Universiteit Leiden
The Netherlands

A Priori truth in the natural world : a non-referentialist response to Benacerraf's dilemma

Novák, Z.

CHAPTER 1

Conceptual Preliminaries: Elucidation of Key Concepts

Introduction

In this work, I shall argue for a naturalistic construal of *a priori* truth and knowledge acquisition. The core tenet that I shall defend in the following chapters is that there is a clear semantical contrast between those claims that we typically distinguish as *a priori* and empirical, which can give rise to a substantive account of the traditional epistemological distinction between *a priori* and empirical knowledge, justification and evidence. The way I shall defend this claim is direct and simple: I shall articulate what exactly this semantical contrast consists in by providing a substantive characterisation of the nature of *a priori* truth.

Although many philosophers are sceptical about the existence of *a priori* knowledge and justification, and there is also much debate over the proper construal of apriority, virtually no one refuses the application of the term for expressing substantive positions in epistemology. This indicates that most philosophers have some, more or less explicit, idea of what apriority consists in. My primary aim in this work is to elaborate an account that preserves the categorical distinction between *a priori* and empirical beliefs, largely observes the received application of the contrasted terms, and, together with our best empirical theories of the world and human cognition, explains the most important characteristics of our purportedly *a priori* claims.

The construal that I shall propose is based on the theoretical hypothesis that the clearest examples of *a priori* claims or beliefs, our claims or beliefs in pure logic and mathematics, are attitudes to truth-apt representations of abstract domains, whose truth

value is nevertheless determined by the obtaining (or absence) of some natural conditions in human heads, in particular, some relations among our representations, which we develop in the course of our cognitive engagement with our direct natural environment. In what follows, I shall call this proposal *a representationist construal* of the relevant *a priori* truths. Although I shall acknowledge the constitutive role of some epistemic facts or events in the emergence and existence of these truths, nevertheless I shall maintain that our actual knowledge of the obtaining of the suggested conditions is definitely not among these facts. Accordingly, my proposal will be, in a sense specified below, realist about these truths.

An important semantical implication of the above hypothesis, which constitutes the second major tenet of this work, is that truth within our paradigm *a priori* discourses (that is, again, in pure logic and mathematics) cannot be referential in character. In other terms, in contrast to the case of our empirical beliefs about the obtaining of some causally effective spatiotemporal conditions, the truth conditions of our logical and mathematical beliefs, as well as those of our presumably justified beliefs about causally inert entities in general, cannot be adequately specified in terms of the relevant intended subject matters.[1] In what follows, I shall call this semantical tenet *a non-referentialist construal* of the relevant *a priori* truths.

After presenting my arguments for the above hypotheses about the cited paradigms of *a priori* truth, in the last section of chapter 7, I shall suggest that being representational might be more than a contingent characteristic of some *a priori* truths. In fact, by reference to this property we can provide a useful and adequate real definition of *a priori* truth in general. According to this construal, a certain truth is *a priori* if and only if the conditions whose obtaining it actually consists in are representational (i.e. obtaining in the natural realm of

[1] In the semantics of pure logic, this tenet applies if it is supposed that logical beliefs have platonic subject matters (e.g. Fregean thoughts or propositions with inferential relations among each other) or have no subject matter at all.

representations within human heads). In harmony with this definition and the existing conceptual link between our notion of *a priori* truth, on the one hand, and our notions of *a priori* knowledge, justification and evidence, on the other, we can stipulate also that a piece of evidence, or the justification that it provides for a true belief that is based on its recognition, or the piece of knowledge that we acquire by this justification, is *a priori* if and only if it is generated by an appropriate causal mechanism within human heads that conveys information of the obtaining or absence of truth conditions within a subject's system of representation to her knowing mind.[2]

Before starting my reasoning, however, in this chapter, I wish to say a few words about the intended meaning of the central terms of the previous proposals. In section 1, I shall start with some remarks on the notion of *apriority*. The main question that I shall address here is whether our understanding of this notion in terms of independence of experience indeed amounts to a purely epistemological distinction, or rather it relies on further contrasts that are best classified as ontological or semantical in character. Another question that I shall raise in this section is whether there is a reasonable way to start an inquiry into the nature of *a priori* truth, knowledge, justification and evidence without having a sharp and commonly shared notion of exactly which truths, beliefs, justifications or epistemic grounds are *a priori* in character.

In section 2, I shall turn to the second key notion of this work: the intended concept of *truth*. Issues about truth take centre stage in present-day analytic philosophy. What are the entities that this concept primarily and derivatively applies to? Does it represent a substantive property of these entities, or it is merely a useful logical device? If it represents a property, then can we say something about the nature and metaphysical status of that

[2] An important consequence of this construal of apriority is that claims *about* the obtaining or absence of these truth conditions in the actual world will also qualify as *a priori*, even if *their* truth conditions remain to be understood in referential terms.

property, or it is a primitive feature that cannot be further characterised? If there is something more to say about it, then what exactly should we say? Are there epistemic properties that are constitutive of the property of truth? Is anyone's knowledge of some truths constitutive of the existence of those truths? If truth is not an epistemic property, then what can be said about the nature of those conditions whose obtaining or absence determines the truth value of the relevant truthbearers? Can that obtaining or absence be inaccessible to the human mind, and if so, then in what sense of inaccessibility? Can the conditions in question obtain necessarily? Can they obtain outside the spatiotemporal world? Can they obtain in a world that is not actual? And how does truth relate to the semantic content of its bearer and the bearer's more basic representational constituents? Finally and, from the perspective of this work, most importantly, how does truth relate to the existence of the intended referent or subject matter of its bearer and the bearer's constituents? In the following chapters, I shall attempt to answer each of these questions in more or less detail. In this chapter, I shall merely present a brief sketch of these answers, which may provide the reader with a preliminary grasp of what I shall argue for somewhat scattered in the remaining part of this work.

In section 3, I shall focus on the third key notion of this work: the adopted concept of *realism*. First, I shall offer an explicit characterisation of what I mean by realism, a specification that will preserve the concept's primarily metaphysical character. On the other hand, I shall recall two reasons for which the resulting metaphysical doctrine concerning a certain object, property, condition or domain may become a semantically significant conception as well. Afterwards, I shall contrast the specified notion of reality with the idea of objectivity, and emphasise that in this work objectivity will be invoked as a major *explanandum*, while reality as an indispensable *explanans*. Having clarified this, I shall briefly examine the quasi-realist observation that a thoughtful anti-realist about a certain domain can to a large extent mimic the linguistic practice of her realist opponent while specifying her view of the semantic features of her

representations of this domain. In contrast to those who believe that the correctness of this observation undermines the communicability or intelligibility of the metaphysical doctrines denoted by the terms 'realism' and 'anti-realism', I shall argue that not only can we understand the alternative senses in which realists and quasi-realists can apply the same words while explicating the content of their own semantical commitments, but there are even ways in which we can communicate these senses, and thus express our metaphysical views in terms of a public vocabulary in a sufficiently transparent manner. Finally, I shall distinguish two senses in which one can adopt realism concerning a given discourse (and thus become, for instance, an 'ethical realist' or a 'mathematical realist'), and observe that adopting realism in the one sense is a logically independent commitment from adopting realism in the other.

