fundamental assumption: roughly speaking, it identifies the meaning of statements with the use of them, and invites us to think that to understand the meaning is to possess a specific sort of practical abilities [29, p.16]. To explain the cogency of this theory, Dummett points out that meaning is an instrument of communication: meaning is unlike to a mental state and has no part which cannot be observed when conveyed [6, p.216]. Now, there indeed are some kinds of understanding whose corresponding abilities can be specified. Wright gives the following examples: the abilities corresponding to understanding in the sense of appraisal of evidence for or against statements, those corresponding to understanding in the sense of recognition of the logical connections from and to statements, etc. can be specified. However, when it comes to undecidable statements, it seems to be impossible to specify the practical abilities we might possess that indicate our grasp of the evidence-transcendent truth-conditions of them. In the case of decidable statements, on the other hand, we may easily specify the abilities for the understanding of the meaning. For instance, to manifest the understanding of the condition under which 'this is salty' holds good, one can put the material at issue into her mouth [29, pp.15-7].

1.2.2 Strict finitism

I will not investigate whether these two kinds of argument conclusively defeat realism or the realist can provide sound replies. Rather, I assume that they have a plausible appeal and, in stead, consider the standpoint that puts anti-realism in doubt: namely, strict finitism.

The anti-realist attacks realism saying that it seems to be impossible to acquire or manifest the realist understanding of meaning. What inspires this attack is ultimately the feature of truth that it is evidence-transcendent. It appears outrageous, to the anti-realist, to insist that truth is the very semantic concept which we can learn as our practical abilities when we acquire, and which we can embody with our abilities when we manifest our understanding of

^{*9} I also agree to this thought. The anti-realist, or the intuitionistic mathematician in particular, would admit that $\forall n \in \mathbb{N} \ [n < k_{99}]$ is significant, although not decidable. This would be because in intuitionistic mathematics, there is a way of understanding an infinite sequence: the intuitionistic mathematician can understand an infinite sequence when there is a 'law' (a 'law-like' description) that governs it. For example, she understands infinite sequence $1, 1, 1, \cdots$, because it can be described as an infinite sequence made only of 1. In other words, I think, the intuitionistic mathematician would expect a general description when she sees a universally quantified statement, and it is the way she understands such a statement. $\forall n \in \mathbb{N} \ [n < k_{99}]$ is significant because the intuitionistic mathematician would acknowledge as a proof of this statement something which conclusively tells the suitable general feature of the infinite sequence d made of π .

meaning. This is why the anti-realist employs the notion of verifiability in principle. But the strict finitist thinks that verifiability in principle is also an outrageous notion to characterise the actual human beings' understanding. Dummett, in 'Wang's Paradox', considers a doubt in the form of the acquisition argument against anti-realism: Must it not be through the instructions about the constructions which we can in practice carry out that we acquire the use of expressions? ([7, pp.248-9]) Wright states doubts about anti-realism in the form of both the acquisition and the manifestation arguments. It is unclear how we are supposed to acquire the grasp of the notion of verifiability in principle when we lack a guarantee of humanly feasible verification or falsification, namely, when there is no practical possibility of proof or disproof. And it is doubtful that we could test whether someone indeed possesses such an 'ethereal' notion that transcends practical possibility [27, p.111]. The central notion of the theory of meaning must be more modest: verifiability in principle should be replaced with the notion of verifiability in practice. This is the strict finitist's basic stance.

Strict finitism is the standpoint which employs the notion of verifiability in practice as its central notion: for the strict finitist, to say a statement holds good is to say that it is verifiable in practice. One could take this position as nothing more than a slightly 'frugal' version of semantic anti-realism, because the phrase 'in practice' in place of 'in principle' is apparently the only conceptual difference. Indeed, the strict finitist will not regard some statements as holding good which the anti-realist does; and this might seem to be a small difference. We, however, can see how remarkable the differences between the two standpoints are by looking at the characterisation of (1) the decision procedures and (2) the significant expressions, for the strict finitist.

(1) The strict-finitistically relevant kind of decision procedure is the kind of procedure which produces the result in a humanly feasible way. For example, consider the predicate Pr in the preceding section (1.1.3). We admit that there are numbers too big to actually determine their primality. Of course what numbers are 'too big' is dependent on the speed (or the efficiency) at which the procedure we choose gives the result, but we think that no matter what procedure we choose, there always are such numbers. For example, we could say that it is decidable in principle whether $2^{1000^{1000}} - 1$ is a prime number or not, but it is too big for us to determine in practice by using the sieve of Eratosthenes – let us use that $2^{1000^{1000}} - 1$ is too big, as our running supposition in this thesis. If n is too big (for a chosen procedure), then Pr(n) is neither verifiable nor falsifiable in practice: it is not acceptable to assert that n is either a prime number or not.

(2) According to my understanding, strict finitism is not different from anti-realism in terms of the criterion of significance of statements: one is said to know the meaning of a statement just in case she knows when the statement is verified. To quote Wright's phrase,

Plausibly, any statement to which we attach a clear sense is such that we have a conception at least of what it would be to *try* to uncover grounds for believing or disbelieving it. [27, p.119, original emphasis]

But strict finitism has another 'layer' which affects the significance of a statement: on the strict finitist conception, some expressions that *are* statements for the anti-realist *are not* statements. As long as something is a statement with the verification-condition, it is significant. But some are in the first place not legitimate as statements, and therefore not significant. As a rule, an expression is not legitimate if it is too complex. I think that it would be helpful to introduce the notion of 'intelligibility' to see what expressions are considered to be legitimate. We use this notion, because now what is legitimate is dependent on how powerful our recognitional capacities are. Our 'recognitional capacities' include not only our ability to think, but also the methods we employ to determine. We also use the notion of 'natural-number denoting expressions', or 'nde's' for short [27, p.167]. We will usually write \bar{n} for an nde of natural number n.

First, we need to notice that some nde's are unintelligible. Roughly speaking, we cannot actually recognise an nde as an nde when it is too complex. Imagine a case where something which looks like a string of symbols is presented in front of an agent. The leftmost character is 1. The second is 4, and the next is 1. Let us say, as long as she can actually confirm, the string looks like the sequence made by putting consecutively the digits of d we defined in the preceding section (1.1.4). But the string continues beyond her sight. So she has no reason to expect that the string really ends. In this case, if she is an anti-realist, and if she is taught that the string eventually ends and given the guarantee that the string only contains the ten kinds of Arabic numeral symbol, then she accepts the string as one long nde for a huge natural number. However, if she is a strict finitist, since the string surpasses the agent's recognitional abilities. The anti-realist accepts an expression as legitimate based on the information which stipulates it is legitimate, but the strict finitist, to accept an expression has to be something which consumes but does not surpass the agent's recognitional capacity.

So, for a mathematical statement P(n) to be intelligible, the number n which the statement is about must be given through an intelligible nde (\overline{n}) . But, furthermore, the statement itself should not be too complex. For example, if $P(\overline{n})$ is a conjunctive statement with one million conjuncts, it may surpass our cognitive ability and be unintelligible, whereas there would be no problem for the anti-realist to accept it as a legitimate statement. Surely, for the strict finitist also, a statement is significant if we know when it is that the statement is verified, but it is needed that the statement is in the first place intelligible. As the result, we should be able actually to check whether any given candidate of proof is really a proof – this is what the strict finitist means when she equates that a statement holds good with that it is verifiable in practice.

Accordingly, on my understanding, a strict finitist theory of reference assigns objects (numbers, in the case of a mathematical language) only to intelligible terms as their references, and feasible decision procedures to predicates. Also, when a statement becomes too complex due to, e.g., an iterated application of a sentential operator, the theory of reference does not assign a reference. It is excluded from the collection of significant statements. Note that what expressions are intelligible is determined outside the theory of meaning, and it is genuinely an epistemological matter.

1.2.3 The structure of this thesis

The description above may be sufficient as the characterisation of strict finitism. However, the strict finitist cannot insist with this alone that her position is established as the correct view about the world. She needs to argue that anti-realism is ill-conceived and that strict finitism has firm grounds. In other words, she has to undertake a 'negative programme' and a 'positive programme' for strict finitism, after Wright's terminology [29, p.29]: the former is a project aiming to defeat the rival theory, and the latter aims to defend her own position. The tasks I engage in in this thesis also are divided into these two directions.

In section 2, I will undertake my negative programme. I will examine and support the strict finitist attack on anti-realism in the form of a manifestation argument presented by Wright's 'Strict Finitist'. We can interpret the major part of his paper as an argument which starts with a critical observation that points out which notions are keys for the anti-realist and must be explained in order for her to be able to manifest the understanding of meaning; and which ends with an assessment that the anti-realist key notions are interlocked, i.e., one notion cannot be explained without recourse to the other key notions. If this argument is correct, it follows that the anti-realist understanding is not manifestable, i.e., anti-realism is apprehensible only to those who already have apprehended the key notions. This is enough evidence to regard anti-realism as an untenable position.

Here I must note that this negative programme is not completely independent of the positive programme. A particular argument which I will endorse that anti-realism is wrong will use the '(weak) decidability' of the 'surveyability' predicate, which is characteristic to strict finitism (2.1.4). The justification that the surveyability predicate is really (weak) decidable will be done in the part of the positive programme (3.1.2.3).

My plan of the positive programme is twofold, and section 3 is the place for the first half of it. I will there address and answer to the doubt about the consistency of the strict finitist conception of meaning. The 'surveyability' predicate which the strict finitism uses is said to incur a serious inconsistency. The anti-realist claims that this inconsistency takes the form of the problem known as Wang's paradox, and it seems that the one who first took up this issue is Dummett $[7]^{*10}$. He analyses the origin of the inconsistency and concludes that the strict finitist cannot avoid this. But Wright gives a way to deal with this paradox. I will reconstruct his idea and try to give a support to it.

In the second part of the positive programme (in section 4), I will investigate the formal theories for strict finitism. By the strict finitist 'formal theories', I mean the semantic theory, the logic system (i.e. proof system) and the mathematics which are suitable for strict finitism as a theory of meaning. Wright's paper 'Strict Finitism' includes an appendix where he embarks on this enterprise, but it does not seem to have had a flourishing development, as far as I know.

^{*10} One could say that Wang's paradox is a variation of what we now call the 'Sorites paradoxes'. But this thesis treats it as an issue specific to strict finitism, and will not examine from a more general perspective.

2. The negative programme: The failure of anti-realism

In this section, we reconstruct and support Wright's arguments that anti-realism is not a tenable position. We will at first see why general anti-realism is mistaken, and then argue that the grounds for the local anti-realism about arithmetic are weak.

2.1 The key notions of general anti-realism are interlocked

Wright's argument against general anti-realism can be boiled down to that the notion of verifiability in principle cannot be explained by using only the notions which the strict finitist comprehends. In 'Strict Finitism', Wright proposes the following conceptual analysis (P) of decidability in principle:

(P) An undecided statement, S is decidable in principle just in case an appropriately large but finite extension of our capacities would confer on us the ability to verify it or falsify it in practice. [27, p.113]

We understand that a statement is (now) undecided just in case we do not (now) know whether it is correct or not. I said that the decidability in principle of a statement is a sufficient condition for the verifiability in principle of the statement. So, here we see basically three kinds of notion: (1) decidability in principle (and verifiability in principle); (2) finitude; and (3) decidability in practice (and verifiability in practice). (P) analyses decidability in principle based on decidability in practice and finitude: decidability in principle is attained as a finite extension of the notion of decidability in practice. We would say that $2^{1000^{1000}} - 1$ is too big for the actual human beings; and therefore that whether $2^{1000^{1000}} - 1$ is prime or not, is decidable only in principle; but if our capacities to calculate were more powerful by a finite extent, it would be decidable *in practice* whether $2^{1000^{1000}} - 1$ is prime or not – this is the basic line of thought of this explanation.

The conclusion of Wright's argument is that finitude and decidability in practice cannot be fully explained to the strict finitist. But decidability in practice is a notion that the strict finitist is supposed to already know and use. So the focus of the argument is on finitude. Notice, however, that this conclusion says that *finitude* cannot be understood by the strict *finitist*. This may sound strange, because strict finitism appears to be the view which only admits what is finite. The strict finitist must understand finitude. What is happening here, I think, is that the notion of finitude is divided into two kinds: one is the finitude which the strict finitist understands, the other is that which she does not. For instance, natural number 2 is no doubt finite for the strict finitist (and for the anti-realist also), but $2^{1000^{1000}} - 1$ may

be too big to be finite for the strict finitist, whereas the anti-realist accepts it as finite. To avoid confusion, it may be useful to call the kind of finitude that the strict finitist admits, the 'first finitude', and the kind which only the anti-realist admits the 'second finitude'. In what follows, we will see how the anti-realist who understands both may try to explain her notion of finitude, which includes both, to the strict finitist who only understands the first finitude, and how the anti-realist explanations fail.

We will see three approaches from the anti-realist camp which Wright discusses. I call them, (1) the usual approach, (2) the notational approach, and (3) the Dedekindian approach. None of them will turn out to be successful. Let us see that they fail first, and then see that especially the Dedekind approach is not only unsuccessful, but rather harmful.

2.1.1 The usual approach fails

One may first try to give an explanation of the anti-realist's notion of finitude by modifying the classical (and hence realist) conception of finitude: a set is finite just in case there is a one-to-one mapping from the set onto an initial segment of the set of natural numbers. The anti-realist version of this definition would be that a set is (anti-realistically) finite just in case one *can* locate a one-to-one mapping, f, from the set onto an initial segment of the set of natural numbers such that for any member, x, of the set, one *can* recognise f(x).

Notice here that to the anti-realist version, the phrases 'can locate' and 'can recognise' are added. They are inserted because on the anti-realist conception of meaning, to assert that something exists is to assert that the existence of it is verifiable (in principle). To say that there is a function is to say that one can (in principle) recognise that there is such a function. One must guarantee that every value of the function is attainable (in principle).

But we should notice also that this formulation mentions the set of natural numbers, a typical example of an infinite entity. We need a modification to avoid mentioning this, because the strict finitist would not understand it. I think that this is what Wright intends when he writes:

[...] a set will count as finite [...] if and only if an *effective* one-to-one function, f, can be located such that for each member x of the set, f(x) can be recognized to be a particular positive integer m, \leq a fixed positive integer n. [27, p.125, original emphasis]

I interpret that a function is effective just in case every value of the function is attainable (or calculable in particular): so this is essentially the same as that f(x) can be recognised. Apart from this point, the proposal Wright puts is the result of rejecting the notion of the set of natural numbers and incorporating the notion of being 'less than or equal to a fixed positive integer'^{*11}.

 $^{^{\}ast 11}\,$ Indeed, Wright's formulation does not require the function to be an 'onto' function. There may be

However, this approach has, Wright points out, a difficulty. The anti-realist does not seem to be able to explain successfully the notions of possibility appearing at the italicised points. When she says 'one *can* locate' a function, does she mean 'one *can in principle* locate' or 'one *can in practice* locate'? When she says 'one *can* recognise f(x)', does she mean 'one *can in principle* recognise' or 'one *can in practice* recognise'? Of course they cannot be 'in principle', because in that case the explanation is circular.

But when we interpret them as possibility in practice, this approach will not classify finite sets in the second sense as finite. Such a set is so huge that we cannot check every value of a function of it. To check that for every x, f(x) is less than or equal to a specific number, one needs to check as many times as the size of the set. But that is beyond what one can do in practice. Therefore a finite set in the second sense does not count as finite according to this approach. If we call a set infinite just in case it is not finite according to this approach, a finite set in the second sense is classified as infinite, as the strict finitist regards.

2.1.2 The notational approach fails

It is difficult to exactly extract the motivation of the next approach from Wright's writing, but it might be described as follows. We notice that while we say $2^{1000^{1000}} - 1$ is too big for the actual human beings, we feel we surely understand how big this number is, because obviously this number is displayed in front of us by $2^{1000^{1000}} - 1$. We admit that this number is too big to handle with procedures, because it takes some time to process one step and it would take too much time to complete all the $2^{1000^{1000}} - 1$ steps. But can't we handle and understand a number if it is an actually intelligible representation in a notational system? What is finite seems to be dependent on the system of notation. So, by extending the extant notations, we may be able to express any 'finite' number. The second definition of what is finite Wright examines is:

a set [of numbers] is finite just in case, even if only by the introduction of a novel notation, we can actually specify an integral upper bound on its size. [27, p.126]

I do not understand what Wright means by 'an integral upper bound on its size', and instead consider 'an upper bound of the set of numbers at issue'.

For example, this approach seems to be proposing that set $\{n : 0 \le n \le 2^{1000^{1000}} - 1\}$ is finite, because an upper bound of this set can be actually intelligibly represented by $2^{1000^{1000}} - 1$.

several ways of justifying and challenging this point. In fact, one might justify by saying that if there is a function which conforms to this description, then we can find a suitable surjection. And one might challenge by questioning to what kind of set such a function is supposed to map: it should be an initial segment of the set of natural numbers. So this formulation has not successfully avoided using the infinite notion. However, anything about this point does not seem to be crucial, because this line of approach has a fundamental difficulty, as we see.

This numeral is expressed by using the notations of exponentiation and subtraction. The approach asserts that by extending the extant notations, we could extend the numbers that are expressible, and that by doing so, we could express anything the anti-realist considers to be finite. Wright says 'we could write ' m_k^j ', for example, to indicate k reiterations of the exponent j' [27, p.126]^{*12}. By using this, we could denote number $2^{1000_{1000}^{1000}} - 1$. This is bigger than $2^{1000_{1000}^{1000}} - 1 = 2^{1000_{100}^{1000}} - 1$, and is too big to actually write in this thesis in the decimal notation. But this is finite in the second sense.

However, 'No progress is made', says Wright [27, p.126]. There are finite natural numbers in the second sense that cannot intelligibly be represented by an extension of extant notation. As I understand it, Wright's argument here could be reconstructed as follows.

What the proposal asserts is that the following proposition, (n), holds for some natural number n (whether n is finite in the first or the second sense):

(n) there is some feasible extension, E, of extant numerical notations such that integers $\geq n$ are actually intelligibly representable in E. [27, p.126]

(n) says that there is a suitable notation that covers all the numbers above some number. But there is no such a notation, because for any notation, there is a finite number in the second sense that cannot intelligibly be represented in the notation. This is because no matter what notation one uses, the bigger the number, the more complex the representation becomes. For example, m_k^j , may indeed be a notation which gives us an advanced power to denote, but there is an actual limit of expressing using this notation. $2^{1000_{1000}^{1000}}$, is indeed intelligibly expressible, and so is $2^{1000_{1000}^{1000}}_{1000_{1000}}$. But we admit that there is nonetheless a limit to what is expressible, e.g., in this thesis in this notation, and the anti-realist insists that even those numbers beyond this limit are finite and she understands them.

To add to this, Wright points out that there is an actual limit to repeating extending extant notation. 'Simplifications are always feasible', one may think [27, p.126]: we could at any moment arbitrarily introduce a new notation in order to abbreviate what we can actually express. I think that we could stipulate that $\alpha = 2_{1000_{1000}}^{1000_{1000}}$ and write $\alpha_{1000_{1000}}^{1000_{1000}}$ to denote a bigger natural number: this would count finite according to the approach at issue. However, when we introduce too many abbreviations, the entire system of notation becomes too complex to be intelligible. One could say that it requires two steps to recognise the number denoted by $\alpha_{1000_{1000}}^{1000_{1000}}$, because this is abbreviated by the stipulation $\alpha = 2_{1000_{1000}}^{1000_{1000}}$ and it uses the notational extension ' m_k^j '. But we would admit that the actual human being cannot deal with $2^{1000_{1000}} - 1$ steps of abbreviations. Thus (n) does not hold for any n. One should notice that we cannot continue advancing extant notations forever: and when we find that we cannot advance our

^{*12} Wright's original sentence lacks the right quotation mark of ' m_k^j , '.

notation any more, the number which our most advanced notation cannot actually represent, is completely beyond us.

2.1.3 The Dedekindian approach fails

The third approach is that which uses the Dedekindian characterisations of infinity and finitude. In the Dedekindian sense, classically, 'a set is infinite just in case it can be one-to-one mapped onto a proper part of itself, and finite just in case it is not infinite' [27, p.127]. The anti-realist would interpret this as follows: in order to justify that a set is infinite,

we have to have found, or know we can find, a method for pairing the elements of the set with those of a proper subset of it without omission of any of the former or repetition of any of the latter; and in order to be entitled to regard a set as finite, we have to be in a position to deny that any such method can be found. [27, p.127]

Notice that Wright supposes that when interpreting anti-realistically, the existence of a suitable function should be established by our recognition of a method of pairing, and that to deny anti-realistically the existence of something is to recognise that there is no suitable method. Also let us write, assuming we know the definition of function, for that set A is anti-realistically Dedekind-infinite, $\exists B \subsetneq A \exists f : A \to B [\forall b \in B \exists a \in A [f(a) = b] \land \forall a_0, a_1 \in A [f(a_0) = f(a_1) \to a_0 = a_1]]^{*13}$.