In section 4, I shall clarify the sense in which I intend to talk about *reference* and *referentialism about truth* in this work. First, I shall endorse what virtually everyone assumes about reference: the idea that it is an asymmetric "word-world" relation between meaningful symbols, on the one hand, and what the users of these symbols intend to think of or speak about while applying these representations in particular contexts in the actual world, on the other. Second, I shall recall a number of alternative construals of this relation, with different implications concerning the exact circle of referring entities, and emphasise that the notion adopted in this work is a relatively inclusive and deflated one, which acknowledges the referential power of definite descriptions, adjectives, adverbs and other meaningful expressions as well, and presupposes neither the reality nor the actual existence of these symbols' intended subject matters or referents. Third, I shall highlight the essential role of our referential intentions in the determination of the nature and identity of actual referents, and explain why I shall largely neglect in this work those conceptions that attempt to save the received referentialist construal of truth in our paradigm *a priori* discourses by adopting a revisionist view about the nature of the relevant intended subject matters. Finally, I shall observe that the suggested deflated notion of reference

stands for a very fine-grained asymmetric "word-world" relation, which must not be conflated with its relatively coarse-grained Fregean counterpart. With these observations in mind, I shall conclude that my primary opponents in this work will be those philosophers who believe that an adequate theory of truth can specify the truth conditions of every truth-apt mental or linguistic symbol in terms of those conditions or entities that the constituents of this symbol refer to in the specified deflated, essentially intentional and fine-grained sense of the term.

In the short characterisation I gave of the construal of *a priori* truth, knowledge, justification and evidence that I should defend in this work, there were three further terms whose meaning may call for preliminary clarification. In section 5, I shall focus on two of them, and specify the sense in which I intend to speak about *natural* conditions, objects, properties and domains, on the one hand, and *abstract* conditions, objects, properties and domains, on the other.

Finally, in section 6, I shall turn to the last crucial constituent of that characterisation, and elucidate the sense in which I shall speak about *representations* in the following chapters. The primary aim of this section is to provide a clear picture of what the suggested representationist construal of *a priori* truth and knowledge presupposes concerning the existence, the nature, and the semantic content of representations. In particular, there will be six major assumptions emphasised and discussed in this part. First, I shall grant that the construal presupposes the existence of both mental and physical representations, vindicate the first half of this commitment against the eliminativist challenges advanced in the philosophy of mind, and finally note that the core message defended in this work could be formulated in a way that is compatible with an eliminativist stance towards mental representations as well. Second, I shall notice that the construal also assumes the compositionality of semantic content, which is often believed to run counter to a connectionist picture of the computational architecture of the mind. In contrast to the latter conviction, I shall adopt the opposite view and maintain that compositional semantic contents can emerge in a connectionist

computational system as well. Third, I shall register that the construal takes the (truth-related part of the) semantic content of our mental and physical symbols to be explicable in terms of the correct declarative use conditions of these representations. Nonetheless, I shall hasten to add that this (broadly) use-theoretic conception of semantic content is meant to be the denial of neither a truth conditional account of meaning nor a realist construal of truth and correct use in general. In fact, as I shall emphasise, the construal assumes both that a truth-apt representation's correct declarative use conditions (in a non-embedded state) are identical with the truth conditions of this representation, and that these conditions are to be understood along the realist lines. Fourth, I shall point out (here again) that the construal abandons the standard referentialist assumption that the declarative use conditions of our mental and physical symbols can be understood in each representational context in terms of those objects and properties that these symbols purport to be about. Fifth, I shall remark that the construal also assumes the existence of narrow semantic contents, and then explain how these contents can be distinguished from each other and the syntactic objects that they are associated with. Finally, I shall observe that the construal distinguishes between the possession and the actual entertainment of representations, and assumes that the truth conditions of *a priori* beliefs obtain (if they do) within the domain of representations possessed in the thinkers' heads. As I shall argue, this assumption leaves room for a realist construal of *a priori* truth. On the other hand, it leaves open whether the domain of representations invoked in this construal can be rightly classified as mental in character.

1. Apriority

The concept of apriority has had a long life in the history of philosophy. In the 17th century, the distinction between *a priori* and *a posteriori* knowledge became a central contrast within epistemology. Rationalist thinkers such as Descartes, Spinoza and

Leibniz argued that real knowledge must be based on reason rather than on experience. Their empiricist opponents heavily disputed this claim. According to them, if there are some truths that can be known independently of experience, they can at most concern the relations of our ideas, rather than the facts of a mind-independent world.

Present-day discussions of apriority are fundamentally shaped by Kant's famous exposition of the *a priori* and *a posteriori* distinction in his introduction to the *Critique of Pure Reason*. Following his rationalist and empiricist ancestors, Kant construed apriority in epistemological terms. According to his definition, a judgement is *a priori* if and only if its justification does not require reliance on sensory experience.[3] The most important novelty of his exposition was the specification of the relation of this contrast to two other (non-epistemological) distinctions, which can (at least derivatively) also characterise human judgements. The first of these is the metaphysical contrast between necessary and contingent truths. Although, according to Kant, our concept of necessity is not analytically related to our concept of a priority, the two concepts are nevertheless extensionally equivalent. *A priori* judgements are necessarily true or necessarily false, while *a posteriori* judgements express contingent truths or falsities. Consequently, necessity can be regarded as a sign of apriority.[4] The second distinction examined by Kant is the semantical contrast between analytic and synthetic judgements. In his construal, analytic judgements are those whose predicate concept is wholly contained within their subject concept, while synthetic judgements are those whose predicate concept amplifies the subject concept. According to Kant, the latter distinction is not coextensive with the one between *a priori* and *a posteriori* claims. While the possibility of *a posteriori* analytic judgements is excluded by the fact that the detection of containment relations among our concepts never relies on sensory experience, some synthetic

[3] Kant (1781/1787), 136.

[4] As is known, beyond necessity, Kant took strict universality as a sign of apriority too. Kant (1781/1787), 137-138.

judgements are clearly necessary, and therefore *a priori* in character. Kant's famous programme in his first *Critique* was to show how these *a priori* synthetic judgements are possible at all.[5]

The programme set the stage for subsequent discussions of *a priori*, and it keeps shaping our ideas of the subject even today.[6] One major issue in contemporary literature is the proper delineation of the concept of apriority. The dominant view today is that apriority is an epistemological concept which is meant to stand for a property primarily attributed to types of evidence or ways of justification. Derivatively, the property in question can be attributed to pieces of knowledge, beliefs, propositions, judgements, sentences, utterances and truths due to the relevant epistemological features of these entities.[7] Sometimes arguments and concepts are also said to be *a priori*. Arguments are regarded as *a priori* if and only if all of their premises are *a priori*, while an *a priori* concept is meant to be one that can be acquired independently of experience. It is commonly understood that an epistemological notion cannot be adequately analysed in purely non-epistemic terms. Most construals of apriority, therefore, include an epistemic concept in the proposed *analysans*. Of course, this policy will not guarantee that the notion analysed becomes clearly disentangled from all non-epistemological connotations. Moderate empiricist construals of *a priori* knowledge, for instance, developed in the era of post-Fregean analytic philosophy usually assumed that the features of analyticity and necessity were essentially linked with the epistemological feature of apriority.[8] Despite the subsequent efforts to keep the epistemological distinction from other (semantical, metaphysical or ontological) distinctions apart, the

[5] Kant (1781/1787), 141-143, 146-148.

[6] Recent works on *a priori* include Hanson and Hunter (1992), BonJour (1998), Bealer (1999), Boghossian and Peacocke (2000), Casullo (2003) and Horowitz (2006).

[7] Some philosophers do not share this view. For instance, as Casullo points out, Quinton (1972) construes apriority as a concept to be analysed in non-epistemic terms. Casullo (2003), 12-13.

[8] Ayer (1946), 16-18.

inclusion of non-epistemic notions in the analysis of the concept of apriority is still prevalent in present-day epistemology.[9]

In this work, I shall adopt the original Kantian (and currently dominant) idea that the distinction between *a priori* and *a posteriori* beliefs concerns the way in which judgements, beliefs, utterances, premises and conclusions, are justified, or in which concepts are acquired in the course of a subject's cognitive development. In particular, I will suppose that a concept or the truth value of a certain truthbearer is knowable *a priori* if and only if it can be, respectively, acquired or justified independently of experience, and that this equivalence holds because the concept of apriority is to be analysed in terms of the notion of *independence of experience*.[10] My primary aim in this section is to examine whether this *prima facie* epistemological construal of apriority is indeed independent of notions that are usually classified as ontological or semantical in character.

In order to see how, following this construal, one could still consistently hold that the difference between those beliefs that are supposed to be *a priori* and those that are not might be partly ontological or semantical, rather than purely epistemological,

[9] For instance, BonJour argues that "a proposition is justified *a priori* when and only when the believer is able, either directly or via some series of individually evident steps, to intuitively 'see' or apprehend that its truth is an invariant feature of all possible worlds, that there is no possible world in which it is false". BonJour (1985), 192. Further analyses of the notion of apriority that include a non-epistemic concept in the *analysans* can be found in Chisholm (1989), 28, and Plantinga (1993), 105. For a critical discussion of these proposals, see Casullo (2003). Suggested examples of *a posteriori* necessary truths include the propositions *that water is H_2O* and *that Hesperus is identical with Phosphorus* while a proposed case of *a priori* contingent truths is the proposition *that the standard meter bar in Paris is one meter long*. Kripke (1972/1980), Putnam (1975c), and Kitcher (1980).

[10] The acquisition of a certain concept or the justification of a certain belief, in turn, is taken to be independent of experience if and only if the evidence on which the development of the concept or the establishment of the belief relies is not experiential in character. Beyond this negative characterisation, purely or partly epistemological analyses of apriority often include some positive claims about how *a priori* propositions *are* knowable or justifiable. According to BonJour, for instance, apriority means *being based merely on the deliverances of reason or rational insight*. BonJour (1998), 11.

consider what fixes the extension of the intended concept of being independent of experience. Supposing that the semantic content of a complex piece of representation is determined by that of its components, we can clarify what independence of experience means by elucidating the content of the notions of independence and experience.

As regards the first of these concepts, there are at least two important senses in which *a priori* justifications, as opposed to *a priori* concept acquisitions, are not meant to be independent of experience. First, a judgement cannot be justified, unless its constituents have been understood by the epistemic agent. As Kant already observed, however, *a priori* judgements can involve empirical concepts, which could not be acquired and understood independently of experience. If this is true, then apriority, in so far as it is applied to justifications, cannot presuppose our understanding of the relevant truthbearers independently of experience.[11] Second, there are philosophers who maintain that *a priori* beliefs can be undermined by experience. If this is right, then apriority cannot consist in immunity to experiential disconfirmation either. In order to ensure the greatest possible consensus over the intended notion of independence in the analysis of *a priori* justification, the best way to proceed is to confine its significance to the sources of positive justification. According to this construal, the justification of a certain belief is independent of experience if and only if it can be confirmed without reliance on experience.

Received interpretations of experience differ from each other in how inclusive they are concerning the notion's extension. According to the narrowest construal, the concept merely denotes perceptual experience of the external world. The

[11] Apriority, applied to concepts, rather than justifications, consists in independence of experience in this stricter sense. Propositions, knowledge claims, beliefs, judgements, sentences, utterances and truths can be *a priori*, derivatively, in two different senses: they can be *a priori* in a looser sense, due to their justifiability independently of experience, and in a stricter sense, due to their understandability independently of experience. Kant calls this stricter sort of apriority *pure* apriority. Kant (1781/1787), 137.

intermediate understanding includes perceptual experience of the subject's own bodily states as well. Finally, on the broadest interpretation, the notion applies to any conscious state or event including, beyond the deliverances of the external and internal senses, any object of reflexive attention.[12] Which, from among these alternative senses, is the one that might be constitutive of the intended notion of *a priori* knowledge and justification? Although the paradigm cases of empirical claims are taken to rely on experience in the narrowest sense of the term, the extension of *a posteriori* knowledge was never confined to justified true beliefs about the external world. Following Kant's classification, judgements relying on introspective evidence (i.e. the deliverances of our inner senses, have also been regarded as empirical).[13]

One may wonder, however, what motivates this terminological convention. In other words, why do philosophers not tend to define apriority in terms of (independence of) experience merely of the external world? The most plausible answer to this question is arguably that the resulting notion of apriority, which would apply to introspective evidence and justification as well, would serve the theoretical purposes of epistemology less than a notion that renders introspection an *a posteriori* source of evidence. Such theoretical purposes can be to have a distinction that coincides with the one between infallible and fallible beliefs, or between epistemologically more and epistemologically less fundamental classes of beliefs, or maybe between more and less reliable types of belief-forming mechanisms.

Beyond these purely epistemological considerations, one may have ontological or semantical motivations as well for refusing the above inclusive construal of apriority. First, one may observe that introspection resembles perception of the external world more than the epistemic faculty used in pure logic and mathematics, in so far as its deliverances support beliefs about contingent and natural, rather than about necessary and abstract,

[12] Boghossian and Peacocke (2000), 2.
[13] Kant (1781/1787), 155-171.

states of affairs.[14] Second, one may find introspection more similar to perception than to logical and mathematical intuition also because the latter faculty, in contrast to the former two, apparently transcends the confines of the natural world.[15]

Now, clearly, this second sort of motivation is heavily loaded with substantive theories of the truth conditions of the contrasted empirical and *a priori* beliefs. A naturalistic construal of *a priori* truth, for instance, would certainly undermine the above difference between perception of the external world and introspection, on the one hand, and logical and mathematical intuition, on the other. The conclusion that I shall draw from these observations is that the adopted negative construal of apriority in itself does not guarantee the purely epistemic character of the resulting concept. The crucial question to be answered remains whether or not our notion of experience can be construed in a purely epistemic way. In other terms, what we should clarify is whether there is a purely epistemic contrast between what we commonly regard as (a type of) experience and what we usually classify as (a sort of) *a priori* evidence. If the answer to this question is negative, then our concept of apriority will hardly qualify as purely epistemic either.