This approach has a merit that it does not mention the set of natural numbers. But Wright argues that this approach also misclassifies. Let us see this step by step.

■2.1.3.1 Infinity may be made finite Wright first points out that there is a threat that the Dedekindian approach misclassifies the infinite sets on the anti-realist conception as finite. Here notice that when the anti-realist uses the Dedekindian approach, the anti-realist's original intention is to present infinity and finitude using the notion of possibility *in principle*, but she must replace them with that of possibility *in practice* in order to explain to the strict finitist. Consider those sets that were infinite because one could only *in principle* locate a suitable function. They may now be misclassified as finite because one cannot *in practice* locate such a function – although this is only a *threat*, not an conclusive argument.

Now, it would be useful to have a table to check how the approaches classify sets.

^{*13} The formulation Wright puts for that A is infinite in this sense is the following: $\exists f \exists B [B \subsetneq A \land \forall x \in A \exists y \in B [f(x) = y \land \forall z \in A [f(z) = y \rightarrow z = x]]]$ (I modified into the modern notation) [27, p.127, n.16]. I do not think that this is a good formulation, because f should come after B because f is a relation between A and B, and especially because this lacks the condition that f is an 'onto' mapping.

	First finitude	Second finitude	Infinity
The usual approach	-	Infinite	-
The Dedekindian approach	-	*	Maybe Finite

In this table, I put '-' for the parts which we do not discuss in this thesis. I put words in Italics where the approaches (may) misclassify. We will see the \circledast part shortly.

Although the above is a threat, it is true that we must avoid misclassifying what is infinite for the anti-realist as finite, because if this happened, a 'finite' extension of our capacity for the strict finitist could be an 'infinite' extension for the anti-realist. This case would heavily damage the intention of the explanation scheme (P).

■2.1.3.2 An improvement Wright thinks that one may try to improve the Dedekindian approach by locally strengthening. One may replace the condition for set A to be finite, $\neg \exists B \subsetneq A \exists f : A \rightarrow B \ [\forall b \in B \exists a \in A \ [f(a) = b] \land \forall a_0, a_1 \in A \ [f(a_0) = f(a_1) \rightarrow a_0 = a_1]],$ with $\forall B \subsetneq A \forall f : A \rightarrow B \ [\exists b \in B \forall a \in A \ [f(a) \neq b] \lor \exists a_0, a_1 \in A \ [f(a_0) = f(a_1) \land a_0 \neq a_1]].$ Notice that the latter intuitionistically implies the former, but not vice versa. On this strengthened condition, in order to say that a set is finite, one needs to be able in practice to locate, for any subset of the set and for any putative mapping, omissions or repetitions of the value of the mapping^{*14}. With this renewal, the threat that the misclassification of anti-realistically infinite sets as finite will be vanished. Whether an anti-realistically infinite set is classified as finite, but it is now unknown whether all the anti-realistically infinite sets are misclassified as finite, but it is now unknown whether all the anti-realistically infinite sets are misclassified as finite.

Thus the table is updated.

	First finitude	Second finitude	Infinity
The Dedekindian approach	-	*	(Unknown)

Let us next discuss the \circledast part.

 \blacksquare 2.1.3.3 Finitude is made 'not finite' Wright argues that the finite sets in the second sense are not going to be classified as finite. As I understand his writing, at least without adopting the 'improvement', it is very difficult to defeat a statement that a set is infinite. The example

^{*14} As I mentioned in the preceding footnote, Wright's formulation of the condition for a set to be finite differs from mine [27, p.127, n.16]. But the conceptual description of the condition that one needs to be able in practice to locate omissions or repetitions to say a set is finite, is shared by Wright and me.

Wright presents us with is set $A = \{n : n \leq 100^{100}\}$ and set $B = \{n : n \text{ is a prime number } \land n \leq 100^{100}\}$. Suppose one claimed that there is a one-to-one mapping from A onto B. It seems that we cannot defeat this claim, because it is needed in general to enumerate the relevant primes to do so: to spot the fact that the putative mapping is not 'one-to-one', or the fact that it is not an 'onto' mapping, we need to have the complete list of the members of B. But this is only in principle possible, not in practice. This means that we cannot negate the supposition that $\exists B \subsetneq A \exists f : A \to B \ [\forall b \in B \exists a \in A \ [f(a) = b] \land \forall a_0, a_1 \in A \ [f(a_0) = f(a_1) \to a_0 = a_1]].$ Therefore we cannot establish that $\{n : n \leq 100^{100}\}$ is finite.

	First finitude	Second finitude	Infinity
The Dedekindian approach	-	Not finite	(Unknown)

2.1.4 The Dedekindian approach is harmful

One of the striking parts of Wright's 'Strict Finitism' is where he argues that the Dedekind approach misclassifies a finite set in the second sense as infinite [27, Sect.6]. We have so far seen that this approach does not classify this kind of set as finite, but he argues that the situation is worse. His argument uses the notion of a 'surveyability predicate'. This is a characteristic notion to strict finitism, but we suppose that the anti-realist also (of course) grasps what this is. We will first see the definition and the features of this kind of predicate (2.1.4.1), although the justification that a surveyability predicate indeed has such features shall be postponed to the 'positive programme' of strict finitism (3.1, especially 3.1.2). Then, we will see the argument for the 'harm' of the Dedekindian approach (2.1.4.2).

■2.1.4.1 Surveyability predicates A surveyability predicate is a predicate of practical intellectual possibility. For example, predicates such as 'x is a natural number actually intelligibly representable in decimal notation' (let us write B for this) and 'X is a proof structure actually intelligibly physically written down' are surveyability predicates. The strict finitist supposes that a surveyability predicate (1) is tolerant, (2) is not borderline-case vague (non-vague for short) and (3) is actually weakly decidable.

A predicate P is said to be *tolerant* just in case 'sufficiently small variations in some associated parameter are apparently insufficient to affect the justice with which it can be applied to something, whereas sufficiently large variations are always so sufficient' [27, p.108]. To rephrase, if an object, a, satisfies a tolerant predicate P, then any object similar enough to a, in the relevant sense, also satisfies P, but an object different enough from a in the relevant sense does not. For example, we admit that natural number 0 can in practice be intelligibly represented in the decimal notation (hence B(0)), so can 1 be, and so can 2 be; but we do not think that $2^{1000^{1000}} - 1$ can be (hence $\neg B(2^{1000^{1000}} - 1)$). To formalise, $B(n) \rightarrow B(S(n))$ and $B(S(n)) \rightarrow B(n)$ hold good for any small enough number n (where S is the successor function), but B(m) should be negated for any big enough number m.

A predicate P (for natural numbers, for example) is said to be *borderline-case vague* just in case there is an integer n of which we could actually recognise that neither P(n) nor $\neg P(n)$ could definitely correctly be asserted [27, p.132]. An example of a borderline-case vague predicate (outside the context of mathematics) is a predicate of colour: we seem to admit that there is something, a, such that we have justification to assert neither that a is green, nor that a is not green. The strict finitist supposes that surveyability predicates are not vague in this sense.

A predicate P is actually weakly decidable just in case for any object x, there is a humanly feasible programme of investigation whose implementation is bound to produce at least ground for asserting either that we can P(x) or that we cannot P(x) [27, p.133]. What should be noted is that Wright regards the assertibility of a statement which appears here as something weaker than the verifiability. The notion of assertibility is introduced as a weaker notion than verification: when we possess a verification of a statement, the statement is of course assertible; but that we possess reason to assert a statement does necessarily mean that we possess a verification^{*15}. The strict finitist supposes that we can in practice know whether a surveyability predicate is assertible or the negation of it is assertible.

■2.1.4.2 Finitude is made infinite On my reconstruction, Wright's argument is that the extension, Σ , of predicate 'we can in practice recognise an nde of x' is infinite according to the Dedekindian account [27, p.129]. Notice that this predicate is a surveyability predicate.

The argument starts with supposing that $\Sigma = \{n : n \leq 1000000^{1000000}\}$. Also, we let A(n) mean that n is actually intelligibly representable in some extant notation; and let $K = \{n : A(n) \land n \leq 1000000^{1000000}\}$. Let B(n) mean that n is actually intelligibly representable in the decimal notation. Then, K is partitioned into $U_1 = \{n : n \in K \land B(n)\}$ and $U_2 = \{n : n \in K \land \neg B(n)\}$, because B is actually weakly decidable: it is another assumption that the actually weak decidability of B is enough for this partitioning.

Define $L = K \setminus \{1000000^{1000000}\}$; and define relation $R \subseteq K \times L$ by

 $Rnm \iff n \in U_1 \wedge m$ is the successor of nor $n \in U_2 \wedge m$ is the immediate predecessor of n.

^{*15} However, it seems in the end that the strict finitist (or the 'theorist' of strict finitism) can be content if she can use the notion of assertibility, even if she cannot use that of verifiability in practice. For example, it suffices to suppose predicate B is actually weakly decidable in order to establish the 'harm' of the Dedekindian approach (2.1.4.2), and the semantic theory which we will see at the end of the thesis treats as if assertibility and verifiability in practice were identical (see 4.2 and 4.3, especially 4.3.1.2).

We claim that R is a one-to-one mapping from K onto L. To establish this, Wright's argument tries to show (1) that $\forall n \in K \exists m \in L [Rnm \land \neg \exists o \in L [Rno \land o \neq m]]$ and (2) that $\forall m \in L \exists n \in K [Rnm \land \neg \exists o \in K [Rom \land o \neq n]]$ [27, p.129].

But I must here insert that it is not clear to me why these jointly imply that R is a one-to-one mapping from K onto L. But at least, if we admit that the double negation elimination about the identity of numbers, i.e., $\forall n_0, n_1 \in K [\neg(n_0 \neq n_1) \rightarrow n_0 = n_1]$ and $\forall m_0, m_1 \in L [\neg(m_0 \neq m_1) \rightarrow m_0 = m_1]$, then (1) and (2) do imply^{*16}. So let us also suppose them.

I reconstruct the arguments for (1) and (2) as follows. [27, pp.130-1]

The argument for (1). Suppose we can actually recognise $n \in K$. Since K is partitioned into U_1 and U_2 , either $n \in U_1$ or $n \in U_2$.

If $n \in U_1$, then B(n) tells us that we shall be able to construct an actually intelligible expression in the decimal notation for n. Since B is tolerant, we should be able to do the same for the successor of n, and hence it is actually recognisable. So take it as m. Then Rnm. Suppose for contradiction that we can locate an $o \in L$ such that Rno actually recognisably holds but o is actually recognisably different from m. Then the definition of R and that $n \in U_1$ tells us that o is the successor of n. This is contradictory to the uniqueness of the successor.

If $n \in U_2$, then A(n) tells us that we shall be able to construct an actually intelligible expression for n in some extant notation (though not in the decimal notation). Since A is tolerant, we should be able to do the same for the immediate predecessor of n, and hence it is actually recognisable. (It would be expressed as $\overline{n} - 1$.) So take it as m. Then Rnm. Suppose for contradiction that we can locate $o \in L$ such that $Rno \wedge o \neq m$. Then the definition of Rand that $n \in U_2$ tells us that o is the immediate predecessor of n. This is contradictory to the uniqueness of the immediate predecessor.

The argument for (2). Suppose we can actually recognise $m \in L$. Since B is actually

Next, the surjectivity is obvious from (2).

^{*16} In this footnote, I prove that (1) and (2) with the double negation elimination imply that R is a one-to-one mapping from K onto L.

I prove first that R is a function, i.e., $\forall n \in K \exists m \in L [Rnm \land \forall o \in L [Rno \to m = o]]$. Consider $n \in K$. Then (1) gives us $m \in L$ with that $Rnm \land \neg \exists o \in L [Rno \land o \neq m]$. So it suffices to show that $\forall o \in L [Rno \to m = o]$. Consider $o \in L$, and suppose Rno. If $m \neq o$, then it holds that $\exists o \in L [Rno \land o \neq m]$, which is contradictory. Therefore $\neg(m \neq o)$. Since we are supposing the double negation elimination, it holds that m = o.

Finally I show that R is injective, i.e., $\forall n_0, n_1 \in K [\exists m \in L [Rn_0m \land Rn_1m] \rightarrow n_0 = n_1]$. Consider $n_0, n_1 \in K$, and suppose that we have $m \in L$ with $Rn_0m \land Rn_1m$. Then (2) gives us that we have $n' \in K$ with $Rn'm \land \neg \exists o \in K [Rom \land o \neq n']$. Here suppose that $n_0 \neq n'$. Then $Rn_0m \land n_0 \neq n'$ tells us that $\exists o \in K [Rom \land o \neq n']$. A contradiction. Therefore $\neg(n_0 \neq n')$. Similarly, the supposition that $n_1 \neq n'$ leads us to a contradiction. Therefore $\neg(n_1 \neq n')$. By the double negation elimination, $n_0 = n' = n_1$.

weakly decidable, either B(m) or $\neg B(m)$. Therefore $m \in U_1$ or $m \in U_2$.

If $m \in U_1$, then B(m) tells us that we shall be able to construct an actually intelligible expression in the decimal notation for m. Since B is tolerant, we should be able to do the same for the immediate successor of m, and hence it is actually recognisable. So take it as n. Now B(n), and hence $n \in U_1$. Therefore Rnm. Suppose for contradiction that we can locate an o such that $o \in K$ and Rom actually recognisably holds but o is actually recognisably different from n. B(m) and the tolerance of B tells us that B(o), because o is the successor or the immediate predecessor of m. Hence $o \in U_1$. However, since $n \in U_1$, this is contradictory.

If $m \in U_2$, then A(m) tells us that we shall be able to construct an actually intelligible expression in some extant notation for m. Since A is tolerant, we should be able to do the same for the successor of m, and hence it is actually recognisable. (It would be expressed as $\overline{m} + 1$). So take it as n. Now A(n). But $\neg B(n)$ holds, because if B(n), then the tolerance of B tells us that B(m), implying that $m \in U_1$ contradictorily. Therefore $n \in U_2$. Therefore Rnm. Suppose for contradiction that we can locate an $o \in K$ such that $Rom \land o \neq n$. Notice that either B(o) or $\neg B(o)$. Therefore either $o \in U_1$ or $o \in U_2$. If $o \in U_1$ holds, then B(o) and the tolerance of B tells us that B(m). Therefore $m \in U_1$, which is contradictory to $m \in U_2$. If $o \in U_2$ holds, then m is the immediate predecessor of o and that of n at the same time, but $o \neq n$. This is contradictory to the uniqueness of the successor.

Thus, it is shown that there is a one-to-one mapping from K onto a subset of it: the set of the numbers that we can in practice intelligibly represent in some extant notation is, according to the Dedekindian approach, infinite. The anti-realist takes this set as finite. It may be controversial whether this is finite in the first sense or the second, but it is anyway not acceptable for the anti-realist that a set she considers to be finite is misclassified as infinite.

2.1.5 The conclusion so far

We have been investigating how the anti-realist could explain her notion of finitude to the strict finitist. A clear explanation of finitude is needed in order to explain the notion of decidability in principle, using (P). But none of the three approaches for finitude was successful. The explanations inevitably involve appeal to possibility in principle (or otherwise the two of them suffer misclassification). The general notion of possibility in principle is called upon in order to explain the special notion of decidability in principle, but when the special notion is not clear, the general notion cannot be. Thus we conclude that the anti-realist cannot explain how to manifest her understanding of the notion of decidability in principle. The anti-realist notions of decidability in principle and finitude are dependent on each other, and hence those who do not understand either of them, such as the strict finitist, cannot attain the understanding of any of them. The anti-realist's key notions are interlocked. One may wonder whether (P) is the only explanation scheme. If there is a different way of analysing the notion of decidability, the series of the arguments so far might lose its impact. Wright thinks about this possibility and has presented the following ideas. In fact, our arguments have been focusing on the difficulty in explaining finitude. Therefore, this series of arguments will not apply when someone comes up with an explanation of decidability in principle which does not appeal to finitude. However, this does not seem likely to happen, because when the anti-realist wishes to characterise the arithmetical statements for which the principle of bivalence holds, she seems to inevitably appeal to the notion of finitude. Wright writes: '[A]ll quantifier- and variable-free arithmetical statements are to go on one side of the distinction, while those which quantify unrestrictedly over all the positive integers go on the other' [27, p.135]. The former statements are decidable (in practice or in principle), while the latter are undecidable. As I understand this, when one tries to make the distinction between the former and the latter, she will characterise the former as 'finite', and the latter as 'not finite'. It seems that the notion of decidability in principle is essentially related to finitude. [27, p.135]

2.2 No loophole for arithmetical anti-realism

Let us say that the conclusion so far is correct, and that, therefore, general anti-realism is untenable. Then the anti-realist has to be a local viewpoint, i.e., a view about a single aspect of the world. The aim of this section is to examine especially whether the anti-realism about arithmetic can be defended, and to conclude that it cannot be.

One may wonder what the difference between general anti-realism and arithmetical antirealism is. General anti-realism is a view that we should take an anti-realist position in every area of dispute, and the notion which can uniformly characterise all the anti-realist positions is decidability in principle. Arithmetical anti-realism is the anti-realist position in the area of the arithmetical statements, and if decidability in principle and finitude were apprehensible, the strict finitist would understand this position based on the notion of decidability in practice, which she already grasps. But '[e]ven if 'decidability in principle' is rejected as mythical, any standpoint can still be reasonably seen as intuitionist [i.e. anti-realist, in our terminology] which [...] opposes a realist view of a certain type of statement while continuing to hold that the acceptability of the principle of Bivalence nowhere requires a guarantee of actual decidability' [27, p.136]. Namely, in an individual case, the anti-realist position can be characterised as a specific view about the acceptability of the principle of bivalence. One could say that decidability in principle is the notion with which we can specify, in every area, the statements for which the principle of bivalence is admitted to hold. If this notion were available, the arithmetical anti-realist would say that the principle of bivalence holds only for the arithmetical statements decidable in principle. But in the area of arithmetic at least, there is another way of specifying such statements.

The anti-realist takes the position that the principle of bivalence holds for, and only for, the statements containing no unrestricted quantifiers^{*17}. For example, she would think that (1) $2^{1000^{1000}} - 1$ is a prime number', and also (2) $\forall n < 2^{1000^{1000}} - 1 [n < k_{99}]$ ' follow the principle, because they contain no unrestricted quantifiers^{*18}. Notice that, on the one hand, arithmetical realism is the view that the principle holds for any statement; and that, on the other, arithmetical strict finitism is the view that the principle holds only for the finite statements in the first sense. Here, it would be helpful to have a table of the stances towards the principle of bivalence.

	Strict finitism	Anti-realism	Realism
With quantifiers	for some only	for some only	for all
Without quantifiers	for some only	for all	for all

I sorted out the stances according to what they say about the statements with or without quantifiers. Now it is clear that, for the anti-realist, the statements for which the principle of bivalence holds are those decidable in principle, and in other words they are those containing no unrestricted quantifiers. So, one could say that while we do not have (yet) a definite characterisation of the first finitude, the anti-realist notion of finitude which include the first and the second finitude can be characterised as 'containing no unrestricted quantifiers'.

To establish arithmetical anti-realism is to complete the following tasks: (1) to defeat realism by establishing that the principle of bivalence does not hold for the arithmetical statements with quantifiers; and (2) to defend itself by establishing that the principle of bivalence holds for the statements without unrestricted quantifiers (the part in the table I underlined). In this thesis, as Wright does, we focus only on (2) [27, Sects.8-9.]. First, we will see with what reason the anti-realist thinks the principle of bivalence holds for those statements, and then I try to give an objection to it.

The anti-realist about arithmetic is now going to establish the bivalence of the statements containing no unrestricted quantifiers. The concepts she can appeal to will be those which the strict finitist apprehends. So, notice that if the following anti-realist ideas are correct, then even the strict finitist should admit the bivalence. The strict finitist must present an

^{*17} Wright here uses 'the statements containing no quantifiers' instead of 'the statements containing no unrestricted quantifiers' [27, p.136]. But it seems that the anti-realist would accept the principle of bivalence for the statements which quantify restrictedly. This is because such statements could (in principle) be written without any quantifiers: e.g. ' $\forall i \in \{0, 1, 2\} [A_i]$ ' is ' $A_0 \wedge A_1 \wedge A_2$ '. In fact, Wright uses 'whose which quantify unrestrictedly over all the positive integers' where I quoted before [27, p.135].