There are similar considerations to the effect that the intended notion of experience appearing in the analysis of *a priori* justification cannot be the broadest one either. Defining apriority in terms of being justified independently of any conscious state or event would render the paradigm cases of *a priori* justification, such as justifications of our mathematical and logical claims,

[14] In section 5 below, I shall specify the sense in which I am to apply the terms 'natural' and 'abstract' throughout this work.
[15] According to BonJour, "the relevant notion of experience should be understood to include any sort of process that is perceptual in the broad sense of (a) being a causally conditioned response to particular, contingent features of the world and (b) yielding doxastic states that have as their content putative information concerning such particular, contingent features of the actual world as contrasted with other possible worlds. […] And thus not only sense experience, but also introspection, memory, kinaesthesia, and clairvoyance or telepathy (should these exist) would count as varieties of experience". BonJour (1998), 8.

trivially empirical. The question, again, is why we should resist the radical empiricist proposal that even these paradigm cases of *a priori* justification are, in fact, empirical. One sort of motive may be, again, intrinsically epistemological: we may hold that in some genuinely epistemological respect the justification of, say, our mathematical or logical beliefs is different from the justification of our perceptual and introspective beliefs. Other motives may be ontological or semantical: we may observe that, due to some facts concerning the relevant truth conditions, the nature of the former class of justification is categorically different from that of the latter. Be that as it may, as before, we may conclude that in absence of a purely epistemic difference between what we commonly regard as (some type of) experience and what we usually classify as (a sort of) *a priori* evidence our concept of apriority can hardly be regarded as purely epistemic in character.[16]

The interesting debate between radical empiricists, who deny the existence of *a priori* knowledge, and apriorists, who maintain that some beliefs can be justified independently of experience, is, clearly, not terminological. Most apriorists would probably concede that there might be some similarities between what *they* call *a priori* and empirical justifications. One of these features might be that both kinds involve conscious events within the subject's mind. No apriorist would ever want to claim that this feature could not be taken as constitutive of aposteriority. On such construal, they would probably admit that all knowledge is empirical. Still, they would insist that if we compared our beliefs about various domains, then we could find a theoretically important difference between those that they want to call *a priori*, on the one hand, and those that they propose to regard as empirical, on the other. Many of them might even insist that the difference in question is purely epistemological in character.

[16] In so far as we maintain that the distinction between *a priori* and *a posteriori* beliefs concerns the way in which judgements, beliefs, utterances, premises and conclusions are justified, or in which concepts are acquired in the course of a subject's cognitive development, our notion of apriority will remain epistemic in a trivial sense of the term.

Similarly, most radical empiricist would probably concede that there might be some contrasts between what *their opponents* call *a priori* and empirical justifications. One of these contrasts might be that the former support beliefs about non-natural states of affairs, while the latter provide warrant for claims about the natural world. No radical empiricist would ever want to claim that this contrast could not be taken as constitutive of apriority. On such construal, they would probably admit that some knowledge can be regarded as *a priori*. Nevertheless, they would insist that, from an epistemological point of view, there is no important difference between the contrasted kinds.

In the following chapters, I shall argue that there is a perfectly good sense in which *a priori* knowledge, justification and evidence can be distinguished from their empirical counterparts. Nevertheless, the distinction that I shall propose will invoke a contrast between the nature of the truth conditions of our *a priori* and empirical beliefs. By reference to this distinction, I shall argue that there must be a corresponding contrast between the nature of those epistemic mechanisms by which we can detect the obtaining or absence of these conditions in the actual world, a contrast whose specification is the task of our empirical sciences.

An inquiry into the nature of *a priori* truth, knowledge, justification and evidence, of course, presupposes that one can (at least vaguely) identify the intended subject matter of this investigation. In view of the above disagreements concerning the extension of our concept of apriority, the best strategy that we can adopt here is to examine the least contentious "paradigm" instances of the contrasted types, develop an adequate proposal of the defining characteristics of the relevant classes on this commonly accepted evidential ground, and then consider whether the more controversial instances, initially set aside, can be classified as *a priori* or as empirical by means of the newly articulated conceptual resources.

In the following chapters, I shall adopt the foregoing strategy too. First, I shall suppose that our perceptual claims about our direct spatiotemporal environment are the least contested examples of empirical representations, while our claims

in pure logic and mathematics are the clearest instances of *a priori* truthbearers. Corresponding to this classification, I shall also assume that perceivable truths, pieces of perceptual knowledge, perceptual justifications and pieces of perceptual evidence are the paradigm cases of empirical truth, knowledge, justification and evidence, while the respective counterparts of these entities in pure logic and mathematics are what we usually regard as the clearest examples of *a priori* truth, knowledge, justification and evidence.

In view of the observable specificities of these paradigms, then, I shall argue that from among the numerous contrasts that obtain between the relevant types of entities there is one that might be invoked in an adequate real definition of apriority as well. Again, the contrast in question will concern the nature of those conditions whose obtaining or absence determines the truth value of the relevant representations, and therewith the nature of those epistemic mechanisms by which we can detect the previous facts in the actual world.

In possession of the proposed analysis of apriority, at the end of chapter 7, I shall finally consider what the suggested construal implies concerning the classification of representations and epistemic grounds beyond the examined paradigms. As we shall see, the account will ratify the *a priori* status of our analytic claims in general, and it will delineate the possible content of synthetic *a priori* knowledge of the actual world as well.

2. Truth

Developing a proper understanding of truth is one of the highest priorities of present-day analytic philosophy. In this section, I shall provide a brief sketch of the account that I am to develop of this subject in greater detail in the remaining part of this work.

First, I intend to say a few words about the primary bearers of this property. In what follows, I shall take truth to be primarily a property of truth-apt mental and physical representations, such as thoughts and their various linguistic expressions, rather than a

property of objectified semantic contents, such as abstract propositions.[17] By this assumption, I do not query that our syntactically identifiable mental and physical symbols are true (or false) partly in virtue (of a certain segment) of their semantic contents (namely, their truth conditions). What I deny is that, because of this role, semantic contents also qualify as bearers of truth and falsity. Having said this, I shall further recognise that, beyond the above fundamental applications, the semantic values under scrutiny can be, derivatively, predicated of some mental attitudes and physical acts (e.g. beliefs, judgements, claims and utterances) involving the relevant primary bearers as well. A common feature of these secondary bearers is that they can be equally regarded as (purportedly correct) declarative applications of the primary bearers involved. This fact fits well with the fifth major tenet mentioned below, according to which truth is identical with the correct unembedded declarative applicability of our truth-apt representations in the context of their actual use.

Second, contrary to deflationist theories of truth, I will maintain that our actual notion of truth is not merely a logical device, but it stands for a substantive property whose nature and metaphysical status can be further characterised. My reason for adopting this view is explanatory in character, and it will be presented in detail in chapter 4, where I examine, among others, the explanatory adequacy of the most elaborated current form of deflationism, Paul Horwich's minimal theory of truth.

Third, beyond committing myself to a substantive construal of the notion of truth, I shall further assume that the property represented by this concept characterises its bearers independently of what anyone actually thinks of this particular circumstance. Note that my claim is not that the truth (or falsity) of a certain bearer can never be the result of our epistemic activities. In fact, as I shall argue in chapter 7, the truth of our *a priori* beliefs is always a product of our epistemic activities. In particular, it presupposes that we have developed certain suitably

[17] In section 6, I shall say more about the ontological and semantical assumptions underlying my claims about alternative sorts of representations.

connected representations in our heads. The current point is rather that these activities do not include anyone's actual recognition or awareness of the relevant truths (or falsities). In a specific sense, which I shall clarify below, this tenet amounts to the endorsement of a general realist construal of truth, which reinforces the view that truth is a correspondence between what is believed, on the one hand, and what actually obtains in the world, on the other.[18] The explanatory consideration I present against deflationism in chapter 4 will reveal my main reason for adopting this general realist construal of truth as well. In chapter 5, I shall examine the most influential "anti-realist" objections to this view, and show that none of them provides us with good reason for revising the previous commitment.

Fourth, I shall suppose that truth, at least when it is knowable for a natural subject, is a natural property that characterises its bearers in virtue of two separate facts: first, the fact that these bearers stand in a certain semantic relation with some conditions that may or may not obtain in the spatiotemporal world; and second, the fact that the conditions in question actually obtain in the spatiotemporal world. In the case of our beliefs about abstract states of affairs, this tenet comes to grips with the standard referentialist construals of truth, according to which the truth or falsity of a certain belief is always to be understood in terms of (the obtaining or absence of) those conditions that the belief in question purports to be about. My reason for adopting this naturalist tenet is also explanatory in character. In particular, I believe that its opponents cannot explain all observable characteristics of (what we take to be) the paradigms of *a priori* knowable truth. Deflationist and anti-realist versions of anti-naturalism will be rejected in chapters 4 and 5 as

[18] It must be noted, however, that the standard referentialist form of the correspondence theory will be rejected here together with the standard referentialist construal of truth. Instead, what will be endorsed is that truth is correspondence between what is said or believed and the obtaining truth conditions of what is said or believed, regardless of whether or not the latter conditions are identical with what the claims or beliefs in question purport to be about.

variants of deflationism or anti-realism about truth. In chapter 6, in contrast, I shall argue against those who defend a realist (i.e. platonist) form of anti-naturalism in the semantics of our paradigm *a priori* discourses. Finally, in chapter 7, I shall show that the proposed naturalistic constraint is compatible with an adequate construal of truth in the semantics of our paradigm *a priori* discourses as well.

Fifth, I shall understand truth as a property that is identical with that of the correct unembedded declarative applicability of its primary bearers in the context of their actual applications.[19] Accordingly, I will assume that the truth conditions of a given representation are identical with those conditions whose obtaining or absence is supposed to govern the correct declarative use of this bearer in the context of its actual applications. The correct declarative use conditions of truth-apt representations will be supposed to be determined by the correct declarative use conditions of their basic constituents and the mode of their composition. The correct declarative use conditions of concepts and their linguistic expressions (i.e. the basic constituents of truth-apt representations) will in turn be identified with those conditions whose obtaining or absence is supposed to govern the correct applicability of these constituents in the context of alternative truth-apt representations. My reason for adopting this strictly realist form of a broadly Wittgensteinian use-theoretic semantics is also explanatory in character, and it will be articulated in chapter 4, in the course of my discussion of Horwich's deflated version of this use-theoretic account.

Finally and (from the perspective of this work) most importantly, in what follows, I shall abandon the standard referentialist construal of truth, the idea that the truth conditions of our truth-apt representations are always to be understood in terms of those conditions that these bearers purport to be about.

[19] By invoking contexts of actual applications, I wish to ensure the explanatory resources for a suitable account of indexicality, lack of truth value and other semantic phenomena emerging from the context dependence of the correct declarative applicability of certain truthbearers.

Contrary to this referentialist understanding, and thus also to Tarski's highly influential semantic theory of truth, I shall maintain that referentiality is not encoded in our general notion of truth.[20] Moreover, if my explanatory considerations in support of the naturalistic construal of knowable truths are correct, then in the case of our paradigm *a priori* discourses about abstract domains standard referentialism cannot even be true. It is not that I query our capacity of developing truth-apt representations of abstract domains whose truth conditions are referential in character. What I shall argue for is merely that our purportedly *a priori* claims in pure logic and mathematics are not such representations. They are not, because they are all analytic, and the truth value of our analytic claims is determined by the obtaining or absence of some conditions in the domain of representations within our heads, rather than by the obtaining or absence of those states of affairs that these claims purport to be about. The truth values of those claims whose truth conditions are rightly supposed to obtain (or fail to obtain) in an abstract domain, on the other hand, will be regarded as necessarily inscrutable to human minds.

Of course, the above construal of truth (and *a priori* truth in particular) can be accepted only if it satisfies all those adequacy conditions that we can reasonably set for a proper theory of these subjects. In chapter 2, I shall say more about these conditions, and in the rest of this work, I shall show that the construal specified above is indeed capable of satisfying them all.

3. Realism

For a long time in the history of philosophy, realism has been understood as a doctrine of the metaphysical status of various entities. Although the term did not lose its original metaphysical connotations, in the late 20th century it became common to

[20] Tarski (1944).

construe realism in semantical terms. Michael Dummett, for instance, argued that the *prima facie* metaphysical disagreement between realists and anti-realists concerning the status of certain problematic domains can be best understood as a debate about the applicability of classical logic within the discourses about these domains. Hilary Putnam also described the doctrine of metaphysical realism as the adoption, in the semantics of the relevant discourses, of a correspondence theory of truth.[21] The adequacy of these and other semantical construals has been heavily debated in contemporary literature. Although a proper discussion of the arguments for and against these understandings is beyond the reach of the current section, a brief characterisation of how I shall use the term in the following chapters might be appropriate.[22]

In what follows, I shall adopt the traditional view and regard realism and anti-realism as doctrines about the metaphysical status of various thinkable entities or domains. In particular, I shall suppose that realists about certain subjects maintain that the subjects in question obtain or exist (if they actually do) independently of whether or not anyone ever does or can think or know of this circumstance. Anti-realists, in contrast, will be supposed to query either the possibility of the obtaining or existence of the relevant entities or the independence of this obtaining or existence of anyone's actual or potential thought or knowledge of it.[23]

[21] Dummett (1963) and Putnam (1981).

[22] An influential defence of the idea that realism is a purely metaphysical doctrine, the correctness of which is entirely independent of our conceptions of truth and valid inference, can be found in Devitt (1984). In chapter 5, I shall examine Dummett's and Putnam's conceptions in detail, and argue that neither of these authors endorses anti-realism about truth in the sense specified below.

[23] The construal does not render realism about conscious mental states inconsistent. There is nothing inconsistent in the claim that our conscious mental states can occur in the actual world independently of anyone's actual thought of this circumstance. Fictive objects and objects "existing" only *within* a conscious mind, on the other hand, cannot be consistently regarded as real. Depending on how one applies the term 'existence', one must either deny the possibility of these objects' existence or maintain that their existence relies on someone's actual or

The second thing that I shall assume here concerning realism is that the adoption of this doctrine concerning a certain entity implies no commitment to the actual obtaining or existence of this entity. What the realist assumes is merely that the entity can obtain or exist in the actual world, independently of anyone's actual or potential thought or knowledge of this circumstance. Thus, for instance, one might adopt a realist (i.e. platonist) construal of some abstract entities without committing herself to the actual existence of those entities, just as much as someone might adopt a realist construal of some spatiotemporal entities without committing herself to the actual existence of those entities. It is one thing that my idea of Napoleon is an idea of a real object, and another that I believe in the (past) existence of this object in the actual world. If my realist construal of Napoleon implied my belief in his actual existence, then I could not consistently entertain the thought that, despite appearances, Napoleon did not exist at all.

The third point that I wish emphasise here is that in this work the concept of reality will be distinguished from the concept of objectivity. Although the two concepts are often interchangeably applied, in what follows, I shall suppose that the question whether or not a certain entity is real concerns the metaphysical status of this entity, while objectivity is a feature of the correct declarative applicability of mental and physical symbols.[24] In particular, I shall say that the declarative applicability of a certain symbol is objective if and only if the symbol's use is subject to a certain sort of evaluation (i.e. it can be correct or incorrect), and no opinion of this issue (i.e. no one's

potential thought or awareness of this circumstance. Eliminativism about a certain domain exemplifies anti-realism stated in terms of the first vocabulary, while phenomenalism or idealism concerning a class of entities exemplifies anti-realism understood in terms of the second.