^{*&}lt;sup>18</sup> For the definition of $n < k_{99}$, see section 1.1.4.

argument which refutes or avoids the anti-realist's conclusion: we will consider what reply the strict finitist can provide (2.2.2).

2.2.1 The anti-realist's argument for the bivalence

The anti-realist's strategy for justifying the principle of bivalence, which Wright examines, can be summarised as follows: (1) the anti-realist proposes a way of interpreting that a statement holds, (2) she presents two key statements, (C) and (D), and (3) she justifies the transition from (C) to (D). (D) is the statement which represents the principle of bivalence under the interpretation proposed in (1). Let us see them in order.

■2.2.1.1 The 'implementation' interpretation As we saw, for the anti-realist, to say that a statement holds is to say that it is verifiable in principle. But as long as we possess a decision procedure for a statement, the condition under which the statement holds good can be interpreted without appealing to verifiability in principle. 'Let ϕ be [...] a full description (= set of instructions) for the decision procedure appropriate to a particular arithmetical statement *S* containing neither quantifiers nor free variables' [27, p.138]. Then, the anti-realist proposes,

[...] S is equivalent to: 'if we ever come to be able surveyably to implement the procedure described in ϕ , we shall be able to effect a verification of S', and that $\sim S$ is equivalent to: 'if we ever come to be able surveyably to implement the procedure described in ϕ , we shall be able to effect a falsification of S' [...] [27, p.139]

To simplify, when we have a decision procedure ϕ for an arithmetical statement S, we can equate that S holds good with that a suitable implementation of ϕ yields an affirmative result, and that S does not hold good with that a suitable implementation of ϕ yields a negative result. The affirmative result amounts to a verification of S, and the negative result a falsification of S.^{*19}

We think that the strict finitist accepts this interpretation, because we assume as follows. I quote Wright's phrase:

It is assumed [...] that it is contingent and, in a large class of cases, comes as a straightforward empirical discovery what our practical computational limitations are, so we can make straightforward sense of the idea of extensions in the relevant capacities; 'a humanly surveyable implementation of the procedure described in ϕ ' can thus be a description that we perfectly well understand but which is, quite contingently, true of nothing. [27, p.146]

 $^{^{\}ast 19}\,$ I mentioned essentially the same idea in section 1.1.3.

Neither the strict finitist, nor we, know exactly the limitations of our practical intelligent abilities. To make it clear is a matter of empirical discovery. When we try to implement a procedure, sometimes we may complete it, but sometimes it may as well turn out to be too long for us to complete. Whether a given procedure is too long for us is, in some cases, not known to us before the implementation: it is discovered only after implementing it. The point is that in some cases, we nonetheless know that a procedure is appropriate to a statement, even before we try implementing it. When we try the procedure and it turns out that it is too long for us, we say 'this is beyond our limitations: if our abilities were more extensive, we could complete this'. We understand the idea of extensions in our abilities. In other words, we understand the idea of 'the implementation of a procedure which is completed in a humanly feasible way', even without trying implementing the procedure, and even after it turned out that the procedure was too long for us. This is how I understand the above assumption.

When we admit this, it is plausible to say that the strict finitist should grasp the idea of the suitable implementation of a procedure. So we think that even she accepts the interpretation of a statement's holding good in terms of the suitable implementation.

■2.2.1.2 Two statements (C) and (D) Similar to the above, the possession of a decision procedure (or the set of the instructions of a decision procedure) ϕ for a statement S is interpreted as follows:

(C) If we ever come to be able surveyably to implement the procedure described in ϕ , we shall be able to effect a verification either of S or of its negation. [27, p.138]

After agreeing to the above assumption (2.2.1.1), the strict finitist should not be able to object the significance of (C). This seems to be a natural reformulation of being in possession of a decision procedure. If a procedure is implemented properly, it should yield an affirmative result or a negative result. As I put, an affirmative result amounts to a verification and a negative one to a falsification. Then a proper implementation must yield a verification of a statement at issue, or a verification of the negation of it.

The anti-realist belief in the principle of bivalence for S is simply an affirmation of the disjunction [27, p.139]:

(D) It is the case *either* that if we ever come to be able surveyably to implement the procedure described in ϕ , we shall be able to effect a verification of *S* or that if we ever come to be able surveyably to implement the procedure described in ϕ , we shall be able to effect a falsification of *S*. [27, p.139]

Again, the strict finitist must grasp this statement.

What is interesting is that (C) and (D) are not necessarily equivalent. This may become clearer if we write them as

- (C) $\phi \to (S \lor \neg S)$
- (D) $(\phi \to S) \lor (\phi \to \neg S).$

Surely, these ϕ and S are not meant to be statements: rather ϕ stands for 'the proper implementation of the decision procedure', and S a 'realisation of a state of affairs'. But for the time being, let us use this notation to represent the exact dependency of them. The point is that it is well-known that $(P \to (Q \lor R)) \to (P \to Q) \lor (P \to R)$ is not valid in intuitionistic logic. A countermodel for this is the following.



Let the valuation function V for this frame be such that:

- $V(P) = \{1, 2\}$
- $V(Q) = \{1\}$
- $V(R) = \{2\}.$

Then $0 \models P \rightarrow (Q \lor R)$ and $0 \not\models (P \rightarrow Q) \lor (P \rightarrow R)$.

Wright touches upon why this transition does not hold, by mentioning Dummett's example [27, p.141]. Dummett's contention is that this does not hold when it is indeterministic which disjunct holds: e.g. in a case where the antecedent of (C) is always accompanied by auxiliary situations and they are the factors that determine which disjunct holds. Dummett's example is a subjunctive conditional about Castro and Carter:

For instance, we may safely agree that, if Fidel Castro were to meet President Carter, he would either insult him or speak politely to him; but it might not be determinately true, of either of those things, that he would do it, since it might depend upon some so far unspecified further condition, such as whether the meeting took place in Cuba or outside. $[6, pp.244-5]^{*20}$

Let us call 'Fidel Castro is to meet President Carter', P, 'Fidel Castro insults President Carter', Q, and 'Fidel Castro speaks politely to Carter', R. Even if subjunctive conditional

^{*20} By the way, in Wright's writing, the president of the United States of America in the example is replaced with 'Reagan', as the result of 'adapt'ing 'Dummett's own example' [27, p.141]. This is not a substantial matter, but in fact, when 'Strict Finitism' was published (1982), the President was Reagan (1981-9). When Truth and Other Enigmas was published (1978), the President was Carter (1977-81). I write, however, as if Castro were alive and Carter were the present President.

 $P \to (Q \lor R)$ is correct, we cannot infer from this that $P \to Q \lor P \to R$. This is because where they met may affect Castro's behaviours: if the meeting took place in Cuba, then Castro might behave rudely, but if it did in the U.S., he might behave politely. P alone may be neither enough to conclude Q, nor enough to conclude R: therefore it may the case that neither $P \to Q$, nor $P \to R$ holds. But if the meeting were going to happen, then the place of the meeting must be determined anyway: therefore P allows us to conclude $Q \lor R$.

It may be useful to look at the above countermodel for the transition from (C) to (D). If it is meaningful to take a model as representing the possible situations about the politics in the real world, this model represents that the fact that the meeting takes place (P) is not the decisive factor for whether Castro insults (Q) or speaks politely (R).

Before moving to the anti-realist idea for justifying the transition from (C) to (D), we should notice that the strict finitist should suspend her reading of disjunction for the time being. (D) allows us to assert a disjunction without knowing which disjunct holds. If the strict finitist were stubborn about her reading and is convinced with that the assertion of a disjunction must be done on the basis of knowing which disjunct holds, then (D) never holds good. This stance amounts to question begging. Therefore, the strict finitist has to examine the anti-realist's idea without supposing such a reading of disjunction. [27, pp.139-40]

■2.2.1.3 The transition from (C) to (D) Dummett points out that even when (C) holds, (D) does not, if there is indeterminacy to which disjunct holds. The anti-realist thinks that, then, if there is no indeterminacy, the transition must hold^{*21}. The anti-realist idea, which Wright considers, is to specify the conditions under which the transition from (C) to (D) is acceptable, and assert that the arithmetical cases satisfy them.

Here, it would be helpful for us to concentrate on the arithmetical case: I summarise and reconstruct the conditions in the context of arithmetic. Suppose (C) holds.

- (i) If ϕ were to be properly implemented, there would be no indeterminism about whether a verification or a falsification of S is effected;
- (ii) There are no additional factors, ψ and θ , (1) which we are not aware of, (2) which may accompany a proper implementation of ϕ , and (3) which satisfy $(\phi \land \psi) \rightarrow S$ and $(\phi \land \theta) \rightarrow \neg S$; and
- (iii) It is not the case that, whether we properly implemented ϕ or not, there are grounds for the consequent of (C), $S \vee \neg S$. [27, pp.142-3]^{*22}

 $^{^{*21}}$ In fact, Dummett himself explicitly took this view later [13, p.349].

^{*22} Wright elaborates his thoughts for the cases of 'future indicative', although it is not clear to me whether they are really specific to future indicative [27, p.142]. He does so by criticising and advancing Dummett's idea [27, pp.142-3] [6, p.245]. I put here the 'future indicative' version of my summary.

⁽i) When P is realised, there is no indeterminism about which one of Q and R will be realised;

(D) seems to hold good when these conditions are met. I think Wright might explain in the context of arithmetic as follows. (i) guarantees that it is fully determined which disjunct holds when ϕ is properly implemented. (ii) tells us that the determination is not done by any outer factors whose realisation we do not expect to accompany the implementation of ϕ . Therefore it seems inevitable that *either* a sufficient condition of Q or that of R is collectively captured by what we expect to obtain at the same time as a proper implementation of ϕ . (iii) now says that we are grasping *either* a sufficient condition of Q or that of R by nothing but the fact that we could properly implement ϕ . [27, pp.143-4]

Then, let us see how the anti-realist could argue that these conditions are met in the arithmetical cases. First, the explanation which Wright gives for (iii) is rather simple: 'there is, or so we may suppose, no prospect of our getting into a position where we may actually verify S or may verify its negation unless we come to be able surveyably to implement the procedure described in ϕ ' [27, pp.143-4]. This essentially states that we do not think that we may actually verify S or its negation unless we properly implement ϕ . I think that this is correct in the most relevant arithmetic cases, namely, when S is undecided, and we have only one decision procedure for S.

As Wright puts it, what are striking are why (i) and (ii) seem to be met. They hold good, Wright says on behalf of the anti-realist, because it is a function *purely and wholly* of the feature of ϕ which disjunct would be verified. The result of a proper implementation of ϕ could not vary in accordance with other accompanying circumstances: the outcome of computation cannot vary. There is, so to speak, an *internal relation* between the feature of proper implementation of the procedure and the outcome which proper implementation would secure. The correct outcome is predetermined. [27, p.144]

Wright supposes that the anti-realist may think (D) is motivated by the above ground. He characterises this line of thought as conferring objectivity on meaning. For the antirealist, once a decision procedure is fixed, the result of a proper implementation of it is fixed even before an implementation is really properly done. One does not need to check whether an ongoing procedure conforms to ϕ or not. Even without human contribution, a decision procedure is determinately associated with either a verification or a falsification of a

⁽ii) There are no circumstances, T and U, such that (1) we do not expect that they accompany a realisation of P and (2) $(P \wedge T) \rightarrow Q$ and $(P \wedge U) \rightarrow R$; and

⁽iii) The consequent of (C), $Q \vee R$, holds only because P holds: in other words, it is P that brings about $Q \vee R$.

My understanding of (ii) is as follows. A circumstance V such that $(P \wedge V) \rightarrow Q$ is a circumstance which is, jointly with P, enough to realise Q. I think that an example of such a V is 'Castro and Carter's meeting takes place in Cuba', and to say that one does not expect that such a V accompanies a realisation of P is to say that she does not notice that there is such a V. If one is aware of such a V, she should incorporate it and consider conjunction $P \wedge V$ as a new P. (ii) guarantees that we have already taken into account all such circumstances.

statement. Wright also says that this is to believe in 'the idea that a sign of any kind can be associated with a meaning [...] in such a way that certain particular uses of it [...] simply do not [...] cohere with its being so associated' [27, p.145]. To me, he seems to mean that for the anti-realist, once a sentence is fixed and a decision procedure is discovered, it is already determined whether it or its negation is correct.

2.2.2 The strict finitist's reply

The anti-realist argument we saw can be a threat to the strict finitist, because it states that arithmetical correctness outruns our actual capacities for deciding: even when we give up deciding whether a statement is correct or not, according to this argument, it is already determined. The notion of arithmetical correctness for the strict finitist might be just like the notion of truth for the anti-realist. It seems that if the strict finitist admits the consideration for the motivation of (D), then she must accept this view.

Wright in 'Strict Finitism' focuses on the objectivity of meaning in which the anti-realist believes, when he examines how this view can be rejected. He thinks that a way of rejecting can be found in Kripke's idea about Wittgenstein's discussion in *Philosophical Investigation* [16] [25]. This is the topic which is now known as 'rule-following'. The basic idea is that it is we who fix the rules of language use; therefore there is no determined meaning 'settled in advance and independently of any investigation we might make' [27, p.148]. But Wright does not wish to be conclusive: after summarising his own views, he writes that '[t]hese are extremely difficult issues', and that 'my intention here has only been to locate an issue' [27, p.151]^{*23}.

I do not think that the strict finitist has to appeal to the concept of rule-following to resist the anti-realist idea. It is because, for the strict finitist, (1) not all the statements for the antirealist are intelligible statements, and (2) not all the decision procedures for the anti-realist are intelligible procedures. The anti-realist wants to say that the principle of bivalence holds for all the statements containing no unrestricted quantifiers, if we have decision procedures. In fact, the strict finitist could admit this, *only superficially*: I think that she could admit that once a given expression is an *intelligible* statement with the verification-condition and an *intelligible* decision procedure is given, the statement follows the principle of bivalence. However, it does not seem that the strict finitist has to fully accept the anti-realist's intention: some anti-realistically legitimate combinations of a statement and a decision procedure would be excluded.

^{*23} According to him, what is written on pages 149 an 150 of 'Strict Finitism' is a summary of 'the argument of the concluding section of chapter XI of' his *Wittgenstein on the Foundations of Mathematics* [27, p.150, n.31]. This book is [26]. He also mentions his paper 'Rule-following, Meaning and Constructivism' as an improvement of the views. This paper is [28].
I think that the unintelligible statements for the strict finitist may include $2^{1000^{1000}} - 1$ is a prime number' or $Pr(2^{1000^{1000}} - 1)$ '. She may reject this as a statement in the first place, based on the ground that expression $2^{1000^{1000}} - 1$ ' is unintelligible. This will happen when the strict finitist sticks to the decimal notation. Also a statement would be unintelligible when it is too long: e.g., a statement formed out of $2^{1000^{1000}}$ conjuncts would be unintelligible. Such a statement could be written as $\bigwedge_{n < 2^{1000^{1000}} - 1} [A_n]$ or $\forall n < 2^{1000^{1000}} - 1 [A_n]$, but it seems that

one needs to appeal to an 'advanced notation' (for the strict finitist) to write short.

My opinion about intelligibility of a decision procedure is, however, more complicated. The strict finitist may have to resist and concede the anti-realist idea at the same time. A decision procedure, I think, will be unintelligible when it is too complex for the actual human beings; therefore even if an intelligible statement is given, the principle of bivalence would not hold for it if the procedure is, for example, too long. But the strict finitist would admit that the sieve of Eratosthenes is simple enough and therefore an intelligible procedure. So a procedure which requires repeating it 1,000,000,000 times would be intelligible. Therefore the principle of bivalence would hold for a statement with such a decision procedure, no matter how impossible in practice it seems to decide whether the result is affirmative or negative.

The point is that what is an intelligible combination of a statement and a decision procedure seems to outrun what is decidable in practice. I in this thesis accept this stance: the principle of bivalence holds for an intelligible arithmetical statement, S, containing no unrestricted quantifiers, if we have an intelligible decision procedure for it – even if we cannot in practice know whether S or its negation holds. However, the restriction of intelligibility which the strict finitist imposes is not light: to repeat, one should notice that even when the strict finitist accepts this stance, her position is never the same as the anti-realist's.

3. The positive programme (1): The basis of strict finitism

In this section, I try to give a support to the basis of strict finitism. As I foreshadowed in section 1.2.3, strict finitism has been accused of being inconsistent: it is said that a surveyability predicate, which is characteristic to strict finitism, is susceptible to a kind of paradox. The paradox is called 'Wang's paradox' by Dummett [7]. This is the main topic of this section – we will reconstruct and try to support Wright's solution to it. But to do so, we need to clearly understand the basic concepts of strict finitism. The reader would remember that we simply defined and assumed several features of a surveyability predicate, when we spoke of the harm of the Dedekindian approach (2.1.4.1). I will first explain the basic concepts and defend the assumptions (3.1), and then try to establish that the strict finitist can avoid the paradox (3.2).

3.1 The basic concepts

We have already spoken about the basic stance and concepts of strict finitism several times (1.2.2, 2.1.4.1, etc.). We will continue to use them. But we need more of them in order to assess what the strict finitist asserts based on her position.

3.1.1 Verifiability in practice

When we attacked general anti-realism, we used an explanation scheme named (P) (in 2.1). We demanded of the anti-realist an explanation of the notion of decidability in principle using finitude and decidability in practice. But we did not doubt the notion of decidability in practice, because it is, for the strict finitist, an accepted notion. But outside of the context where the strict finitist attacks, it may be needed to explicate this notion, since verifiability in practice is the central notion of strict finitism.

What we will see in this section is not the full explanation of this notion. Wright gives an analysis of 'actual verification' [27, Sect.3]; and surely this is not the same notion as verifiability in practice, since Wright's target does not include the modality of 'possibility'. But it would be helpful to look at this, because it gives us a more detailed explanation of the implementation interpretation of 'holding good'.

After a series of consideration, Wright defines the 'I-class' for a statement:

Let S be any statement for which there is an investigative procedure, I; and let the I-class of statements for S contain every and only statements R satisfying the following conditions:

a any rational agent who considers that the upshot of performing I is a justified belief

in S commits himself to believing R;

- b one admissible explanation of the agent's coming falsely to believe S on the basis of I would be the falsity of R; and
- c S does not entail R nor vice versa. [27, p.119, original emphasis]

To simplify, we could understand clause (a) as saying that a decision procedure, I, for a statement, S, is associated with a set of statements called '*I*-class'; that I is divided into several parts (or steps); and that a member R in the set stands for the belief that a part of I has been implemented correctly. So, to believe that I has been implemented correctly is to believe all the statements of the set. Clause (b) intuitively means that, because I is a decision procedure for S, if one is wrong in believing that S is correct, then it means that she is wrong in believing in at least one member of the set. For example, if I is an attempt to construct a mathematical proof, an R is such that its negation would stand for the existence of errors, such as oversight, illusion or misunderstanding. Clause (c) is added in order to exclude irrelevant cases. A statement can be now said to be significant just in case it is 'potentially associated with a non-empty I-class': notice, however, that 'potentially' is added, because there is no guarantee that a decision procedure must emerge. [27, p.119]

Wright's analysis of verifiability in practice and decidability in practice is as follow:

S is capable of *actual verification* if and only if there is some investigative procedure, I, such that (i) we can actually implement I and, on that basis, achieve, if we are rational, what we will consider to be a grounded belief in S; but (ii) subsequent grounds sufficient to call S into question would, if we are rational, have to be allowed to call into question simultaneously the truth of at least one member of the original I-class. (An undecided S may thus be regarded as *decidable in practice* just in case we have recognized that either it or its negation is capable of actual verification.) [27, p.119, original emphasis]

While clause (i) essentially corresponds to the implementation interpretation, clause (ii) states a requirement for a decision procedure. After S is admitted as holding good (or failing), we needs to doubt the implementation of the procedure whenever she doubts that S really holds (fails). One could say that this feature bolsters the view that a statement must be established based on a cognitive process.

3.1.2 The features of a surveyability predicate

We assumed in 2.1.4.1 that a surveyability predicate is (1) tolerant, (2) non-vague, and (3) actually weakly decidable. Now we try to justify each of them by examining Wright's ideas.

\blacksquare3.1.2.1 The tolerance Tolerance is an important feature of a surveyability predicate, (1) because this is one of the key assumptions we used for the argument against the Dedekindian

approach (2.1.4), and (2) because as we will see, the totality of the natural numbers which the strict finitist takes to be legitimate, is supposed to has 'indefinitely many' numbers thanks to this property (3.1.4).