[24] In a derivative sense, the feature can be predicated of the relevant symbols or their particular applications as well. Supposing the objectivity of mathematical truths and falsities, one can also say that the mathematical representations expressing those truths and falsities, or their actual applications by us (i.e. our claims and beliefs), are objective.

actual or potential judgement of this correctness) is conceptually prevented from being mistaken. A certain claim made by a certain subject in a certain context, for instance, will be understood as a particular declarative application of a certain truth-apt representation in that context, which is subject to the relevant kind of evaluation (it can be correct, i.e. true, or incorrect, i.e. false). To say that the truth value of this claim is objective is to maintain that no one's actual or potential opinion of the correctness of this (or any other) application of the relevant representation is conceptually prevented from being mistaken.[25]

Objectivity is a relatively transparent feature of the applicability of our mental and physical symbols. In most cases, we can easily decide whether anyone's actual or potential use of a certain symbol is constitutive of how the symbol in question has to be applied (in which case the relevant subject's cognitive or linguistic practice would be conceptually prevented from being mistaken). In contrast, our view about the metaphysical status of various entities or domains depends on highly theoretical explanatory considerations. For instance, hardly any philosopher of mathematics would query today the objectivity of mathematical truths. In contrast, the same thinkers are rather divided concerning the metaphysical status of mathematical objects and properties. In what follows, I shall observe that objectivity is a general (i.e. discourse-independent) feature of the applicability of our truth-apt representations (and thus a general feature of truth and falsity), I shall regard this feature as a major *explanandum* for a theory of any sort of truth, and finally, in chapter 4, I shall argue that a proper account of this *explanandum* requires a substantive realist construal of the truth conditions of our truth-apt representations. Putting it briefly, in this work, I

[25] The fact that the use of a certain meaningful symbol is never conceptually prevented from being mistaken, of course, does not exclude the possibility of perfect applications, or in other terms, the occurrence of an ideal thinker or speaker, who never fails to know whether or not the declarative use conditions of the symbols applied by her obtain in the actual world.

shall invoke objectivity as a major *explanandum*, while reality as an indispensable *explanans*.[26]

Despite the previous endorsement of the traditional metaphysical construal of the contrast between realism and anti-realism, I shall add that there can be at least two reasons for which realism about a certain domain can also be regarded as a genuine semantical position. First, since the metaphysical distinction at hand concerns thinkable conditions, objects, properties and domains that can be identified as subject matters of particular thoughts and beliefs, it is always possible to regard realism about these entities indirectly as a conception of the subject matters or purported referents of those thoughts or beliefs. Second, in so far as the contrasted metaphysical doctrines concern a conceivable domain whose states of affairs can be the truth conditions of certain thoughts and beliefs, the doctrines can, again, be equally regarded as components of a conception of truth in the semantics of discourses involving those thoughts and beliefs. Note, however, that even if these links may explain the emergence of alternative semantical construals of realism, the possibility of these readings does not imply that the notion would lose its essentially metaphysical character. The fact that the debate between realists and anti-realists concerns the metaphysical status of some semantically relevant conditions does not lessen the metaphysical nature of this controversy.[27]

One hotly debated issue in contemporary literature on realism is whether the metaphysically thick notions of existence, reference or aboutness, correspondence, facts, objects, properties, and domains of quantification, in terms of which the contrast between realist and anti-realist positions has often been

[26] In chapter 2, I shall say more about the methodological considerations underlying this argumentative strategy.

[27] The same observation appears in Devitt (1991), 46. Despite their semantical characterisation, a number of influential examples of anti-realism (e.g. error theory, emotivism, expressivism, instrumentalism) can be understood as metaphysical doctrines concerning truth conditions and intended referential domains. Mackie (1977), Field (1980), Ayer (1946), Blackburn (1993), Gibbard (1990).

characterised, can be intelligibly distinguished from their metaphysically thin counterparts whose application in alternative discourses implies no metaphysical commitments on our parts concerning the relevant subject matters.[28] If the distinction is unintelligible, then the characterisations of the contrast between realism and anti-realism made in terms of these notions cannot be sufficiently illuminative either. This would be a rather sobering conclusion, since in absence of an unambiguous metaphysical vocabulary, it might become futile to argue for one or another metaphysical position.

An anti-realist about arithmetical objects, for instance, who denies the reality of the intended abstract domain of pure arithmetic, is supposed to maintain that there can be no numbers and no arithmetical facts outside the human mind that our numerals and arithmetic sentences could refer to. Most of them, however, do not want to deny that there are three prime numbers between 70 and 80. Clearly, if the term 'there are' is applied in the same sense in the metaphysical claim as in the arithmetical one, then the two commitments cannot be reconciled with each other. In other words, an anti-realist could not consistently believe in the truth of this apparently correct existential statement of pure arithmetic.

Some anti-realists have famously accepted this conclusion, and argued that our received mathematical theories, whose truth would require the existence of platonic objects and properties, or the obtaining of some platonic conditions, are actually false and the reason for which we still accept them is not because they are

[28] For a useful discussion of this problem, see Blackburn (1993) and Fine (2001). As may be apparent from my earlier claims about the concept of truth, contrary to Fine, Blackburn, Horwich and many others who observed the possibility of a "quasi-realist" or a deflated reading of the above metaphysical and semantical terms, I believe that the application of the predicate 'is true' does not allow such a construal. While there is a sense in which we can think about, believe in, refer to, and quantify over some facts, objects, properties or domains that are merely invented by human minds, the conditions whose obtaining or absence is supposed to determine the declarative applicability of our truth-apt mental and physical symbols always obtain (if they do) in the real world.

true, but because they have some other utility in our theorising about the world.[29] Most anti-realists about abstract mathematical objects and properties, however, do not accept this conclusion. They argue that their metaphysical view is fully compatible with the truth of their mathematical beliefs. They maintain that they have, in fact, two different concepts in mind that happen to be expressed equally by the term 'exists'. From the fact that the metaphysically neutral concept can be truly applied to a thinkable object it does not follow, they observe, that the metaphysically thick concept also applies to it on pain of contradiction. One may wonder, however, if the metaphysically heavy notion of existence can indeed be intelligibly stated (or denied) about particular thinkable objects and properties, and, once the answer to this question is positive, also whether the conditions under which the notion can be correctly applied to those subject matters can be informatively communicated or characterised.

These questions can be equally raised concerning all other concepts that used to play a central role in the specification of the metaphysical notion of realism. In particular, one may wonder whether the metaphysically heavy notions of aboutness/reference, correspondence, facts, objects, properties, and domains of quantification can indeed be intelligibly applied in various representational contexts, and whether the conditions governing these applications can be informatively characterised.

Since the re-emergence of metaphysical concerns in late-20th-century analytic philosophy, it had been standardly taken for granted that the proper response to the above questions is affirmative, and the alternative characterisations of realism in terms of the above notions can (more or less adequately) specify the content of the intended metaphysical positions. More recently, however, this opinion has been challenged again by

[29] Field (1980), for instance, argues that the utility of our accepted mathematical theories resides in their conservativeness (the fact that their addition to nominalistically interpretable theories does not extend the circle of nominalistically interpretable consequences of these theories), and the fact that they make the derivation of those consequences easier than it would be otherwise.