For the time being, we will consider only using the decimal notation. Let us write $x \in \mathcal{A}$ for that number n is intelligibly representable in notation \mathcal{A} . We call the decimal notation, \mathcal{D} . Notice that $n \in \mathcal{D}$ if and only if one of the nde's of n is intelligibly representable in the decimal notation. Also, notice that $x \in \mathcal{D}$ is a surveyability predicate.

Wright's idea of justifying the tolerance of a surveyability predicate is as follows – although he puts a disclaimer that this is a 'diagnostic', not an 'argument' ([27, p.165]). We assume a principle:

For any *d*-digit expression, x, made of Arabic numerals and for any *d*-digit expression, y, made of Arabic numerals which is different from x in at most one place, one can in practice verify that x is an nde in the decimal notation, if and only if she can in practice verify that y is an nde in the decimal notation.^{*24}

Using this principle, one may argue for that, for any small enough number $n, n \in \mathcal{D} \to S(n) \in \mathcal{D}$, where S(n) is the successor of n, as follows. Suppose $n \in \mathcal{D}$. Then we can in practice recognise what the decimal representation of n is. Call it \overline{n} , and let it be d-digit. By the principle above, any d-digit expression made of Arabic numerals is an intelligible decimal expression. Therefore, if \overline{n} is not a string only made of 9's, then $S(n) \in \mathcal{D}$. If \overline{n} is a string made of d 9's, then the agent grasps what the decimal expression of S(n) would be: it is the string beginning with 1, followed by d 0's. [27, pp.164-5]

This is his idea, and I think this is basically an acceptable argument, with a modification and some additional supports. First of all, I do not think we need the principle above. The intelligibility of the successor can be assured without it. Rather, by having this principle, we may receive unnecessary 'counter-arguments': one may try to offer an argument against that once we have an intelligible *d*-digit decimal expression, any *d*-digit expression made of Arabic numerals is an intelligible decimal expression^{*25}. But we do not need this principle in order to conclude that the successor is also intelligibly decimally representable.

^{*24} In fact, Wright introduces two principle. But I think we could summarise them into one. His two principles are: '[...] first, that any *n*-fold sequence of single decimal numerals is neither easier or harder to take in than any other *n*-fold such sequence differing from it in at most one place; second, that if x is ϕ -able by X, and y is no harder to ϕ than x, then y is ϕ -able by X.' [27, p.164]

^{*25} One could say that (1) $10^{1,000,000}$ is very easily seen as an intelligible decimal expression, but the first one million and one digits of π may not be so; or that (2) even if, on the one hand, $10^{1,000,000}$ is very easily seen as an intelligible decimal expression, in order, on the other, to assert the string made of one million and one 9's is also an intelligible decimal expression, one needs to apply the principle very many times – the actual human being might not be able to complete this process. But these complaints could be presented only when we appeal to the principle.

Without the principle, the above argument could be presented as follows.

Suppose $n \in \mathcal{D}$. We can now in practice grasp what the intelligible decimal representation of n is. Call it \overline{n} , and let it be d-digit. Therefore, if \overline{n} is not a string only made of 9's, then $S(n) \in \mathcal{D}$. If \overline{n} is a string made of d 9's, then the agent grasps what the decimal expression of S(n) would be: it is the string beginning with 1, followed by d0's.

I think that to defend this argument, it is enough to discuss the following points. (1) One might wonder why we can assume that the number of digits which \overline{n} has, is a specific number d: Is there no case where this d is too big to handle? I think this is a legitimate setting because we are supposing that $n \in \mathcal{D}$. It would be a easier task to grasp how long an expression is (i.e. how many digits a decimal expression has) than to grasp the number of each digit of the expression.

(2) One might wonder why the disjunction that either \overline{n} is a string of d 9's or it is not, is assertible. One way of answering is to appeal to the anti-realist argument we saw in 2.2.2: the principle of bivalence holds for a statement with a decision procedure. It is true that the anti-realist argument was limited to the mathematical statements, and a statement about the arrangement of an expression might be out of range, but one might argue that the conditions for the transition can be met in this case also (2.2.1.2, 2.2.1.3). However, the strict finitist admitted the argument as a compromise. So it may be better to avoid using it, if we can. And I think we can. It is because, again, we are supposing that \overline{n} can in practice be grasped: this should include that we can in practice inspect all the digits of \overline{n} and judge whether there occurs a numeral which is not 9.

(3) One might wonder how we could in practice grasp the actually intelligible decimal expression of S(n), when \overline{n} is not a string only made of 9's. Or, one might wonder how we could do this, when \overline{n} is a string made only of 9's. I answer to both at the same time: it is because it is algorithmically determined what the actually intelligible decimal expression of the successor of \overline{n} is. My point is that the process does not change according as \overline{n} is made only of 9 or not. We execute one unique process as follows:

- 0. Look at the rightmost numeral of \overline{n} ;
- Determine whether the numeral you are looking at is 9 or not, keep the answer, and then replace the numeral you are looking at with its successor numeral^{*26};
 - 1.1 If the answer is yes, then go to 2.;
 - 1.2 If the answer is no, then end this process;

^{*26} The 'successor numerals' are defined as expected: that of 0 is 1; that of 1 is 2; that of 2 is 3; that of 3 is 4; that of 4 is 5; that of 5 is 6; that of 6 is 7; that of 7 is 8; that of 8 is 9; and that of 9 is 0.

- 2. Determine whether there is a numeral to the left of the numeral you are now looking at;
 - 2.1 If yes, then look at the left number to the numeral you are now looking at, and go to 1;
 - 2.2 If no, write 1 to the blank you are looking at, and end this process.

If one agrees to that the result of this process is an actually intelligible decimal expression when \overline{n} is not a string only made of 9's, then she must agrees to the same statement for the case where \overline{n} is.

This is how I defend the 'tolerance statement' $(x \in \mathcal{D} \to S(x) \in \mathcal{D})$, for any small enough x). But I think that as long as the statement which stands for the tolerance is accompanied by a similar procedure, a like argument would justify it. For example, the statement

For any small enough number n, if $n \in \mathcal{D}$ and n is not 0, then the immediate predecessor of n is actually intelligibly representable in the decimal notation

should be able to be defended.

Let me, however, put the following note. The above argument alone never guarantees the principle which Wright presented and which I dismissed. It only asserts that if n is actually intelligibly representable in the decimal notation, then so are its adjacent numbers. Whether a remote number could be in practice decimally represented is another matter.

■3.1.2.2 The non-vagueness First, let us clarify the distinction between non-vagueness and (actual weak) decidability: they are similar, but not the same. While a predicate P is not vague just in case there is no such an n that we can actually recognise that neither P(n) nor $\neg P(n)$ holds, it is (actually weakly) decidable just in case $\forall n [P(n) \lor \neg P(n)]$.

My attitude towards the non-vagueness of a surveyability predicates is very simple: I think we can see clearly Wright's reasoning, and we can accept it. I reconstruct Wright's argument that $x \in \mathcal{D}$ is not vague as follows.

Suppose we can actually recognise that neither $n \in \mathcal{D}$ nor $n \notin \mathcal{D}$ could be asserted. But to say that we actually recognise that $n \in \mathcal{D}$ could not be asserted is to say that we actually recognise that n has no actually intelligible decimal representation: namely, it is to say we actually recognise that $n \notin \mathcal{D}$. This is contradictory to say that we actually recognise that $n \notin \mathcal{D}$ could never be asserted. Therefore the supposition is contradictory. [27, p.132]

I see no difficulty in this argument, and I accept this.

One thing interesting to note is that it seems that one could find a vague predicate in the accepted intuitionistic mathematics. Define a set of infinite sequences of natural numbers

 $\mathcal{M}_2 = \{ \alpha : \forall n \in \mathbb{N} [\alpha(n) \in \{0,1\} \land \alpha(n) \le \alpha(n+1)] \}$. Also define an infinite sequence α as follows:

For each $n \in \mathbb{N}$, if $n < k_{99}$, then $\alpha(n) = 0$; and if $k_{99} \leq n$, then $\alpha(n) = 1$.

Then $\alpha \in \mathcal{M}_2$. In this case, at least one predicate about functions from natural numbers to \mathcal{M}_2 would be vague. Let us define a predicate P about such functions by

$$P(f) \iff \exists n \in \mathbb{N} [f(n) = \alpha].$$

Then we would find a function g such that we can (in principle) recognise that neither P(g)nor $\neg P(g)$ could be asserted. Define function $g : \mathbb{N} \to \mathcal{M}_2$ by:

- $g(0) = \underline{0}$, and for each $n \in \mathbb{N}$, $g(n+1) = \underline{0}n * \underline{1}$, where
- (1) <u>n</u> (for some $n \in \mathbb{N}$) is the infinite sequence α such that $\forall i \in \mathbb{N} [\alpha(i) = n]$;
- (2) αn (for an infinite sequence α and an $n \in \mathbb{N}$) is sequence $\langle \alpha(0), \alpha(1), \cdots, \alpha(n-1) \rangle$; and
- (3) $s * \alpha$ (for a finite sequence s and an infinite sequence α) is the infinite sequence β with $\forall i \in \mathbb{N} [i < length(s) \rightarrow \beta(i) = s(i) \land length(s) \leq i \rightarrow \beta(i) = \alpha(i - length(s))].$

Suppose that P(g) holds. Then we could in principle find an n such that $g(n) = \alpha$. Notice that either n = 0 or $n \neq 0$. Therefore either $\alpha = \underline{0}$ or there is an n such that $\alpha = \underline{0}n * \underline{1}$. Namely, either $\forall n \in \mathbb{N} [n < k_{99}]$ or $n = k_{99}$. But neither is not (now) even in principle provable. Therefore P(g) is not assertible. Suppose that $\neg P(g)$ holds, i.e., $\neg \exists n \in \mathbb{N} [g(n) = \alpha]$ holds. This means that we can in principle know for each $n \in \mathbb{N}$ that $g(n) \neq \alpha$. But, in fact, this is something we cannot (now) even in principle prove. Therefore $\neg P(g)$ is not assertible. Hence P is vague.^{*27}

■3.1.2.3 The actual weak decidability We now look at the actual weak decidability of $x \in \mathcal{D}$. Wright seems to be discussing this issue from page 133 to 135, and from page 160 to 162 of 'Strict Finitism'. However, Wright's ideas are quite unclear: after a complicated consideration, he writes 'nothing has been settles' [27, p.135]. But it is nonetheless true that the actual weak decidability is a key assumption for establishing the harm of the Dedekindian approach. If the strict finitist withdraws the actual weak decidability, then she would also lose the argument against the Dedekindian approach itself. It is certain at least that Wright does not wish to rescind his argument. He writes right after the quote above:

And the prospect remains that, from a strict finitist point of view, every well-defined set of integers large enough to contain the extension of some tolerant surveyability predicate

^{*27} This example is inspired by [24].

as a subset will rank as Dedekind-infinite. [27, p.135]

Hence, I think, it would be more productive to try to salvage and reconstruct what he thought might be a support to the actual weak decidability. In what follows, I do not assert that I have correctly interpreted his writing, but instead only that the idea I present is plausible. (By the way, in order to provide a support to the actual weak decidability, we could appeal to the anti-realist idea that we saw in 2.2.2, which I accepted as a compromise. But I rather try searching for another solution, not relying on a compromise.)

It seems to me that the basic ground for believing in the actual weak decidability of $x \in \mathcal{D}$ is that a persistent failure should amounts to the negation of the application of the predicate. I would provide the following two scenarios as example. (1) A mathematical layerson tries to know the decimal expression of $2^{1000^{1000}} - 1$. She visits a calculator website and inputs $2^{1000^{1000}} - 1$: let us say the result is an 'error'. She visits another website and tries again: the result is, this time, an 'infinity', which is obviously a synonym of 'error'. She gives up finding the answer: she regards it as beyond today's calculation technology. (2) An elementary school teacher teaches her pupil what multiplication is and has the pupil try calculating $2 \times 2 \times 2 \times 2 \times 2 \times 2$. She fails many times. She says that the answer is 32, then says that it is 60, and says that is 32 again. From the teacher's perspective, it is obvious that she never succeeds in calculating 2^6 . The teacher realises that this calculation is beyond the pupil's calculation ability this day. These are scenarios where a persistent failure amounts to the negation of $n \in \mathcal{D}$ for some number n. Wright seems to be suggesting that this conversion happens when there is no other excuse for the failure. So, advancing his idea, one could say that the negation holds as an explanation of a persistent failure: the attempts in the scenarios to express a number fail *because* it is impossible to express.

Wright also seems to be thinking about success. But he appears to be negative on this point: it is controversial that the emergence of a result in finding a decimal expression simply amounts to the application of the predicate. This is clear when we think of the pupil in the second scenario. Wright, first, seems to be thinking that one needs 'operational distinctions for success and failure', because

[...] unless it is clear at any particular point whether X has so far succeeded in ϕ -ing x or not, no ground has yet been provided for thinking that we must be able to get into a position where we can (weakly) decide between the alternatives. [27, p.133]

It is not quite clear from here whether he thinks that a failure is 'not having succeeded' or not. But I think that, in any case, we can agree to the view that one needs operational *criteria*, at any stage in an on-going process, for (1) having succeeded, (2) having failed and (3) having to wait for the next stage. I think that we could identify two difficulties, which Wright is concerned with, in setting the operational criteria for success. (a) The first difficulty is that, at least in the case of predicate $x \in \mathcal{D}$, having criteria for success could be question begging. Surely, while the pupil in scenario (2) does not know the decimal expression of 2^6 , we know that it is 64. So, the criterion for success for the pupil may be to produce the result $2 \times 2 \times 2 \times 2 \times 2 \times 2 = 64$. But nobody knows the decimal expression of $2^{1000^{1000}} - 1$, which appeared in scenario (1). Therefore in this case we cannot have criteria for success in the same way we have in the case of 2^6 . We can have such criteria for success only after we have succeeded in the attempt. Wright seems to be giving another way of describing this situation: the criteria for success will be relative to information independent of the agent (as in scenario (2), where the pupil is not given the answer) [27, pp133-4].

A remedy which we might be able to find in Wright's writing is the idea that the task of making the criteria is an *epistemic* matter [27, p.161]. I interpret him as suggesting that criteria for success do not necessarily have to specify the correct decimal expression explicitly, and that we should be content with criteria as long as the 'epistemic' consideration which gives us them is reliable. For example, I think that we could propose a purely algorithmic criterion for success: we could stipulate that an attempt of executing a process is not successful at any stage before its end; and that an attempt is successful at the end, if we are sure about the correctness of the execution of it.

To add to the above, I think that the conversion of failure into negation could be understood from this epistemic perspective. Frankly speaking, we should, and sometimes indeed do, adopt decision procedures, with certain operational criteria for failure, such that when we have failed in them some definite times, we are allowed to conclude that statements at issue are falsified. Certainly, a situation in which

[...] on the one hand, nothing emerges which we feel to be, dependably, the right result; but on the other hand, things do not go *so* badly, under so wide a variety of circumstances, that we feel justified in claiming that to identify the right result is beyond us [...] [27, p.134]

may happen. But we should say that this is a symptom of the negation of a statement at issue. We should have an epistemic answer to what times this situation can happen before we decide that the statement is negated.

The point which I interpret as Wright's second difficulty is (b) that, in the cases of success, according to criteria we admit, there seems to be no behavioural difference between genuine success and success which is not. To quote,

In contrast with raising a bar above one's head, for example [of a physical case], there

is no finite performance which conclusively shows that somebody finds a particular expression *intelligible*; and while genuine recognition of a reduplication of a proof shows that one found the original surveyable, there is no sharp behavioural distinction between an identification based on genuine recognition and one fluked on a guess. [27, p.162]

A surveyability predicate is about our recognitional abilities. Therefore, as long as the criteria for success are about the behaviours of an agent, we would not be able to distinguish whether she has intellectually recognised or not. It seems to be a 'mental' matter whether an agent has really recognised, or it may be a matter of 'luck' that an agent has guessed right.

I cannot find a remedy in Wright's text, but to me this essentially is an epistemic matter. I do not think that one can deny the following: (1) if it is epistemically (or according to the philosophy of mind) correct that there is no decision procedure of a surveyability predicate at issue such that we can check whether an agent is correctly following it by observing her behaviours; and (2) if the strict finitist is epistemologically convinced of the idea that a decision procedure must be behavioural; then there is no decision procedure for the surveyability predicate. The only wayout for the strict finitist seems to be *either* to engage in the epistemological enterprise to find a behavioural decision procedure of a surveyability predicate; *or* to adopt and be content with a 'loose' epistemology in the sense that mental states are not required to intellectually recognise. The second will allow a 'mindless' machine to legitimately recognise a decimal expression, and a 'lucky' student to legitimately write a correct proof. I take this epistemological stance, and I see no further difficulty.

So far, we have been seeing (1) that a decision procedure should have operational criteria for success, failure, and need for waiting the next step; (2) that a procedure must be given by an epistemic consideration; (3) that the criteria for success do not have to mention the result explicitly; (4) that a procedure needs to have the criteria for conversion of failure into negation; (5) that the strict finitist may have to embark on such an epistemic consideration herself, if she cannot be content with behavioural decision procedures; and (6) that the strict finitist can choose to be content with such procedure.

The conclusion which I wish to draw from the above obversations is that the decision procedures which meet those characterisations may not be decidable in practice, but will have the property of actual weak decidability. Simply put, an agent who executes a decision procedure, for a statement P(a), which fits the characterisations will necessarily be judged to be either successful or failing in the execution, and if she is successful, then we have grounds for asserting P(a), and she fails, then we have grounds for asserting $\neg P(a)$. This happens, in this setting, because a decision procedure and the operational criteria for success and failure are set by epistemological considerations, and what is epistemically affirmative gives grounds for asserting an affirmative statement, and what is epistemologically negative gives grounds for asserting a negative statement. Whether any surveyability predicate is actually weakly deciable is a matter of epistemology and epistemological choice.

3.1.3 The notation

It is important for the strict finitist to specify the notational system she uses. We spoke about notation several times: we considered the idea that a notational simplification is always possible, and whether we could extend the range of what is 'finite' by extending the extant notational systems in section 2.1.2. Also, we said that $2^{1000^{1000}} - 1$ is unintelligible for the strict finitist, if she sticks to the decimal notation, \mathcal{D} , in 2.2.2. One needs to notice that what is intelligible is dependent on what notational system we use, and that we only admit what is intelligible as a proper target of consideration. A natural number, for the strict finitist, is a number which we can in practice represent^{*28}. Indeed, according as the notational system an agent adopts, the totality of natural numbers varies [7, p.249]. So, the strict finitist should make clear what notation she admits as legitimate.

It would be here helpful to introduce and consider the following notions. (1) Wright uses the notion he calls *completeness*:

Let us say that a system of notation for zero and the positive integers is said to be *complete* just in case every integer actually intelligible representable in that notation by a numeral n has n predecessors likewise so representable. [27, p.156]

Formally, a system of notation \mathcal{A} is complete, just in case $\forall n [n \in \mathcal{A} \rightarrow \forall m \leq n [m \in \mathcal{A}]]$. Let us call a system of notation which is not complete, *incomplete*. Dummett seems to be considering essentially the same notions as these [7, pp.249-50]. We admit that \mathcal{D} is complete without discussion: for example, 1,000,000,000 $\in \mathcal{D}$, and any natural number from 0 to 999,999 is actually decimally representable.

Dummett thinks that there is no great advantage of taking, over \mathcal{D} , an advanced notation such as that which admits addition, multiplication and exponentiation (written as \mathcal{T}). His reason for this is (1) that such a system is incomplete, and (2) that such a system will not be closed under all effective arithmetical operations definable over it [7, p.250]. I basically agree with his stance that there will be no gain even if we choose an incomplete system over a complete system.

One thing I regret is that, as a reason for (1), Dummett only says that, while $10^{10^{10}}$ is intelligibly representable in \mathcal{T} , 'plainly, the totality [of the numbers intelligibly representable in \mathcal{T}] does not contain as many as $10^{10^{10}}$ numbers' [7, p.250]. Unfortunately, I can show neither that

^{*28} Dummett writes: '[...] by 'natural number' must be understood a number which we are in practice capable of representing' [7, p.249].

 \mathcal{T} is really incomplete, nor that the system which admits the decimal expressions, subtraction and exponentiation (in which $2^{1000^{1000}} - 1$ is intelligibly representable) is incomplete.

However, it seems that if we assume two seemingly plausible statements, we can show that there is at least one incomplete system of notation : such a system is that which only admits the decimal expressions and exponentiation (written as \mathcal{D}^{exp}).

Let us assume (1) $2^{1000^{1000}} - 1 \notin \mathcal{D}$; and (2) there exists a prime number c such that $2^{1000^{1000}} < c < 3^{1000^{1000}}$. Evidently, $3^{1000^{1000}} \in \mathcal{D}^{exp}$. We show that $c \notin \mathcal{D}^{exp}$. First, suppose that $c (= c^1) \in \mathcal{D}^{exp}$. Then, $c \in \mathcal{D}$. Now by the completeness of \mathcal{D} , it holds that $\forall n [\exists m \ge n [m \in \mathcal{D}] \to n \in \mathcal{D}]$. Take the contraposition: $\forall n [n \notin \mathcal{D} \to \forall m \ge n [m \notin \mathcal{D}]]$. Therefore, by (1), it holds that for any $n \ge 2^{1000^{1000}}$, $n \notin \mathcal{D}$. Therefore $2^{1000^{1000}} < c$ tells us that $c \notin \mathcal{D}$. A contradiction.

Therefore, we might be able to say that completeness and incompleteness are not empty notions.

So, let us in this thesis assume that the strict finitist adopts \mathcal{D} as her standard notational system.

3.1.4 Indefiniteness: The strict finitist numbers

Hereafter, following Wright's (implicit) terminology, we use 'indefinite' ('indefinitely many') to describe how many objects the strict finitist can actually deal with [27, Appendix]. As I mentioned in 3.1.3, the legitimate natural number for the strict finitist is that for which she can in practice intelligibly represent (or recognise)^{*29}. Since we have settled down what notational system the strict finitist uses (that is the decimal notation, \mathcal{D}), we can say as follows: there are *indefinitely many* natural numbers which are actually intelligibly representable in \mathcal{D} . We write \mathbb{D} for the totality of the strict finitistically legitimate natural numbers, namely, the extension of $x \in \mathcal{D}$. Therefore, of course, $\forall n \ [n \in \mathbb{D} \leftrightarrow n \in \mathcal{D}]$, and \mathbb{D} is indefinite. We assume that the strict finitist can deal with a totality of objects as long as she can in practice define a one-to-one mapping from it to \mathbb{D} . So, if there a totality \mathbb{E} and there is an actually recognisable order isomorphism from \mathbb{E} to \mathbb{D} , we could deal with \mathbb{E} as if it is the totality of legitimate numbers. Therefore, for the strict finitist in this thesis, the totality of number is unique up to order isomorphism.

We here see two features of \mathbb{D} . (1) Dummett calls this indefinite totality \mathbb{D} , weakly finite and weakly infinite, and Wright follows it. A totality is said to be weakly finite just in case,

^{*29} By the way, Wright writes: 'For the strict finitist, only those first-order numerical expressions may be thought of as having sense, and therefore reference, which are actually capable of intelligible employment by the community' [27, p.153]. So he counts, as nde's, 'first-order numerical expressions' in addition to the usual numerical terms. I must confess that I have no idea what he intends to mean by this phrase.

for some finite ordinal n (in classical or intuitionistic sense – since the strict finitist would not understand this 'finitude'), there exists a well-ordering of the totality with no n-th number; and a totality is said to be *weakly infinite* just in case there exists a well-ordering of the totality with no last member [7, p.258] [27, p.154]^{*30}. Indeed, we seem to be able to admit this characterisation. For the weak finitude, it holds that $\forall n \in \mathbb{D} [n < 2^{1000^{1000}} - 1]$. To assert the weak infinity, we would appeal to the tolerance of the surveyability predicate $x \in \mathcal{D}$. As we saw in 3.1.2.1, we think it holds that $\forall n \in [n \in \mathcal{D} \to S(n) \in \mathcal{D}]$. This means that, for any $n \in \mathbb{D}$, we can in practice find the successor of n in \mathbb{D} . Therefore, it seems that sequence $\langle 0, 1, 2, \cdots \rangle$ has no last member.

(2) Wright points out as follows.

[...] in order for X to recognize a smallest decimal numeral which he cannot understand, he will have to have some *independent* way of identifying it as denoting the successor of a number whose decimal numeral he does understand. And what goes for X goes for the community at large; so recognition of sharp boundaries to what we can survey, or understand, is going to require recourse to some kind of technology. [27, p.164]

He is here suggesting that it is difficult to recognise the 'sharp boundary' of \mathbb{D} . As we saw above, \mathbb{D} is weakly finite and weakly infinite. But because $x \in \mathcal{D}$ is tolerant, it seems to be impossible actually to recognise n (the 'sharp boundary' of \mathbb{D}) such that $n \in \mathbb{D} \land S(n) \notin \mathbb{D}$.

Wright, when he writes the above, seems to be presenting the idea that the sharp boundary of \mathbb{D} can actually be recognised in the end, if we appeal to technology^{*31}. An individual agent, X, alone cannot recognise the sharp boundary; neither can a collective agent to which X may belong recognise; but a collective agent can rely on technology they possess in order to provide the decimal expression of the natural number, n, such that they cannot in practice recognise it, but they can do the decimal expression of the immediate predecessor of it – Wright seems to be endorsing this view.

However, one may say that when we appeal to technology to produce the decimal expression of a number, then we are entitled to say that we know the expression. So, it seems that we can draw the following ideas from the suggestion above: (1) for any n, whether $n \in \mathbb{D}$ or not depends on an individual agent at issue, the collective agent to which she belongs to, and the technology she possesses; (2) if there is nothing (no other agent or no technology) to which an agent at issue can appeal in order to actually intelligibly represent a number n, then

^{*30} Wright, at one place, refuses to call D weakly finite and weakly infinite, and calls 'structurally indeterminate' [27, p.158]. But he describes D as weakly infinite again [27, p.161]. I cannot now reasonably interpret this transition. I assume that he in the end admits this characterisation.

^{*&}lt;sup>31</sup> But I should inform that for Wright, this point is a mere lesson of a preliminary consideration for the tolerance of a surveyability predicate, and he presents what he thinks is the real motivation for the tolerance right after this place [27, pp.164-5].

 $n \notin \mathbb{D}$; (3) therefore, there is no way an individual agent can actually recognise the smallest decimal expression which she, the collective agent to which she belongs, or any technology she possesses, cannot produce; and (4) in other words, ultimately, she cannot actually recognise the sharp boundary of \mathbb{D} .

3.2 Against inconsistency

In this section, we discuss the doubt about the consistency of strict finitism. The doubt takes the form of the paradox known as 'Wang's paradox'. Wright, in section 11 of 'Strict Finitism', examines this phenomenon, and argues that the strict finitist can avoid it [27, Sect.11]. This paradox is essentially the same type of paradox as that known as the 'Sorites paradox', but we call this 'Wang's paradox' following Dummett's paper 'Wang's paradox', which seems to be the first to have brought this to our attention as a challenge to strict finitism^{*32}.

In what follows, we only consider Wright's ideas: Dummett also examines this, but his conclusion is that the strict finitist will inevitably suffer this paradox, if she is serious about her position, and it only shows that the surveyability predicate (in Wright's and our terminology) is inconsistent [7]. First, I will display what this paradox is like and how this can be presented (3.2.1); then I will reconstruct and represent Wright's solution to this (3.2.2); and finally I will try to provide a support to his idea and defend the consistency of strict finitism (3.2.3).

3.2.1 Wang's paradox and its presentation

At first, let us see how the paradox can be presented. The following formulation is by Wright, but we can see essentially the same presentation in Dummett's paper [27, pp.155-6] [7]. We define the following notions. Let k be a number such that k successive pairwise steps of universal instantiation and Modus Ponens constitute a proof structure which the actual human beings can deal with, while 2k steps do not. While Wright uses a different way of presenting a proof structure, we take it that one 'pairwise step of universal instantiation and Modus Ponens' stands for the following structure [27, p.155].

$$\begin{array}{c} \forall x \left[P(x) \to Q(x) \right] \\ \hline P(n) & \hline P(n) \to Q(n) \\ \hline Q(n) \end{array}$$

Also, m is called *Small* (we write Sm(m) for this) just in case m + k such pairwise steps constitute a structure which the actual human beings can deal with^{*33}.

 $^{^{*32}}$ By the way, Wright call this the 'Sorites paradox' [27, Sect.11].

^{*&}lt;sup>33</sup> Wright uses 'small' in small letters [27, p.155].

Now, Wang's paradox about humanly feasibility of proof structure is said to be presented as follows. 0 is Small, because a structure with 0 + k = k pairwise steps is humanly feasible. It is plausible that $\forall n [Sm(n) \rightarrow Sm(S(n))]$ (i.e., the tolerance of predicate Sm). This is plausible enough, the proponent of the paradox would say, because this is at least as plausible as the statement that 'a decimal numeral consisting of an initial '1' follows by n '0's is actually intelligible whenever n '9's is actually intelligible' is [27, p.155]. Thus, by k pairwise steps of universal instantiation and Modus Ponens, we can prove that Sm(k), i.e., a structure with k + k = 2k pairwise steps is humanly feasible, contrary to the hypothesis. [27, pp.155-6]^{*34}

What this paradox is supposed to show is the following: predicate 'x successive pairwise steps of universal instantiation and Modus Ponens constitute a proof structure which the actual human beings can deal with' seems to be a surveyability predicate; so, the strict finitist may suppose that the extension of this predicate is weakly finite and weakly infinite at the same time; but this supposition is contradictory. The proponent of the paradox would say that, therefore, the position based on the extension of a surveyability predicate is essentially inconsistent; that the strict finitist, whose mathematics will be of this sort, cannot have a consistent mathematics; and that therefore strict finitism is not acceptable.

3.2.2 Wright's solution

Wright's arguments against the above accusation can be boiled down to an argument to stop the generalisation by presenting a counterexample. His idea is that a totality which is weakly finite and weakly infinite but not susceptible to Wang's paradox will be a 'safe' totality for the strict finitist: such a totality will be that on which the strict finitist will develop her mathematics.

He starts with checking the notion of proof. A proof of Wang's paradox must be acceptable for the strict finitist, in order to be a threat to strict finitism. We would say that the strict finitist accepts a proof with indefinitely many steps, and may say that she can agree to counting a 'pairwise steps of universal instantiation and Modus Ponens' as one step. Wright thinks that the strict finitist can concede further. We call a proof *surveyable* just in case we are actually able to say of the proof what its every *n*-th line will be – 'even if the labour of writing it out correctly and convincing ourselves that that is what we had done would be beyond us' [27, p.157]. Wright says that the strict finitist can and should respect a surveyable proof in this sense. Surely, this 'surveyability' of proof seems to be an 'in principle' kind of notion: a proof is called surveyable, if we can in principle write down any line. To call this 'surveyability' would sound uneasy, but let us be content with his naming.

^{*34} Wright writes: 'Thus by k pairwise steps of universal instantiation and Modus Ponens we can prove that k + k = 2k, is small, contrary to hypothesis' [27, pp.155-6]. But I believe that this 'k + k = 2k, is small' is a typo.

Also, we use the notion of a supersessor (or supersede): a system \mathcal{B} of notation supersedes a system \mathcal{A} of notation just in case both \mathcal{A} and \mathcal{B} are complete and there is a number $n \in \mathcal{B}$ such that n is an upper bound of $\{m : m \in \mathcal{A}\}$ [27, p.156].

Wright argues that a proof of Wang's paradox is surveyably constructed for predicate ' $x \in \mathcal{L}$ ' (for some notation \mathcal{L}), only if \mathcal{L} has a recognisable supersessor. In general, the proof essentially would be a structure with the following features:

- (1) It is a chain of universal instantiation of $\forall n \ [n \in \mathcal{L} \to S(n) \in \mathcal{L}]$ and Modus Ponens;
- (2) It starts with an acceptable sentence ' $k \in \mathcal{L}$ '; and
- (3) It ends with an unacceptable sentence ' $s \in \mathcal{L}$ '.

At first, let me note that we could suppose that \mathcal{L} is complete. This is because there would be no motivation for constructing a proof of Wang's paradox against an incomplete notation. For the proponent of the paradox, what is important is to show that the totality of numbers which the strict finitist adopts is inconsistent; and the strict finitist adopts a complete notation for it.

Now the important thing here is, Wright seems to be suggesting, to think about in what notation k and s could be expressed. I reconstruct his argument [27, p.159]. k is actually intelligibly representable in \mathcal{L} , because, of course $k \in \mathcal{L}$ is correct. s is not actually intelligibly representable in \mathcal{L} , because, again, $s \in \mathcal{L}$ is not acceptable. Notice that for the proof to be acceptable for the strict finitist, every line of the proof must be intelligibly representable (in principle). Therefore, every line of the first several steps of the proof,

must be recognisable for the strict finitist. Therefore S(k) and S(S(k)), in this example, must be actually intelligibly representable too. By the same argument, for the proof to be acceptable for the strict finitist, any number n such that $k \leq n \leq s$ must be actually intelligibly representable, in some system, \mathcal{L}' , of notation. \mathcal{L}' is a recognisable supersessor of \mathcal{L} , because (1) every number n such that $k \leq n \leq s$ must be actually intelligibly representable for the strict finitist in \mathcal{L}' ; and (2) we can show that s is an upper bound of $\{n : n \in \mathcal{L}\}$ as follow.

Consider $n \in \mathcal{L}$. Suppose for contradiction that s < n. Since \mathcal{L} is complete, $\forall m \notin \mathcal{L} \forall n \geq m [n \notin \mathcal{L}]$. Hence $\forall n \geq s [n \notin \mathcal{L}]$. Therefore the supposition that s < n contradicts to $n \in \mathcal{L}$.

Therefore, if a system of notation which has no recognisable supersessor, Wang's paradox

cannot surveyably be constructed. So the strict finitist may use such a system for the basis of her mathematics.

3.2.3 A support to Wright's solution

What is unfortunate is that Wright does not provide firm grounds for the existence of a system of notation without a recognisable supersessor. He writes:

Intuitively, there must be such systems of notation, since the price of relative brevity in the numerals in a system is relative richness of primitive symbolism, and our memories are limited. [27, pp.156-7]

One may feel that we need a more detailed argument. In this portion of the thesis, I try to defend that there is a system of notation without a recognisable supersessor.

I assert that the decimal notation, \mathcal{D} , has no recognisable supersessor from the strict finitist's perspective. My idea, in short, is that as long as the strict finitist sticks to the view that a legitimate number is that which satisfies predicate $x \in \mathcal{D}$, \mathcal{D} would have no recognisable supersessor. The strict finitist definition of numbers will be the key.

Before proceeding further, I would like to emphasise that Wang's paradox is supposed to show that strict finitism is inconsistent from the strict finitist's perspective. To quote Wright's phrase, it is thought that 'Sorites [Wang's] paradoxes continue to be generatable within the framework of strict finitism' [27, p.155, my emphasis]. The threat of the paradox is pressed to the strict finitist, saying that inconsistency seems to be presentable only using what she admits as legitimate.

Now, our strict finitist uses \mathcal{D} as the standard notation and thinks that \mathbb{D} is the suitable totality of numbers. So let us restate Wang's paradox in the context of \mathcal{D} . The proponent of Wang's paradox would appeal to the following three statements.

(1) $\forall n [n \in \mathcal{D} \rightarrow S(n) \in \mathcal{D}];$ (2) $0 \in \mathcal{D};$ and (3) $2^{1000^{1000}} - 1 \notin \mathcal{D}.$

The strict finitist would accept them, because (1) is the tolerance of predicate $x \in \mathcal{D}$, (2) is evident, and (3) is the running supposition.

Wright's idea we saw was that for this proof to be acceptable, there must be a system, \mathcal{D}' , of notation such that

(1)
$$0 \in \mathcal{D}';$$

- (2) $2^{1000^{1000}} 1 \in \mathcal{D}'$; and
- (3) \mathcal{D}' supersedes \mathcal{D} .

This is because, to restate, if there is no such system, the strict finitist will not accept the proof: she will not understand, in particular, the unacceptable statement that $2^{1000^{1000}} - 1 \in \mathcal{D}$.

One may expect that the system which admits the decimal expression, subtraction and exponentiation (let us write \mathcal{E}) is such a notation, for example. But I argue that \mathcal{E} is not a supersessor of \mathcal{D} . Suppose for contradiction that \mathcal{E} is a supersessor of \mathcal{D} . Then we can actually recognise a number, s, in \mathbb{D} such that $s \in \mathcal{E}$ and $\forall d \in \mathbb{D} [d \leq s]$. But, here, notice that this s is said to be in \mathbb{D} . We can actually recognise this s only if we can actually recognise the maximum of \mathbb{D} . This is against the tolerance of $x \in \mathcal{D}$. So, \mathcal{E} is not a supersessor of \mathcal{D} .

As I mentioned, this idea is dependent on the strict finitist definition of numbers. The definition, of course, uses the notion of the decimal notation, and I assert that as long as the strict finitist sticks to this definition, she does not have to accept a proof of Wang's paradox. I do not take this as question begging: it is not question begging for the strict finitist to stick to her definition of numbers, in order to defend the consistency of the definition itself. This is because Wang's paradox is supposed to show in general that, if the strict finitist picks up a system, \mathcal{A} , of notation, which satisfies

- (1) $\forall n [n \in \mathcal{A} \to S(n) \in \mathcal{A}];$
- (2) for some number $k, k \in \mathcal{A}$; and
- (3) for some number $s, s \notin \mathcal{D}$,

in order to make use of the extension of predicate $x \in \mathcal{A}$ for her mathematics, inconsistency will arise. What I am asserting is that when the strict finitist takes \mathcal{D} , an inconsistency will not arise the way the proponent of the paradox argues.

Let me clarify my ideas by seeing the following two points. The first point is that the strict finitist could think about a totality of numbers other than \mathbb{D} . To see this, let us consider the following system, $\Delta^{\times 2}$, of notation. This is a notation where the only legitimate expressions are those in the form of

 $n \times 2$,

where $n \in \mathcal{D}$. I think that the extension of predicate $x \in \Delta^{\times 2}$ is indefinite, because any expression actually intelligibly representable in $\Delta^{\times 2}$ is actually intelligibly representable in \mathcal{D} , and vice versa. Therefore, our strict finitist can make use of the extension of $x \in \Delta^{\times 2}$ as another totality of numbers^{*35}. Therefore, the strict finitist would not think that there is a number s such that $s \in \Delta^{\times 2}$ and for any number d, $d \leq s$. This is because the strict finitist now thinks about the numbers representable in $\Delta^{\times 2}$. There should not be the maximum of what is representable in $\Delta^{\times 2}$.

 $^{^{*35}}$ For the strict finitist in this thesis, the totality of numbers is unique up to order isomorphism. See 3.1.4.

However, the standard notation is \mathcal{D} , and other totalities of numbers than \mathbb{D} must be considered in light of \mathbb{D} . My second point is that, when the decimal notation is available, an expression, a, is an actually intelligible numeral just in case $a \in \mathcal{D}$. Let us think about the system, $\mathcal{D}^{\times 2}$, of notation, where an expression, a, is legitimate just in case either $a \in \mathcal{D}$ or $a \in \Delta^{\times 2}$. I think that the natural expectation about the order of the extension of $x \in \mathcal{D}^{\times 2}$ is as follows.

:
? =
$$n \times 2$$

:
 5
 $4 = 2 \times 2$
 3
 $2 = 1 \times 2$
 1
 $0 = 0 \times 2$

Our intuition might tell that there is a 'number' which is actually intelligibly representable in the form of $n \times 2$, but not actually intelligibly representable in the decimal notation. This may be true when one compares two systems ($\Delta^{\times 2}$ and \mathcal{D}), but I do not think it is an accurate strict finitist way of thinking, when the strict finitist looks at one system $\mathcal{D}^{\times 2}$. Since the standard notation is \mathcal{D} , any expression actually intelligibly represented in the form of $n \times 2$ should at first enter the picture as a mere candidate of numeral: the strict finitist must think to what the mapping to \mathbb{D} would map it. To admit it as a numeral, we must locate the number which the expression denotes. To locate it, the strict finitist must try to calculate, in order to represent it in the decimal notation. Either she succeeds or she fails^{*36}. If she fails, she cannot in practice create a bijection from \mathbb{D} to the extension of $x \in \mathcal{D}^{\times 2}$. This means that this extension is larger than she admits. Simply, there is no way she has to admit that there is a number which is not actually intelligibly representable in \mathcal{D} .