Simon Blackburn's quasi-realist considerations in the semantics of various discourses.[30]

According to Blackburn's original idea, a quasi-realist is a type of sophisticated anti-realist, who observes that her metaphysical commitments do not prevent her from an honest imitation of the linguistic practice of her realist opponent. In meta-ethics, for instance, she can honestly maintain that our ethical notions and beliefs refer to genuine ethical properties and facts, or that these beliefs are true if and only if they correspond to these facts. She can endorse this realist verdict, because the metaphysically significant terms of this formulation admit of a deflated, metaphysically non-committal reading as well, which is fully compatible with an anti-realist construal of the subject matter of ethical beliefs. If this quasi-realist observation is true, then it is hard to see how the contrast between realists and anti-realists could be properly characterised. The minimal conclusion that appears to follow from this result is that we must be quiet about these metaphysical issues, and give up the project of attributing metaphysical beliefs on the basis of others' linguistic utterances.[31] However, one may go on and argue that, since intelligible conceptual distinctions cannot fail to receive proper linguistic expressions, the ultimate moral that we should draw from the previous considerations is that the metaphysical contrast between realism and anti-realism is ultimately unintelligible.

The significance of this quietist challenge to the argumentation that I shall advance in the chapters that follow can hardly be overestimated. If the contrast between realism and anti-realism concerning a certain domain is indeed unintelligible or inexpressible, then the most important claims of this work cannot be properly understood either.

[30] Blackburn (1993).

[31] Similar conclusions have been drawn by Fine (1984), Putnam (1987) and Dworkin (1996). In his review of Blackburn's book on quasi-realism, on the other hand, Rosen (1998) argues that the quasi-realist observations need not contradict realism in the semantics of the discourses under consideration.

Accordingly, in defence of the metaphysical vocabulary that I shall rely on in the remaining chapters two claims have to be established: first, that we can distinguish the metaphysically thick notions of existence, correspondence, reference, objects, properties, and domains of quantification from their deflated counterparts; and second, that we can also communicate these distinctions, and express our metaphysical beliefs in a sufficiently transparent way. As regards the first of these points, one may observe that a proper understanding of Blackburn's argumentation presupposes the distinguishability of metaphysically thick concepts from their deflated counterparts. If we had no idea of what the realist (as opposed to a quasi-realist) intends to say by maintaining that a certain domain exists independently of our thoughts or knowledge of it, or that the true sentences of a given discourse refer or correspond to some facts in the actual world, then we could hardly understand the challenge that these formulations can be equally interpreted in an alternative, quasi-realist way, which makes them entirely acceptable to anti-realists as well.[32]

As regards the second point, at first sight, the observed anti-realist reinterpretability of the received characterisations of realism may seem to undermine the communicability of the relevant metaphysical positions. If all sentences that a realist would utter in order to characterise her own commitments could be truly endorsed (under a different interpretation) by anti-realists as well, then it would seem rather plausible to conclude that the linguistic expression of these, otherwise perhaps intelligible, conceptions can never be transparent enough. This conclusion, however, does not follow from the correctness of Blackburn's observations. The fact that there is a quasi-realist reading of the

[32] Less sophisticated evidence for our ability to distinguish between the thick and thin notions of existence, facts etc. is that we recognise the difference between the metaphysical status of Napoleon's and Little Red Riding Hood's grandmothers, even if we are ready to endorse that both of these individuals had a grandmother. Apparently, the simple fact that we truly quantify over something does not mean that we believe in its thick existence.

received realist manifestos does not mean that there can be no linguistic manifestation of the difference between a genuine and a quasi-realist construal of certain domains of facts, objects and properties, and their semantic relations to our thoughts and linguistic expressions.

One obvious way to decide whether someone is a realist or a quasi-realist concerning a certain domain is to ask her if there is a sense in which she would deny the existence of the facts, objects and properties within that domain. Supposing that the person addressed understands the difference between the metaphysically thick and the deflated thin notions of existence, facts etc., which has been observed to be a precondition of a proper understanding of Blackburn's argumentation, we may expect her to reply negatively if she is a genuine realist about the domain, and to say 'yes' if she is merely a quasi-realist about the relevant facts, objects and properties. If this is true, then we may conclude that our understanding of the quasi-realist challenge to the intelligibility and communicability of the metaphysical notions of realism and anti-realism actually guarantees not only the intelligibility, but also the communicability of these positions.

Supposing now that realism is, after all, an intelligible metaphysical doctrine, which concerns the metaphysical status of thinkable objects, properties, conditions and domains, and can be characterised by the claim that these thinkable entities exist (if they do) independently of our actual or potential thoughts or knowledge of this circumstance, I shall finally spell out a systematic ambiguity in which this metaphysical notion is often applied to alternative discourses, as in the case of mathematical or ethical realism. If the truth conditions of our claims within certain discourses can differ from what these claims purport to be about, then there will be two senses in which one can adopt realism (or anti-realism) with regard the discourses under consideration. First, one can subscribe to realism about those objects, properties and conditions that the beliefs in question purport to be about. Second, one can adopt realism about the truth conditions of these beliefs. I will call these metaphysical doctrines, in their generic forms, *realism about subject matters* and *realism about truth*

conditions, respectively, while in their specific forms in the context of particular discourses, for instance, *realism about logical or mathematical objects and properties* and *realism about logical or mathematical truths*, respectively.[33]

In the following chapters, I shall argue that in the case of our paradigm *a priori* claims about abstract domains (as well as in the case of our justifiable claims about causally inert entities in general), truth conditions cannot be identified with intended referents, and therefore a realist construal of the former entities is compatible with any view of the metaphysical status of the latter.

4. Reference

Reference is mostly explicated as a semantic relation between meaningful mental or physical symbols (e.g. concepts, thoughts, or their linguistic expressions), on the one hand, and what the users of these symbols think of or talk about while applying these representations in particular contexts, on the other. There is much disagreement concerning what can be truly said about the nature of this relation, but most people accept that, intuitively, it is an asymmetric "word-world" relation that is essentially associated with intentionality or aboutness.

Of course, the elaboration of a suitable account of the nature of this relation requires a relatively clear view about which symbols under what circumstances can be regarded as referring entities. Unfortunately, opinions about this issue are at least as divergent as those of the extension of the concept of apriority. It is commonly recognised that not all meaningful expressions in our language can refer to something. The key symbols of our non-indicative sentences (e.g. 'why', 'yes' and 'thanks') have arguably no referring power at all. Some words, however, are taken as referring expressions by virtually everyone. The least

[33] For a similar distinction in philosophy of mathematics, see Shapiro (2000), 24-33.

contested examples of the category are proper names, but most philosophers acknowledge the referring capacity of indexicals and natural kind terms as well. Definite descriptions are also amongst the popular examples, though Russellians famously oppose the classification underlying this opinion, arguing that the semantic relation of definite descriptions to the world is less direct than, and radically different from, the one obtaining between a logically proper name and its actual bearer. There is more disagreement about the referring capacity of those symbols which cannot be taken to stand for concrete particulars. For instance, many philosophers reject the view that adverbs and adjectives occurring in a predicate position can refer, for instance, to some properties or relations. Even less intuitive is to assume such a capacity in the case of some logical symbols, such as 'and' and 'nothing'.

Beyond delimiting the circle of symbols that are capable of referring to something, one may wonder also whether reference is conditional upon the actual existence of the intended entity. For instance, one may ask whether it can be reasonably supposed that empty names, names of fictive entities, expressions representing uninstantiated universals, or subject terms in true negative or false positive existentials refer even in the absence of the intended referent-candidates.