Now, we should replace the definition of supersessor, because inequality must be considered in \mathbb{D} . A system, \mathcal{B} , of notation is a supersessor of a system, \mathcal{A} , of notation just in case

- (1) \mathcal{A} and \mathcal{B} are complete;
- (2) There is an actually recognisable order embedding from the extension of $x \in \mathcal{A}$ to \mathbb{D} ;
- (3) There is an actually recognisable order embedding from the extension of $x \in \mathcal{B}$ to \mathbb{D} ; and
- (4) There is an $n \in \mathbb{D}$ with $n \in \mathcal{B}$ such that, for any $m \in \mathbb{D}$ with $m \in \mathcal{A}$, it holds that $m \leq_{\mathbb{D}} n$.

^{*&}lt;sup>36</sup> We argued that predicate $x \in \mathcal{D}$ is actually weakly decidable.

If the conception of numbers which I described is strict finitistically correct, and if the above definition of supersessor is acceptable, it is justifiable that there is no actually recognisable supersessor of \mathcal{D} . If there is such a supersessor \mathcal{D}' , there will be an $n \in \mathbb{D}$ with $n \in \mathcal{D}'$ such that, for any $m \in \mathbb{D}$, it holds that $m \leq_{\mathbb{D}} n$. This means that there is a number n which is actually intelligibly decimally representable, and which is larger than any actually intelligibly decimally representable number. Then, the successor of n cannot be actually intelligibly decimally representable. This contradicts to that an agent cannot actually recognise the sharp boundary of \mathbb{D} , which we saw in 3.1.4.

4. The positive programme (2): The formal theories of strict finitism

The aim of this section is to make a starting point from which the strict finitist formal studies can flourish in the future. In the previous section (3.), we saw the basic concepts of strict finitism and the reasons not to take strict finitism as inconsistent. Regarding them as enough conceptual grounds of strict finitism, we in this section try to build up a formal framework on them.

As I mentioned in 1.2.3, we can find Wright's precursory attempts for the formal theories in his 'Strict Finitism' [27, Appendix]. Our investigations into the strict finitist formal theories proceed by reconstructing and examining them.

4.1 The language

First, we define the language on which we work. It should be a language where arithmetic could be developed.

Our language, L, is a language that has

- (i) = as the only predicate symbol;
- (ii) $S, +, \cdot$ as the only function symbols;
- (iii) $\overline{0}$ as the only numeral symbol;
- (iv) x_0, x_1, \cdots (indefinitely many) as variables; and
- (v) $\land, \lor, \rightarrow, \neg, \exists, \forall$ as the only logical connectives.

In addition to these, we have a category of expressions called *natural-number denoting expressions* (or *nde*'s for short). The *nde-formation rules* of L are as follows:

- (i) Any numeral symbol is an nde;
- (ii) If \overline{n} is an nde, then $S(\overline{n})$ is an nde;
- (iii) If \overline{n} and \overline{m} are nde's, then $\overline{n} + \overline{m}$ is an nde; and
- (iv) If \overline{n} and \overline{m} are nde's, then $\overline{n} \cdot \overline{m}$ is an nde.

We will also use the notion of *canonical nde*'s.

- (i) Any numeral symbol is a canonical nde; and
- (ii) If \overline{n} is a canonical nde, then $S(\overline{n})$ is a canonical nde.

We use the nde's as the terms of L. The formulae of L are as follows:

- (i) If $\overline{n}, \overline{m}$ are nde's, and x, y are variables, then $\overline{n} = \overline{m}, \overline{n} = x, x = \overline{n}, x = y$ are (atomic) formulae;
- (ii) If ϕ and ψ are formulae, then $\phi \wedge \psi$ is a formula;
- (iii) If ϕ and ψ are formulae, then $\phi \lor \psi$ is a formula;
- (iv) If ϕ is a formula, then $\neg \phi$ is a formula;
- (v) If ϕ and ψ are formulae, then $\phi \rightarrow \psi$ is a formula;
- (vi) If ϕ is a formula and x is a variable, then $\exists x \phi$ is a formula; and
- (vii) If ϕ is a formula and x is a variable, then $\forall x \phi$ is a formula.

Hereafter, we freely use and omit brackets for ease of reading.

In 3.1.3, we set the decimal notation, \mathcal{D} , as the strict finitist's standard notation. But we use a language where nde's are not decimal expressions: in L, they are in the forms of $\overline{0}, S(\overline{0}), SS(\overline{0}), S(\overline{0}) + S(\overline{0})$, etc. This is simply because the nde's in these forms are formally easier to deal with. We call the notation where the only legitimate nde's are the canonical nde's above, the *successor notation*. The notation of L is that which allows the successor notation with addition and multiplication.

We nonetheless keep the conception of numbers that the strict finitist can only grasp indefinitely many (or less) numbers, and a system of notation of nde's is acceptable as long as there is an actually recognisable mapping from the nde's in the notation to those in \mathcal{D} . Therefore, an nde in L is admitted as a legitimate numerical expression, only if the agent can in practice recognise the decimal representation of it. This point will be formally realised in the form of an M_W , in the next section (4.2). (To foreshadow, however, we will see that the notations of addition and multiplication may have to be excluded from our language, in 4.4.1.3)

4.2 Semantics

A position in the semantic realism debate has its theory of meaning, and the theory of meaning has its theory of reference (1.1.2). A semantics can be seen as a formal representation of the features of the theory of reference: (usually) it does not really assign the referents to the syntactic expressions, but rather specifies the meaning of, e.g., the logical constants.

4.2.1 Model

First, we define what a model is. In the strict finitist settings, a model is intended to represent all the possible developments of the arithmetical capacities of an individual or collective agent equipped with machines. I reconstruct Wright's writing [27, pp.167-8, (ii) - (iv)].

A Wright model is a sextuple $W = \langle T_W, M_W, E_W, S_W, P_W, K_W \rangle$. We write \mathcal{W} for the set of Wright models. (We could omit the subscript if confusion would not arise.) T_W is a tree, and the elements of T_W are called *elementary arithmetical accumulations* (accumulations for short) of W. T_W ramifies into indefinitely many branches at any accumulation $B \in T_W$. For any $C \in T_W$, we write $\hat{\Sigma}(C)$ for $\{D \in T_W : D \leq C\}$. For any $B, C \in T_W$, we call $\hat{\Sigma}(C)$ a sequence on which B lies, just in case $B \in \hat{\Sigma}(C)$. We write $\Sigma(B)$ for the set of the sequences on which B lies.

Intuitively, when $B \in T_W$ is given, $M_W(B)$ represents the set of numerical expressions (i.e. nde's) which are actually intelligible for the agent at development stage B. $E_W(B)$ represents all the actually intelligible equations between numerical expressions at B; $S_W(B)$ represents all the actually intelligible equations concerning the successor relation at B; $P_W(B)$ represents all the actually intelligible equations concerning addition; and $K_W(B)$ represents all the actually intelligible equations concerning addition. Naturally, the information represented by $S_W(B)$, $P_W(B)$ and $K_W(B)$ shall be included in $E_W(B)$, but we do not wish them to be merged into $E_W(B)$, because of the restriction which we state below.

Formally, M_W is a mapping that assigns a set of nde's to each accumulation in T_W , such that if A is the root of T_W , $\overline{0} \in M_W(A)$. E_W and S_W are mappings that assign to each accumulation $B \in T_W$ a subset of $M_W(B) \times M_W(B)$. P_W and K_W are mappings that assign to each accumulation $B \in T_W$ a subset of $M_W(B) \times M_W(B) \times M_W(B) \times M_W(B)$.

Wright suggests that the strict finitist has to impose a *manageability* restriction on accumulations. The motivation is that

[...] a larger accumulation will not really represent the full increase in information over a smaller one which it includes unless it is *surveyable*; unless it is possible for us to know what, assuming all has gone well, the net gains have been, even though none of us could singly accomplish, or check out, those gains. [27, p.167]

He calls the restriction at issue the 'manageability' restriction, in order to save 'surveyability' for proofs. He writes that this restriction would be stronger than the actual decidability of a member of an accumulation, but he does not settle how the restriction should be formalised [27, p.167]. So, below is my proposal.

The manageability restriction For any model $W \in \mathcal{W}$,

- (1) for any accumulation, B, and any successor, C, of B, we can actually recognise that four of five functions M_W, E_W, S_W, P_W, K_W take the same values for B and C, and we can actually detect X such that the one remaining function, f, satisfies $f(C) = f(B) \sqcup \{X\}$; and
- (2) for any accumulation, B, and for any nde $\overline{n} \in M(B)$, we can actually recognise that there is a successor, C, of B such that $S(\overline{n}) \in M(C)$.

Wright points out that, under the manageability restriction, it is not guaranteed that for

any accumulations, B and C, and for any equation statement, ϕ , such that $\phi \in E(C)$, there is an accumulation, D, in a sequence in $\Sigma(B)$ such that $\phi \in E(D)$ [27, p.168]. To rephrase, even though $\phi \in E(C)$, there may not be an accumulation, D, after B, such that $\phi \in E(D)$. We could interpret the situation where there is no such an accumulation as indicating that an agent has the 'pool' of recognitional resources for her life-span, and that when she has used up the resources needed for recognising ϕ as a correct statement, by spending resources on other facts instead, she will never be able actually to recognise it^{*37}. The branch where there is no accumulation, D, with $\phi \in E(D)$ is a branch representing a 'sad' possible intellectual history where she misses the correctness of ϕ .

The argument Wright gives for this is not easy to interpret. The following is the best I could offer as an explanation, not an interpretation, for it now. While a model has indefinitely many accumulations, a branch of the tree also has indefinitely many accumulations. A case where there is a branch with an accumulation C such that $\phi \in E(C)$ and another branch without such an accumulation is legitimate, if this happens, because the strict finitist can actually make a one-to-one mapping from the accumulations of the first branch onto those of the second.

4.2.2 Verification-conditions

As we saw in 1.2.2, the central notion of strict finitism is verifiability in practice. Now, since a Wright model is intended to represent all the possible situations, we use the notion of *actual verification* (*verification* for short) inside a model^{*38}. Namely, we here define the conditions under which statements are actually verified at a given accumulation, but one could think that our semantics states the conditions under which statements are verifiable in practice as a whole.

I first state the verification-conditions for atomic formulae, and then those of complex formulae. We write $W \models_B \phi$ for 'In model W, formula ϕ is actually verified at accumulation B of W'. For any $W \in W$, for any $B \in T$, and for any atomic formula ϕ , $W \models_B \phi$ if and only if

- (i) for some $\overline{p}, \overline{q} \in M(B)$, ϕ is $\overline{p} = \overline{q}$ and $\langle \overline{p}, \overline{q} \rangle \in E(B)$;
- (ii) for some $\overline{p}, \overline{q} \in M(B)$, ϕ is $S(\overline{p}) = \overline{q}$ and $\langle \overline{p}, \overline{q} \rangle \in S(B)$;
- (iii) for some $\overline{p}, \overline{q}, \overline{r} \in M(B), \phi$ is $\overline{p} + \overline{q} = \overline{r}$ and $\langle \overline{p}, \overline{q}, \overline{r} \rangle \in P(B)$; or
- (iv) for some $\overline{p}, \overline{q}, \overline{r} \in M(B), \phi$ is $\overline{p} \cdot \overline{q} = \overline{r}$ and $\langle \overline{p}, \overline{q}, \overline{r} \rangle \in K(B)$.

A formula containing variables is actually verified at an accumulation, if and only if its universal closure is actually verified there. For any $W \in \mathcal{W}$, for any $B \in T$, and for any complex

^{*37} Wright describes this by: 'the totality of actually verifiable statements is not stable but *shrinks* as our knowledge advances' [27, p.168, original emphasis].

 $^{^{*38}}$ For the conceptual analysis of actual verification, see 3.1.1.

formula ϕ , $W \models_B \phi$ if and only if

- (i) ϕ is $\psi \wedge \chi$, and it holds both that $W \models_B \psi$ and that $W \models_B \chi$;
- (ii) ϕ is $\psi \lor \chi$, and it holds either that $W \models_B \psi$ or that $W \models_B \chi$;
- (iii) ϕ is $\neg \psi$, and $\forall C \in T [W \not\models_C \psi];$
- (iv) ϕ is $\psi \to \chi$, and $\forall \sigma \in \Sigma(B) \forall C \in \sigma [W \models_C \psi \Longrightarrow \exists \sigma' \in \Sigma(B) \cap \Sigma(C) \exists D \in \sigma' [W \models_D \chi]];$
- (v) ϕ is $\exists x \ [Fx]$, and $\exists \overline{n} \in M(B) \ [W \models_B F\overline{n}]$; or
- (iv) ϕ is $\forall x [Fx]$, and $\forall \overline{n} \forall \sigma \in \Sigma(B) \forall C \in \sigma [\overline{n} \in M(C) \implies \exists \sigma' \in \Sigma(B) \cap \Sigma(C) \exists D \in \sigma' [W \models_D F\overline{n}]].$

We define several notions as follows. I interpret and modify what Wright writes [27, pp.170].

- (i) A formula ϕ is valid in model W (written $W \models \phi$) just in case for any $B \in T_W, W \models_B \phi$.
- (ii) A formula ϕ is valid (written $\models \phi$) just in case for any $W \in \mathcal{W}$, ϕ is valid in W.
- (iii) A formula ϕ is assertible in model W just in case for some $B \in T_W, W \models_B \phi$.
- (iv) A formula ϕ is assertible just in case for any $W \in \mathcal{W}$, ϕ is assertible in W.
- (v) $\Gamma \models \phi$ if and only if $\forall W \in \mathcal{W} \forall B \in T_W \forall \vec{n} \in M_W(B) \left[\forall \psi \in \Gamma[W \models_B \psi(\vec{n})] \Longrightarrow W \models_B \phi(\vec{n}) \right].$

Now, notice that neither is the negation of a statement, A, defined as $(A \to \bot)$ as some other systems do, nor is \bot an accepted symbol in this system.

Also, the definition of the negation itself may appear strange, as Wright mentions: it would allows us to negate a statement just because we cannot actually assess it [27, p.170]. But the above definition should be better, Wright says, than

 $W \models_B \neg \psi$, if and only if $\forall \sigma \in \Sigma(B) \forall C \in \sigma [W \not\models_C \psi]$.

This alternative is what the anti-realist employs for the negation in intuitionistic logic. Wright says that this is not plausible because it allows us to negate statement ψ at accumulation Beven if there is a branch in the model such that ψ is actually verified somewhere on it. This happens when the agent has no recognitional resources to actually verify ψ in one possible history, but in another she has. We, according to the definition, call such ψ 'assertible' in this model. It may seem to be stranger to negate something assertible. [27, p.170]

On another note, I think that, given the definition of assertibility, the notion of verifiability in practice and that of assertibility are treated as if they are identical in this formal framework. As I mentioned in 4.2.1, a model represents all the possible intellectual developments of an agent. What holds at an accumulation is what is actually verified by an agent. Therefore, it seems that a statement, ϕ , is verifiable in practice just in case there is a model which has an accumulation at which it is actually verified. This is equivalent to the condition under which a statement is assertible. It is true that, when we saw the notion of assertibility (2.1.4.1, 3.1.2.3), this notion was said to be weaker than that of verification. But I think that the present formal treatment is suitable, because Wright and we regard 'assertibility' of surveyability predicate 'x is actually intelligible decimally representable' about an object is enough for the argument against the Dedekindian approach (2.1.4).

■4.2.2.1 The lemma Wright and we accept the following, as the 'lemma'.

The lemma $\forall W \in \mathcal{W}[W \models_B P \rightarrow \forall \sigma \in \Sigma(B) \forall C \in \sigma [B \leq C \rightarrow W \models_C P]]$ [27, p.170]

Intuitively, this states that once an agent has actually verified a statement, she keeps the actual verification ever after. This is something the intuitionist can proves for her semantics. But the strict finitist, as Wright says, may not be able to prove this. It is because induction on the complexity of formulae may not be acceptable.

In fact, the definition of validity in a model which I put above is not what Wright puts, but we can prove that they are equivalent, if this lemma is available. Wright's definition is as follows.

 $W \models \phi$ just in case there is a branch, b, such that $\forall B \in b [W \models_B \phi]$. [27, p.170]

Proof I will show that $\forall C \in T_W [W \models_C P]$ is equivalent to that there is a branch in W such that $\forall B \in b [W \models_B P]$.

The direction of \implies is easy: it suffices to take an arbitrary branch.

For the converse, suppose b satisfies that $\forall B \in b [W \models_B P]$. Then $W \models_A P$, where A is the root of T_W . Consider an accumulation $C \in T_W$. Since T_W is a tree, we can say that $A \in \hat{\Sigma}(C)$: $\hat{\Sigma}(C)$ is a sequence on which A lies. Therefore, the lemma allows us to say that $\forall D \in \hat{\Sigma}(C) [W \models_D P]$. Hence $W \models_C P$.

Let us discuss the features of this semantics through thinking the proof system Wright proposes: we will look at the strict finitist logic, and then look back at this semantics. It would be helpful to see first, in particular, that the axiom we will call '(9)' must be rejected.

4.3 Logic

Usually, a logic, or a proof system, is said to specify what the 'valid inferences' are, but in the semantic realism debate, a logic has a status as the guiding principle of how to conceive of the world, because it dictates what must be happening in the world when suppositions about the world are given. When we are given a semantics, we have a logic suitable for it [11, pp.13-4]. If we are given a semantics derived from the position in the semantic realism debate which defeats the other positions, then the logic suitable for the semantics should be the best way of thinking what are happening in the world. This logic, in a sense, reflects the world. The realist Frege stated that the purpose of logic is 'to discern the laws of truth' [15, p.325]. Given that the world is the collection of the statements that hold, or are 'true' for the realist, the logic is the system which provides us with the rules of inference that tell us what are happening in the world.

One can take the fact that the worldviews have their respective logic as a piece of evidence of what is described above. Traditionally, realism has been strongly connected to classical logic, and anti-realism intuitionistic logic [11, p.9]. So, one might expect that the strict finitist who refuses anti-realism will have to pursue a logic different to both.

4.3.1 The partial justification of intuitionistic logic

Wright in 'Strict Finitist' tries to justify axioms and inference rules for the strict finitist logic. It seems that his condition under which an axiom, or an inference rule, is justified is as follows.

The condition of justification An axiom is justified just in case it is valid. An inference rule

$$\phi_0, \cdots, \phi_{n-1}$$

 ψ

is justified just in case, for any $W \in \mathcal{W}$, if $\phi_0, \dots, \phi_{n-1}$ are valid in W, then ψ is assertible in W. [27, p.171]

Wright thinks that the strict finitist can justify all the axioms and all the inference rules of intuitionistic logic, in the following form [27, pp.171-3]. The axioms are

(1) $P \rightarrow (P \lor Q);$ (2) $Q \rightarrow (P \lor Q);$ (3) $(P \land Q) \rightarrow P;$ (4) $(P \land Q) \rightarrow Q;$ (5) $P \rightarrow (Q \rightarrow P);$ (6) $P \rightarrow (\neg P \rightarrow Q);$ (7) $(P \rightarrow Q) \rightarrow ((P \rightarrow (Q \rightarrow R)) \rightarrow (P \rightarrow R));$ (8) $P \rightarrow (Q \rightarrow (P \land Q));$ (9) $(P \rightarrow Q) \rightarrow ((P \rightarrow \neg Q) \rightarrow \neg P);$ (10) $(P \lor Q) \rightarrow ((P \rightarrow R) \rightarrow ((Q \rightarrow R) \rightarrow R));$ (11) $\forall x [Fx] \rightarrow Ft;$ and

(12)
$$Ft \to \exists x [Fx].$$

The rules are

(i) Modus Ponens, i.e.,

$$\frac{P \to Q \qquad P}{Q}$$

(ii)

$$Fy \to P$$
$$\exists x [Fx] \to P$$

(iii)

$$\frac{P \to Fy}{P \to \forall x \, [Fx],}$$

with the eigenvariable condition for (ii) and (iii), that y does not appear freely in the consequences.

However, I have to oppose his opinion, because it seems to me that (i) and (iii) can be justified more easily than he thinks, and (9) cannot be justified, although I do not have a different opinion about any of the other items. In what follows, I will first show that (i) and (iii) can be easily established: but, when I do so, I will show an additional useful theorem, which I call the 'semantic Modus Ponens' (4.3.1.1). Then, I will present a refutation of (9) (4.3.1.2). After that, let us discuss the features of the semantics and the logic, by establishing other basic results (4.3.1.3).

■4.3.1.1 Modus Ponens and (iii) Wright writes, 'With the aid of our pirated lemma, Modus Ponens clearly passes the test' [27, p.171]. But in fact, we can justify (i) without appealing to the lemma.

Proof Suppose that $W \models_B P$ and $W \models_B P \to Q$. We will show that there is an accumulation at which Q is verified. The definition of $P \to Q$ says that $\forall \sigma \in$ $\Sigma(B) \forall C \in \sigma [W \models_C P \Longrightarrow \exists \sigma' \in \Sigma(B) \cap \Sigma(C) \exists D \in \sigma' [W \models_D Q]]$. Look at the leftmost sequence, σ , on which B lies. Then $W \models_B P$ gives us a $\sigma' \in \Sigma(B)$ such that there is an accumulation $C \in \sigma'$ such that $W \models_C Q$.

But if we indeed appeal to the lemma, we can show a slightly stronger and useful result. The above proof only states that if P and $P \to Q$ hold at B, then there is somewhere an accumulation C at which Q holds. The best we know about the location of C is that B and C are on the same branch. Now, determine whether $B \leq C$ or $C \leq B$; and call the largest C': I will hereafter write $C' = \max\{B, C\}$ for this process. Then, by appealing to the lemma, $W \models_{C'} Q$ holds. In sum,

• If $W \models_B P$ and $W \models_B P \to Q$, then there is an accumulation $C \ge B$ such that $W \models_C Q$,

when we accept the lemma. Let us call this the *semantic Modus Ponens*, because this mentions the location of the accumulation, whereas the proof-theoretic Modus Ponens (i.e., (i)) does not.

As for (iii), he presents an argument which justifies (iii) appealing to (5) [27, p.172]. But if I am correct, we can establish this without doing so.

Proof Suppose that $W \models_B P \to Fy$, i.e. $W \models_B \forall y [P \to Fy]$. We will show that $W \models_B P \to \forall x [Fx]$. Consider a $\sigma \in \Sigma(B)$ and a $C \in \sigma$ with $W \models_C P$. Let $C' = \max\{B, C\}$. Notice that $\Sigma(B) \cap \Sigma(C') = \Sigma(C')$. So it suffices to show that $\exists \sigma' \in \Sigma(C') \exists D \in \sigma' [W \models_D \forall x [Fx]]$. In fact, we will show that $W \models_{C'} \forall x [Fx]$, i.e., $\forall \overline{n} \forall \sigma' \in \Sigma(C') \forall D \in \sigma' [\overline{n} \in M(D) \Longrightarrow \exists \tau \in \Sigma(C') \cap \Sigma(D) \exists T \in \tau [W \models_T F\overline{n}]]$.

Consider an \overline{n} , a $\sigma' \in \Sigma(C')$ and a $D \in \sigma'$ such that $\overline{n} \in M(D)$. Let $D' = \max\{C', D\}$. Then, by $W \models_B \forall y [P \to Fy]$ and $B \in \sigma'$ and $D' \in \sigma'$, we can find a $\sigma'' \in \Sigma(B) \cap \Sigma(D') (= \Sigma(D') = \Sigma(C') \cap \Sigma(D))$ and an $E \in \sigma''$ such that $W \models_E P \to F\overline{n}$. Let $E' = \max\{D', E\}$. The lemma gives us $W \models_{E'} P$ and $W \models_{E'} P \to F\overline{n}$. Therefore, by the semantic Modus Ponens, it is guaranteed that we can find an $T \ge E'$ such that $W \models_T F\overline{n}$. Notice that $C' \in \hat{\Sigma}(T)$ and $D \in \hat{\Sigma}(T)$; therefore we can actually recognise that $\hat{\Sigma}(T) \in \Sigma(C') \cap \Sigma(D)$. Hence $W \models_{C'} \forall x [Fx]$. Therefore $W \models_B P \to \forall x [Fx]$. \Box



■4.3.1.2 The refutation of (9) Wright sees 'no obvious way of counter-exemplifying' axiom (9) [27, p.173]. But I think we can present a counter-example to it.

Counter-model Take $\overline{0} + \overline{0} = \overline{0}$ as P, and $\overline{0} = \overline{0}$ as Q. Define a model W as follows: let $A \in T$ be the root of T, and let $\overline{0} \in M(A)$, $E(A) = S(A) = P(A) = K(A) = \emptyset$; call the leftmost successor of A, B_0 , and for any accumulation C such that $C \leq B_0$ or $B_0 \leq C$, let $\overline{0} + \overline{0} \notin M(C)$; call the second successor of A from the left, B_1 , and let $\overline{0} + \overline{0} \in M(B_1)$; and call the leftmost successor of B_1 , C_1 , and let $\langle \overline{0} + \overline{0}, \overline{0} \rangle \in E(C_1)$.

 $\overline{0} + \overline{0}$ never appears



In this model, both $W \models_{B_0} P \to Q$ and $W \models_{B_0} P \to \neg Q$ vacuously hold, because P cannot be expressed at any accumulation D in any sequence on which B_0 lies. On the other hand, $\neg P$ does not hold at any place, because $W \models_{C_1} P$. Therefore (9) does not hold at B_0 , and hence (9) is not valid in this model. Therefore (9) is not valid.

I think that this example well reflects Wright's conception of assertibility nonetheless. To establish the assertibility of P, it is enough to see it verified at an accumulation. But the only way of negating the assertibility – when (9) is rejected by the counter-example – is to survey all the sequences and establish that P does not appear in any accumulation. Let us consider what are verified at an accumulation B, V_B . When we take this as a set, this will be infinite 'in some sense' – since the strict finitist would not understand 'infinity'. But the existence of the negated statement seems to require that one can actually survey each accumulation B, and determine whether whether $P \in V_B$ or not. This might seem to be an excessive demand, but the actual weak decidability of a surveyability predicate implies that an agent can actually determine whether a statement of the predicate is assertible or not: assertibility is bivalent. So, the rejection of (9) is, in my opinion, natural as the result of a formal representation of the strict finitist standpoint. We can say some features of negation as follows. For any model, W, and for any formula, P, (1) either P is assertible (i.e. we can actually find B with $W \models_B P$), or P is not assertible (i.e. there is no B with $W \models_B P$); (2) $\neg P$ holds somewhere $\iff \neg P$ holds everywhere – so, we could say 'negation is global'; (3) P is not assertible $\iff \neg P$ holds somewhere; (4) P is assertible $\iff \neg P$ holds nowhere.

■4.3.1.3 Other semantic results The rejection of axiom (9) and the reasons for it give us interesting questions to ask. We know that intuitionistic logic is sound and complete for Kripke semantics. So with (9) rejected, at least the strict finitist proof system will not have all the intuitionistically valid statement proved. Then, will all the strictly finitistically valid statements be proved in intuitionistic logic? My answer is no, because $\neg P \lor \neg \neg P$ is strictly finitistically valid, while intuitionistically invalid. But it will be informative to think first about other statements which include negation.

First, the law of excluded middle, i.e. $P \vee \neg P$, is not strict finitistically valid.

Counter-model Take $\overline{0} = \overline{0}$ as P. Define a model W as follows: let $A \in T$ be the root of T, and let $\overline{0} \in M(A)$, $E(A) = S(A) = P(A) = K(A) = \emptyset$; call the leftmost successor of A, B; let B satisfy that $\langle \overline{0}, \overline{0} \rangle \in E(B)$.

$$\begin{array}{ccc} B & & \langle \overline{0}, \overline{0} \rangle \in E(B) \\ \\ \\ \\ A & & \langle \overline{0}, \overline{0} \rangle \notin E(A) \end{array}$$

Then, $W \not\models_A P$. But $W \models_B P$ tells us that $\neg P$ does not hold anywhere. Therefore $W \not\models_A \neg P$. $\therefore W \not\models_A P \lor \neg P$.

Also, we can show that the double-negation elimination, i.e., $\neg \neg P \rightarrow P$, is not strict finitistically valid.

Counter-model Take $\overline{0} + \overline{0} = \overline{0}$ as P. Define a model W as follows: let W has the same A and B as the previous example; let the second successor of A from the left, C; for any accumulation $C' \ge C$, let $\langle \overline{0}, \overline{0} \rangle \notin E(C')$



Then, we can show that $W \not\models_C \neg \neg P \rightarrow P$. Suppose otherwise. Then, $W \models_B P$ tells us that $\neg P$ holds nowhere; hence $W \models_C \neg \neg P$; then we must be able to find a $\sigma \in \Sigma(C)$ and $D \in \sigma$ with $W \models_D P$; but this is contradictory to the definition of C.

This is an interesting result, because, as I said as a 'feature of negation', for any model, for any statement, P, we know that P is assertible $\iff \neg P$ holds nowhere. Of course, $\neg P$ holds nowhere $\iff \neg \neg P$ is assertible. So, it is indeed correct that P is assertible $\iff \neg \neg P$ is assertible. Therefore the assertibility of $\neg \neg P$ guarantees the assertibility of P. But this result says that when we consider an accumulation at which $\neg \neg P$ holds, P may not occur in the branch where the accumulation is located. In other words, while negation has an global effect on a model, neither conditional nor affirmation does not. This result, I think, makes this contrast explicit.

Now it is indeed correct that for any model, W, and for any statement, P, P is assertible $\iff \neg \neg P$ is assertible. It may be worthwhile to give an explicit proof.

Proof The direction of \implies is obvious.

For the converse, we appeal to the bivalence of assertibility: either P is assertible or P is not assertible. If P is assertible, then the conditional $\neg \neg P$ is assertible $\implies P$ is assertible, vacuously holds good. On the other hand, if P is not assertible, then the antecedent of the conditional to be established is contradictory.

We can show that for any model, for any statement, P, the double-negation of the law of excluded middle, i.e., $\neg \neg (P \lor \neg P)$, is valid.

Proof We appeal to the bivalence of the assertibility of P: either P is assertible or not. If P is assertible, then we can actually find an accumulation at which P holds. Then, $P \lor \neg P$ holds at the accumulation. Therefore $\neg(P \lor \neg P)$ holds nowhere. Hence $\neg \neg(P \lor \neg P)$ holds everywhere. On the other hand, suppose P is not assertible. Then $\neg P$ holds everywhere. Therefore $P \lor \neg P$ holds everywhere. Hence $\neg \neg(P \lor \neg P)$ holds everywhere.

Now, let us see the proof of the validity of $\neg P \lor \neg \neg P$. The results so far might seem to

suggest that for any statement, if it is strict finitistically valid, then it is intuitionistically valid. However, this is not the case.

Proof We appeal to the bivalence of the assertibility of P. If P is assertible, $\neg P$ holds nowhere: therefore $\neg \neg P$ holds everywhere; hence $\neg P \lor \neg \neg P$ holds everywhere. If P is not assertible, then $\neg P$ holds everywhere; therefore $\neg P \lor \neg \neg P$ holds everywhere. \Box

This result makes use of the feature of negation that it is global, in combination with the bivalence of assertibility. The bivalence guarantees the assertibility of *either* of P, or the nation of P. Once the negation is assertible, it holds everywhere. On the other hand, the affirmation of P gives rise to the assertibility of $\neg \neg P$: and this assertibility of a negative statement guarantees that it holds everywhere.

4.4 Mathematics

In the framework of the semantic realism debate which this thesis adopts, mathematics is one aspect of the world: an axiom system of mathematics is the system that allows us to probe that aspect of the world by means of the accepted logical inferences. The axioms would be the 'bold' fundamental statements that purport to declare that they grasp the basic features of the world. The theorems will be the pieces of knowledge we have attained about the mathematical states of affairs that hold good in the world.

To see what the strict finitist mathematics looks like is to pin down the strict finitistically acceptable axioms about mathematical expressions, and to grasp the implications of them. In what follows, we will first (4.4.1) discuss what Wright proposes about arithmetic in the appendix of 'Strict Finitism' [27, pp.173-5]; and then try to imitate intuitionistic mathematics (4.4.2).

4.4.1 Arithmetic

We start with arithmetic as the most basic part of mathematics. Wright considers the axioms of intuitionistic arithmetic. Peano axioms are

(A) $\exists x [x = \overline{0}];$ (B) $\forall x \exists y [y = S(x)];$ (C) $\forall x [0 \neq S(x)];$ (D) $\forall x, y [S(x) = S(y) \rightarrow x = y];$ and

(E) $(\forall x [Fx \to FS(x)] \land F\overline{0}) \to \forall x [Fx].$

Axioms of identity are

(F) $\forall x [x = x];$ (G) $\forall x, y [x = y \rightarrow t(x) = t(y)];$ and (H) $\forall x, y [x = y \rightarrow (Fx \rightarrow Fy)],$

where t is any term forming operator. Axioms for addition and multiplication are

- (I) $\forall x [x + \overline{0} = x];$
- (J) $\forall x, y [x + S(y) = S(x + y)];$
- (K) $\forall x [x \cdot \overline{0} = \overline{0}];$ and
- (L) $\forall x, y [x \cdot S(y) = x \cdot y + x].$

Wright makes three observations about these (candidates of) axioms. Let us see them in order.

■4.4.1.1 Wright's first observation The first observation is about what will be implied if we adopt the axioms about successor as valid. Wright writes:

[...] the assertibility of the axioms for zero and successor, (A) to (D) will require that the manageability of an accumulation, B, always tolerates the addition to M(B) of $S(\overline{n})$ for any \overline{n} in M(B) and that such an addition is invariably actually possible. [27, p.174, the notation is modified]

I understand this as suggesting that the manageability restriction imply that for any accumulation, B, and for any nde, $\overline{n} \in M(B)$, there is an actually recognisable successor, C, of B such that $S(\overline{n}) \in M(C)$; and I in fact proposed my version of manageability restriction already incorporating this requirement. I agree with that especially admitting (D) most clearly requires this condition.

■4.4.1.2 Wright's second observation Wright thinks that the axiom, (E), of induction cannot be validated. To explain, he puts a putative proof of the validity of (E). I reconstruct it.

'Proof' Suppose $W \models_B F\overline{0}$ and $W \models_B \forall x [Fx \to FS(x)]$.

Then, in general, it holds that for any nde \overline{n} , for any $\sigma \in \Sigma(B)$ and for any $C \in \sigma$ with $\overline{n} \in M(C)$, we can find a $\sigma' \in \Sigma(B) \cap \Sigma(C)$ and a $D \in \sigma'$ such that $W \models_D F\overline{n} \to FS(\overline{n})$. Therefore, if $W \models_B F\overline{n}$ holds, then the lemma and the semantic Modus Ponens gives us an accumulation $E \ge D (\ge B)$ such that $W \models_E FS(\overline{n})$.

Now, ex hypothesi, $W \models_B F\overline{0}$. Therefore we can find an $E_0 \ge B$ with $W \models_{E_0} FS(\overline{0})$. Because $\hat{\Sigma}(E_0) \in \Sigma(B)$ and $S(\overline{0}) \in M(E_0)$, by using the same argument, we can find an $E_1 \ge E_0$ such that $W \models_{E_1} FSS(\overline{0})$. By repeating this process, we get accumulations, E_0, E_1, E_2, \cdots , such that $B \le E_0 \le E_1 \le E_2 \le \cdots$ and $W \models_B F\overline{0}, W \models_{E_0} FS(\overline{0}), W \models_{E_1} FSS(\overline{0}), W \models_{E_2} FSSS(\overline{0}), \cdots$ hold.

Therefore $W \models_B \forall x [Fx].$

Wright points out that this 'proof' has a difficulty. Namely, we can advance to the conclusion that $W \models_B \forall x [Fx]$ only if it is guaranteed *either* that the only nde's that appear at any $B \in T$ are canonical, or that for any non-canonical nde, \overline{n} , that appears at any accumulation, B, $W \models_C \overline{n} = \overline{c}$ holds for some canonical nde \overline{c} and for some accumulation C with $C \leq B$ or $B \leq C$. The former is already violated by other axioms of arithmetic. If the latter case held, Wright says, not every actually intelligible nde could feature in T, although one may want to pay this cost to save induction. [27, p.174]^{*39}

I agree with the view that the putative proof has an unacceptable leap as Wright says. To me, however, it is not clear why the latter modification makes it the case that not every actually intelligible nde can appear in T, and why it is problematic. But putting this point aside, it appears to be more plausible to modify the induction axiom. I propose (E') as follows:

(E') $(\forall x [Fx \to FS(x)] \land F\overline{0}) \to \forall \overline{c} [F\overline{c}],$

where \overline{c} ranges over the canonical nde's. Also, to make (E') hold, it would need to hold that

for any accumulation $B \in T$, and for any canonical nde, $\overline{c} \in M(B)$, we can actually find a successor, C, of B such that $S(\overline{c}) \in M(C)$.

But this is already incorporated into the manageability restriction.

■4.4.1.3 Wright's third observation Wright has opinions against (J) and (L): he thinks they should be negated. The reason is that acceptance of them will require that the agent whose recognitional ability is represented by a model at issue can actually compute any big numbers. Let us see what we will have to accept when we accept (L), for instance.

Suppose (L) is valid and holds at $B \in W$. Then for any nde's \overline{n} and \overline{m} and any $\sigma \in \Sigma(B)$ and any $C \in \sigma$ such that $\overline{n}, \overline{m} \in M_C$, it is guaranteed that there is a sequence $\sigma' \in \Sigma(B) \cap \Sigma(C)$ such that we can actually find an accumulation $D \in \sigma'$ with $W \models_D \overline{n} \cdot S(\overline{m}) = \overline{n} \cdot \overline{m} + \overline{n}$.

Here, it seems plausible that the verification of $\overline{n} \cdot S(\overline{m}) = \overline{n} \cdot \overline{m} + \overline{n}$ requires (1) a computation of the canonical nde $\overline{c_0}$ such that $\overline{c_0} = \overline{n} \cdot S(\overline{m})$, (2) a computation of the canonical nde $\overline{c_1}$ such that $\overline{c_1} = \overline{n} \cdot \overline{m}$, and (3) a verification that $\overline{c_0} = \overline{c_1} + \overline{n}$. But it also seems plausible that these computations cannot be actually done in the cases where the numbers denoted by \overline{n} and \overline{m} (i.e. n and m) are too big. Therefore, there must not be, in any model, an accumulation where the nde's for too big numbers like n and m appear. This requirement does not seem to be met in any model, because the manageability of any accumulation allows the addition of

^{*&}lt;sup>39</sup> By the way, this is the very place where Wright mentions the notion of canonical nde's for the first time. In this thesis, I introduced this notion when I defined the language (4.1).

any single actually intelligible nde; and because any such addition is always humanly feasible. Therefore, (L) should be rejected; and so will (J) be. [27, pp.174-5]^{*40}

My current opinion about this view is that I am sympathetic to that we cannot accept (L), but I have no solid reason to reject (J). Wright's argument above, I think, is surely plausible to some extent. Let us write \mathcal{D}^+ for the decimal notation supplemented with addition, and \mathcal{D}^{\times} for that with multiplication. To reject (L) thinking that it requires actually impossible computation, is essentially the same as to reject \mathcal{D}^{\times} (or, the successor notation with multiplication) thinking that it requires actually impossible computation. As long as we use the decimal notation (or the successor notation) as the standard, both of them require us to be able actually to compute and grasp the natural numbers that are actually intelligibly representable using multiplication. So, if one rejects taking \mathcal{D}^{\times} as outrageous, then she should have reason to reject (L).

However, now I have no *a priori* conclusive reason to reject \mathcal{D}^{\times} , or \mathcal{D}^+ . I have never been able to justify the rejection of either. I would reject by the same reason the notation which allows the notation of subtraction and that of exponentiation, but it is only because we have the running supposition that $2^{1000^{1000}} - 1$ is actually decimally unintelligible. Whether a notation is acceptable may depend on our, *a posteriori*, computation abilities. The safest opinion here seems to be the conditional claim that if one has reason to reject \mathcal{D}^{\times} (or \mathcal{D}^+), she should reject (L) (or (J)).

One thing here to notice is that, the strict finitist's arithmetic will become, if we agree to Wright's view, an arithmetic where neither multiplication nor addition is allowed. Surely, it will still contain (I) and (K), but they are merely stipulations about how to compute +0 and $\times 0$. To me, a system with such scarce recourses for arithmetic does not seem to be worthy of being called 'arithmetic'. I suggest that we first defend the actual intelligibility of \mathcal{D}^+ and allow the strict finitist to compute addition, but this task is, unfortunately, beyond this thesis.

4.4.2 Intuitionistic mathematics

In this final portion of the thesis, I put speculations about how the strict finitist could imitate intuitionistic mathematics. If one wishes for a new framework to rise in mathematics, she must clarify to what extent the results of the extant frameworks will be maintained. The strict finitist claims that her standpoint is the right one, whereas anti-realism is not. So one would wonder how much the strict finitist mathematics could preserve intuitionistic mathematics. Below, I present one way the strict finitist might imitate a very simple example of the objects which the intuitionistic mathematicians call *spread*.

^{*40} I use the notation of this thesis. Also, since the formula for $\overline{c_0}$ seems to be a typo, I changed it.
■4.4.2.1 The idea In this thesis, I understand as follows: we call a *spread-law*, a decision procedure with which, for any infinite sequence, α , of natural numbers, the anti-realist (or the intuitionistic mathematician) can in principle determine whether α is admitted or rejected; and for any spread-law, its spread is the set of infinite sequences of natural numbers which the anti-realist can in principle recognise to be admitted by the spread-law. For example,

$$\mathcal{M}_2 = \{ \alpha : \forall n \in \mathbb{N} \left[\alpha(n) \in \{0,1\} \land \alpha(n) \le \alpha(n+1) \right] \}^{*41}$$

is a spread, because for any infinite sequence, α , of natural numbers, it is decidable in principle whether for any natural number, n, $\alpha(n) \in \{0, 1\}$ and $\alpha(n) \leq \alpha(n+1)$, or not.

The key of my idea is that the strict finitist can deal with a tree with indefinitely many nodes, just as the tree of a Wright model is. When a spread is as simple as \mathcal{M}_2 , even the strict finitist seems to be able to deal with it. To imitate \mathcal{M}_2 , define a tree T_2 as follows: call the root A; let A have three successors, B_0 , B_1 and C_0 ; let C_0 have the *left successor* D_1 and the *right successor* C_1 ; and as long as n is actually decimally recognisable, let C_n have the left successor D_{n+1} and the right successor C_{n+1} . Notice that if n is actually decimally recognisable, so is n + 1 because of the tolerance.



Then, define a *label function* f from the nodes of T_2 to indefinitely long sequences (or shorter than them) as follows:

(1) $f(A) = \langle \rangle;$

 $^{^{*41}}$ We touched upon this example in 3.1.2.2.

- (2) $f(B_0) = \langle \overline{\underline{0}} \rangle;$
- (3) $f(B_1) = \langle \overline{\underline{1}} \rangle;$
- (4) $f(C_0) = \langle \overline{0} \rangle;$
- (5) As long as n is actually decimally recognisable, $f(C_{n+1}) = f(C_n) * \langle \overline{0} \rangle$; and
- (6) As long as n is actually decimally recognisable, $f(D_{n+1}) = f(C_n) * \langle \overline{1} \rangle$;



I intend that for any actually decimally representable n, \overline{n} is an nde (or the nde in the decimal notation in the standard case) of n, and \overline{n} stands for the infinite sequence made of \overline{n} . But \overline{n} is nonetheless a single sign. So a sequence is indefinitely long at most even if it contains \overline{n} : e.g., $\langle \overline{0}, \overline{0}, \overline{1} \rangle$ is a 3-digit sequence.

I assert that the strict finitist can actually recognise each node of this T_2 , because (1) T_2 has only indefinitely many nodes of an (at most) indefinitely long sequence, and (2) each node, C', satisfies the manageability condition in the sense that, if it is a successor of C, then the label of C' is made by adding one and only one element to the label of C. Also T_2 , I think, gives the strict finitist what she should understand about \mathcal{M}_2 . It tells (1) $\overline{0}$, (2) $\overline{1}$, and (3) for each actually decimally representable $n, \overline{0}n * \overline{1}$.

■4.4.2.2 The formalisation In the semantic realism debate, to treat things as objects is to have terms which denote them and build a semantic theory. So let us try formalising the semantics related to the sequences which belong to \mathcal{M}_2 . In the following, I discuss using only canonical nde's instead of including nde's in general. This is because, as we saw in 4.4.1.3, we have not yet attained a satisfactory theory about addition (and multiplication): most of our

nde's are canonical^{*42}.

We define sequence-denoting expressions ('sde's' for short) specific to the current example case. We use the language, L, we defined in 4.1 and amplify it. The sde-formation rules of Lare as follows: when we abbreviate $S(\overline{0})$ as $\overline{1}$,

- (i) $\overline{\underline{0}}$ is an sde;
- (ii) $\overline{\underline{1}}$ is an sde; and
- (iii) for any canonical nde, \overline{c} , $\overline{\underline{0}}\overline{c} * \overline{\underline{1}}$ is an sde.

Also, we have to have indefinitely many *sde-variables*, ξ_0, ξ_1, \cdots . These sde's and sde-variables now count as our terms as well. Note that, as we saw, there are only indefinitely many sde's now. Using sde's, we can derive a new kind of nde:

(1) If α is an sde, \overline{c} is a canonical nde, and x is an variable, then $\alpha \overline{c}$ and αx are nde's.

Intuitively, $\alpha \overline{c}$ denotes the *c*-th member of α . Next, we define sde-related atomic formulae:

(A) If α, β are sde's, and ξ, ζ are sde-variables, then $\alpha = \beta, \alpha = \xi, \xi = \alpha, \xi = \zeta$ are atomic formulae.

Then, tentatively, let us modify the definition of a Wright model, $W = \langle T_W, M_W, E_W, S_W, P_W, K_W \rangle$, so that, for any $B \in T_W$,

- (1) $M_W(B)$ can contain $\overline{\underline{0}}$, if $\overline{\overline{0}} \in M_W(B)$;
- (2) $M_W(B)$ can contain $\overline{\underline{1}}$, if $\overline{\underline{1}} \in M_W(B)$;
- (3) for each canonical nde \overline{c} , $M_W(B)$ can contain $\overline{0}\overline{c} * \overline{1}$, if $\overline{0} \in M_W(B)$, $\overline{1} \in M_W(B)$ and $\overline{c} \in M_W(B)$;
- (4) for each sde α , and for each canonical nde \overline{c} , four sets $E_W(B), S_W(B), P_W(B), K_W(B)$ can contain tuples which include $\alpha \overline{c}$ as a member if $\alpha \in M_W(B)$ and $\overline{c} \in M_W(B)$; and
- (5) for each sde α , and for each variable x, four sets $E_W(B), S_W(B), P_W(B), K_W(B)$ can contain tuples which include αx as a member if $\alpha \in M_W(B)$.

We are going to define the verification-conditions for the sde-related formulae as follows. But for this task, we need the definitions concerning inequality. We introduce new predicate symbols, < and \leq . For any canonical nde's $\overline{c_0}, \overline{c_1}$, for any variables x and $y, \overline{c_0} < \overline{c_1}, \overline{c_0} < x, x < \overline{c_0}$ and x < y (also those with \leq) are now atomic formulae. For any model W, and for any accumulation $B \in T_W$, we define

(i) for any canonical nde's c_0 and c_1 , $W \models_B c_0 < c_1 \iff$ there are less than indefinitely

^{*42} To confess, another reason is that I cannot present a plausible definition of inequality with nde's in general.

many canonical nde's d_0, \dots, d_n such that $\langle c_0, d_0 \rangle, \langle d_0, d_1 \rangle, \dots, \langle d_n, c_1 \rangle \in S_W(B)$;

(ii) for any canonical nde's c_0 and c_1 , $W \models_B c_0 \leq c_1 \iff either$ there are less than indefinitely many canonical nde's d_0, \dots, d_n such that $\langle c_0, d_0 \rangle, \langle d_0, d_1 \rangle, \dots, \langle d_n, c_1 \rangle \in$ $S_W(B), or \langle c_0, c_1 \rangle \in E_W(B);$

We define, for any model W, and for any accumulation $B \in T_W$,

- (i) for any canonical nde \overline{c} , $W \models_B \overline{\underline{0}}\overline{c} = \overline{0} \iff \overline{c} \in M_W(B)$ and $\overline{\underline{0}} \in M_W(B)$;
- (ii) for any canonical nde \overline{c} , $W \models_B \overline{\underline{1}}\overline{c} = \overline{1} \iff \overline{c} \in M_W(B)$ and $\overline{\underline{1}} \in M_W(B)$;
- (iii) for any canonical nde's $\overline{c_0}$ and $\overline{c_1}$, $W \models_B (\overline{\underline{0}c_0} * \overline{\underline{1}})\overline{c_1} = \overline{0} \iff \overline{c_0} \in M_W(B), \overline{c_1} \in M_W(B), \overline{\underline{0}} \in M_W(B), \overline{\underline{1}} \in M_W(B)$ and $W \models_B \overline{c_1} < \overline{c_0}$; and
- (iv) for any canonical nde's $\overline{c_0}$ and $\overline{c_1}$, $W \models_B (\overline{\underline{0}} \overline{c_0} * \overline{\underline{1}}) \overline{c_1} = \overline{1} \iff \overline{c_0} \in M_W(B), \overline{c_1} \in M_W(B), \overline{\underline{0}} \in M_W(B), \overline{\underline{1}} \in M_W(B)$ and $W \models_B \overline{c_0} \leq \overline{c_1}$.

This is my proposal of the strict finitist formal theory of \mathcal{M}_2 . For the arguments about the plausibility and the theorems of this formalisation, however, I have to wait for another opportunity for investigation.

5. Summary and the topics left untouched

All the above is the substantial part of this thesis. Now, let us review what we have been discussing and what we have been trying to establish (5.1). This thesis will end with some problematic points I have to leave unsolved and a short list of the literature I should have incorporated.

5.1 Summary

In this thesis, I investigated the plausibility of the standpoint called strict finitism and attempted to establish the cogency of this position. In section 1, I introduced this position as that which opposes realism and anti-realism in the semantic realism debate. I started with displaying the general framework of the semantic realism debate as a research area of philosophy about the reality of the world (1.1.1), then described realism as semantic realism, a specific semantic stance about statements (1.1.2), and introduced anti-realism enter as a traditional alternative to semantic realism (1.1.3). After that, in order to explain how strict finitism can enter the picture, I gave a general explanation of the two kinds of anti-realist attack on realism (1.2.1). Strict finitism then got the characterisation as a semantic standpoint which attacks anti-realism, using the same type of attack which the anti-realist uses to attack on realism (1.2.2).

I argued for strict finitism with two directions: one is to attack anti-realism (the negative programme); the other is to defend strict finitism itself (the positive programme). Section 2 was where I engaged in the negative programme: I attacked anti-realism, by reconstructing and examining Wright's paper 'Strict finitism' [27].

In the former part of this section (2.1), I tried to show that general anti-realism does not hold. This is the view that we should take the anti-realist stance for any aspect of the world. To defeat this, we examined the central notion of anti-realism, i.e., that of decidability in principle; and argued that it cannot be explained so that the strict finitism could understand. This notion may be analysed into that of finitude and that of decidability in practice^{*43}. But finitude cannot be explained without appealing to the notion of possibility in principle, the general notion of decidability in principle. Possibility in principle cannot be clear unless so is decidability in principle. Therefore the notions which the anti-realist employs could be understood only by those who have already understood them. To establish this conclusion, we examined three approaches which the anti-realist may take, and saw all of them fail (2.1.1, 2.1.2, 2.1.3). Further, we saw that when we suppose some strict finitist notions, the third

 $^{^{*43}}$ See the explanation scheme (P) in 2.1.

approach incurs a serious difficulty (2.1.4).

In the latter part of this section (2.2), I tried to block the way arithmetical anti-realism could hold. Using Wright's ideas as the guide, I specified how the anti-realism about arithmetic may insist on the bivalence of all finite statements: for the scheme of the crucial transition, see 2.2.1. Wright's reply to this anti-realist's strategy seemed to me to be an overreaction, and in fact he, in 'Strict Finitism', does not wish to be conclusive about this issue [27, p.151]. My reply to the anti-realist's ideas was that even if her ideas are correct, the strict finitist does not have to admit the bivalence of all the finite statements, because some are unintelligible and not legitimate statements in the first place (2.2.2).

I embarked on the first part of the positive programme in section 3. This was the place where I tried to give firm grounds to strict finitism by justifying the features of the concepts the strict finitist uses and by defending the consistency of strict finitism. In 3.1, I discussed what verifiability in practice is (3.1.1), why a surveyability predicate is tolerant, non-vague, and actually weakly decidable (3.1.2), what notation the strict finitist will use (3.1.3), and what the strict finitist numbers are (3.1.4). Using these concepts, I argued against the contention that strict finitism should suffer what is called 'Wang's paradox' (3.2). After viewing how this paradox could be presented (3.2.1), we saw the solution to this paradox proposed by Wright (3.2.2). This is a defensive argument from the strict finitist's perspective, and I regarded his strategy as basically correct. However, it lacked a crucial piece: Wright's solution appealed to a system of notation with a specific feature, but he only said that such a systemy should naturally exist, and provided no substantial argument. It was my task in section 3.2.3 to give an argument to show that the decimal notation possesses the required property.

The formalisation of the strict finitist view was the aim of section 4. This was the second part of the positive programme of strict finitism: so far, this standpoint had been treated as a contentious alternative to anti-realism, but given the confirmed grounds (in 3.), we could now investigate what the formal theories of strict finitism would be. We at first defined a language for arithmetic (4.1), and traced the ideas for the formalisation presented by Wright: we saw the definition of a model for semantics (4.2.1), and the verification-conditions for the strict finitist semantics (4.2.2). In the framework of the semantic realism debate, a worldview determines its semantics, and a semantics does its logic: so we saw the logic which the strict finitist semantics dictates (4.3). Wright presented conditions according to which an axiom and an inferential rule are justified, but I pointed out that one of the axioms he thought to be justified was, in fact, not justified. The final part of section 4, and of the thesis, was about the strict finitist mathematics (4.4). We investigated arithmetic as the basic part of mathematics (4.4.1.3). I ended the substantial part of the thesis with my speculations on the strict finitist theory to simulate intuitionistic mathematics (4.4.2).

5.2 The topics left untouched

To end my thesis, I have many issues which I must leave unsure or unsatisfactory. Below, I will name only a few. (1) I could not propose a total counter-argument against the antirealist's contention that finite statements are bivalent (2.2.1, 2.2.2). While judging Wright's reply as an overreacting, I could only propose a compromise with the anti-realist ideas.

(2) I could not find plausible arguments for the incompleteness of the decimal notation supplemented with addition, multiplication and exponentiation. Dummett says this is clearly incomplete; and so it appears to me. But I had to leave this as an unsettled matter (3.1.3).

(3) I fear that my argument that the decimal notation has no actually recognisable supersessor may be too unclear and bold. It was an argument stemming from the idea that the strict finitist can stick to her standard notation (i.e. the decimal notation). But I may have to be more careful about how to treat the other notations (3.2.3).

(4) Also, I should mention the formal theories. I have to leave the 'lemma' as a mere assumption (4.2.2.1). We saw that induction could be justified within the formal theory, if we focus only on the canonical nde's (4.4.1.2). But I could not find a plausible way of arguing for the connection between the formalised induction and the induction in the meta-language.

(5) The last issue I name here is the strict finitist theory of arithmetic: as it is now, the strict finitist arithmetic cannot deal with any substantial addition (4.4.1.3). I think that the strict finitist should take this issue seriously. Apparently, any hard-nosed strict finitist should admit that 2 + 2 = 4 is a perfectly safe arithmetical statement, but within the current theory it is not, because the legitimacy of this statement requires 4 + 4 = 8 and 8 + 8 = 16, etc., and eventually numbers which are not actually intelligible should enter the theory. This issue must be connected with my poor treatment of the notation which is not our standard. The strict finitist should invent a conception of the notations which does not simply discard non-standard notations.

I have to end this thesis without examining many other articles than Wright's 'Strict Finitism' ([27]) and Dummett's 'Wang's Paradox' ([7]). Especially, it is known that Wright published an article ([30]) about Dummett's 'Wang's paradox' again, 25 years after 'Strict Finitism', and Dummett gave a response ([14]) to it.

Also, I had no room to discuss the contributions offered by other writers. I here note only three articles: (1) Yessenin-Volpin's 'The Ultra-Intuitionistic Criticism' in 1970, as one of the earliest proponents of strict finitism [23]; (2) Mawby's dissertation *Strict Finitism as a Foundation for Mathematics* in 2005, as an extensive examination of strict finitism as a standpoint in the semantic realism debate [17]; and (3) Nelson's 'Internal Set Theory' in 1977, as a clue to a mathematical approach to regard small numbers as standard natual numbers and too big numbers as non-standard natural numbers [19].

References

- Brouwer (1952), 'Historical Background, Principles and Methods of Intuitionism', reprinted in Heyting (ed.), L. E. J. Brouwer: Collected Works, Vol.1, 1975, North-Holland Publishing Company, pp.508-15.
- [2] Devitt, Realism and Truth, 2nd edn., 1991, Princeton University Press.
- [3] Dummett (1959), 'Truth', reprinted in his Truth and Other Enigmas, 1978, Harvard University Press, pp.1-24.
- [4] Dummett (1963), 'Realism', reprinted in his *Truth and Other Enigmas*, 1978, Harvard University Press, pp.145-65.
- [5] Dummett (1967), 'Platonism', reprinted in his *Truth and Other Enigmas*, 1978, Harvard University Press, pp.202-14.
- [6] Dummett (1973), 'The Philosophical Basis of Intuitionistic Logic', reprinted in his Truth and Other Enigmas, 1978, Harvard University Press, pp.215-47.
- [7] Dummett (1975), 'Wang's Paradox', reprinted in his *Truth and Other Enigmas*, 1978, Harvard University Press, pp.248-68.
- [8] Dummett (1976), 'What is the Theory of Meaning? (II)', reprinted in his *The Sea of Language*, 1993, Oxford University Press, Ch.2, pp.34-93.
- [9] Dummett, Truth and Other Enigmas, 1978, Harvard University Press.
- [10] Dummett (1982), 'Realism', reprinted in his *The Sea of Language*, 1993, Oxford University Press, Ch.11, pp.230-76.
- [11] Dummett, The Logical Basis of Metaphysics, 1991, Harvard University Press.
- [12] Dummett, Truth and the Past, 2004, Columbia University Press.
- [13] Dummett (2007), 'Reply to Wolfgang Künne', in Auxier and Hahn (eds.), The Philosophy of Michael Dummett, Open Court Publishing Company, 2007, pp.345-50.
- [14] Dummett (2007), 'Reply to Crispin Wright', in Auxier and Hahn (eds.), The Philosophy of Michael Dummett, Open Court Publishing Company, 2007, pp.445-54.
- [15] Frege (1918-9), 'Thought', in Beaney (ed.), *The Frege Reader*, translated by Geach and Stoothoff, 1997, Blackwell Publishers, pp.325-45.
- [16] Kripke, Wittgenstein on Rules and Private Language: An Elementary Exposition, 1982, Harvard University Press.
- [17] Mawby (2005), Strict Finitism as a Foundation for Mathematics, his dissertation submitted to Glasgow University.
- [18] Moore (ed.), Infinity, 1993, Dartmouth.
- [19] Nelson (1977), 'Internal Set Theory: A New Approach to Nonstandard Analysis', in Bulletin of the American Mathematical Society, Vol.83, No.6, pp.1165-98
- [20] Parsons (2005), 'Truthmakers, the Past, and the Future', in Beebee and Dodd (eds.) Truthmakers: The Contemporary Debate, 2005, Clarendon Press, pp.161-74.
- [21] Shieh (1998), 'Undecidability in Anti-realism', in *Philosophia Mathematica*, Vol.6, pp.324-333.
- [22] Yamada (2012), 'Two Realisms: An Analysis from the Viewpoint of 'Worldview" [in Japanese], in *Tetsugaku-Ronso*, No.39, pp.98-109.
- [23] Yessenin-Volpin (1970), 'The Ultra-Intuitionistic Criticism and the Antitraditional Program for Foundations of Mathematics', in Kino, Myhill and Vesley (eds.), *Intuitionism and Proof Theory*,

1970, North-Holland Publishing Company, pp.3-45.

- [24] Veldman (2016-7), 'Notes on Intuitionistic Mathematics', the material for his course in 2016-7.
- [25] Wittgenstein, *Philosophical Investigations*, translated by Anscombe, Hacker, and Schulte, revised 4th edition by Hacker and Schulte, 2009, Blackwell Publishing.
- [26] Wright, Wittgenstein on the Foundations of Mathematics, 1980, Harvard University Press.
- [27] Wright (1982), 'Strict Finitism', reprinted in his *Realism, Meaning and Truth*, 2nd edition, 1993, Blackwell, Ch.4, pp.107-75.
- [28] Wright (1986), 'Rule-Following, Meaning and Constructivism', reprinted in his Rails to Infinity: Essays on Themes from Wittgenstein's Philosophical Investigations, 2001, Harvard University Press, Ch.3, pp.53-80.
- [29] Wright (1993), 'Introduction' of his *Realism, Meaning and Truth*, 2nd edition, 1993, Blackwell, pp.1-43.
- [30] Wright (2007), "Wang's Paradox", in Auxier and Hahn (eds.), The Philosophy of Michael Dummett, 2007, Open Court Publishing Company, pp.415-44.