Disagreements about exactly which symbols under what circumstances can be claimed to refer to something may have two major sources: one is epistemological, the other semantical. If we apply the term 'reference' in the same sense and still disagree about its applicability in a given situation, then our disagreement must have some epistemological ground. Some of us have presumably failed to recognise the obtaining of the declarative use conditions of the symbol in the actual world. Otherwise, the relevant clash in use can be explained by invoking the semantic fact that we have actually different notions of reference in our mind. Clearly, the controversies indicated above cannot result merely from epistemic mistakes. While applying the term 'reference', philosophers often speak about different asymmetric "word-world" relations that involve intentionality or aboutness.

In one extreme, there are those who hold that reference is a substantive semantic link between a certain class of singular expressions and some concrete particulars that exist in the actual world. For a strict Russellian, for instance, the class of referring expressions includes merely the logically proper names of our language. On this construal, symbols like 'the tallest man in the world', 'is larger than' or 'Little Red Riding Hood' do not qualify as referring entities.

On the opposite side, there are those who construe reference in a very inclusive way. Frege, for instance, assigned referents to all those expressions whose meaning was supposed to contribute to the determination of the truth value of a sentence.[34] Deflationists, in turn, reject any substantive construal of this relation, and maintain that in the suggested deflated sense every expression applied in declarative contexts has a certain referent, which is specified by the pertinent instance of the general schema "a refers to a", whether or not the entity in question exists in the actual world.35 According to this construal, the fact that Little Red Riding Hood is a fictive individual who cannot occur in the actual world does not exclude that we can refer to her by applying her name in various declarative representational contexts.

In this work, I shall apply the terms 'reference' and 'referent' in a relatively inclusive sense.[36] On the construal that I shall adopt here, our ordinary acts of thinking of and speaking about some conditions, objects, properties or domains can all be called reference to these entities independently of whether or not the latter are real and obtaining or existing in the actual world. The term 'referent' will in the meantime be applied interchangeably with that of 'subject matter'. As a consequence, I shall suppose

[34] Frege (1893).

[35] Recent proponents of a deflationist construal of reference include Brandom (1994), Horwich (1998b), Horwich (2005), Field (1994) and Field (2001).

[36] In the following chapters, when I speak about reference, I shall always mean the semantic relation. In order to speak about what a referring symbol actually refers to, I shall apply the term 'referent'.

not merely that definite descriptions, adjectives, adverbs, and logical terms (or the corresponding mental symbols) have a referential link to what they purport to be about, but also that this link holds independently of whether or not the intended entities obtain or exist in the actual world. In other terms, in contrast to the case of truth, where I think our cognitive and linguistic practices undermine the viability of a deflationary construal of the subject, in the case of reference I shall subscribe to such a construal, allowing that a meaningful symbol can refer to entities that are not real or (if real) not existing in the actual world as well.[37]

In line with this understanding, I shall say that a construal of (a certain type of) truth is referential exactly when it specifies the truth conditions of the truth-apt representations under scrutiny in terms of the conditions, or the objects and properties, that these representations, or their basic constituents, purport to be about. For instance, in the case of mathematical truth, a construal is referential if and only if it implies that the truth value of our mathematical thoughts and sentences depends on whether the intended mathematical objects actually exist and possess the intended mathematical properties. It will be in this sense that I shall hold that the truth conditions, and thus the truth values, of our paradigm *a priori* claims about abstract domains (and our presumably justified claims about causally inert entities in general) cannot be suitably regarded as referential in character. In other terms, again, the point that in the following chapters I shall argue

[37] A deflationary understanding of reference does not exclude the reality and actual existence of some referents. More importantly, the adoption of this construal is fully compatible with a strictly realist explanation of the emergence of fixed referential relations between our meaningful symbols and their intended referents. The account may even presuppose our capacity to refer to actually existing entities in a narrower (Russellian or other realist) sense of the term, and require someone's acquaintance with various aspects of the actual world for "grounding" the most basic instances of substantive reference. For an early example of such a substantive theory of reference, see Devitt (1981). In chapters 5 and 7, I shall develop the outlines of such an explanation. If correct, the account will show how our mental and physical symbols can acquire fixed referential relations to fictive or actually non-existent entities as well.

for is that the truth conditions of these representations cannot be specified in terms of those objects and properties that the symbols occurring in them purport to be about.

A further important characteristic that I shall assume of reference in this work is that our referential intentions accompanying the meaning-conferring and meaning-communicating applications of our mental and physical symbols play an essential role in the determination of the nature and identity of what these symbols actually refer to. For instance, the fact that by applying the symbol 'Little Red Riding Hood' we intend to think of or speak about a fictive individual guarantees that no individual in the actual world can be the referent of this representation, or in other terms, that Little Red Riding Hood is essentially (and thus necessarily) fictive (i.e. not real) in character. By the same token, the fact that by applying a symbol in pure logic and mathematics we intend to think of or to speak about an abstract object or property guarantees that our claims or beliefs in pure logic and mathematics cannot refer to anything in the spatiotemporal world, or in other terms, that the subject matter of our representations in pure logic and mathematics is essentially, and thus necessarily, abstract (i.e. non-spatiotemporal) in character.

The intuitive ground of this assumption is relatively clear. At least it seems quite obvious that an appropriate account of, say, why the term 'tree' refers to trees rather than bushes must in some way invoke the circumstance that when the meaning of this term was once taught to us our attention was turned toward trees rather than bushes. Of course, the fact that by applying the symbols of pure logic and mathematics most people intend to refer to some abstract entities does not mean that everyone does or has to do the same. Referential intentions accompanying the application of a certain symbol may vary from person to person and from time to time. In many cases, however, these intentions are relatively well harmonised: by applying a certain symbol most people intend to refer to the same kind of entities most of the time. The greater is the uniformity among these referential intentions, the more compelling it becomes to regard the

commonly intended entities as the standard referent of the relevant symbols.

Some people, however, may still diverge from these common practices, and dispute what most others hold about the nature of some referents. In chapter 3, I shall call these dissidents "revisionists" about the relevant subject matters. In philosophy of mathematics, for instance, it is commonly believed that the subject matter of pure mathematics is abstract in character. Nevertheless, some thinkers have famously queried this consensus, and maintained that mathematics is in fact about the spatiotemporal world. According to the above terminology, they are revisionists about the nature of mathematical objects and properties.[38]

The main problem with this and most other versions of revisionism about various subject matters is that they run counter to an existing (and dominant) cognitive and linguistic practice.[39] Of course, sometimes this practice may turn out to be inconsistent. If by applying the term 'Madagascar' some people intend to speak about the island (or *this* territory over here), some others about the original bearer of the name on the African mainland (or *that* territory over there), and at the same time each of these people about the same territory as all the others, then it is clear that any consistent reconstruction of the defining characteristics of this subject must neglect at least some of the

[38] The classical proponent of this view is Mill (1843), while its most influential recent advocate Kitcher (1984). A further example of revisionism, which in the last few decades became rather influential in philosophy of mathematics, is structuralism. On a structuralist interpretation of arithmetic, for instance, the referent of our numerical term 'number one' is not a particular abstract object, but instead either any (abstract or spatiotemporal) object that occupies a certain position in a certain type of structure (viz. an ω-structure) or the position itself that the previous objects were meant to occupy.

[39] I suppose that our referential intentions accompanying the standard application of a certain symbol can always be communicated by well-chosen claims about the symbol's intended referent. The fact, for instance, that we reject the idea that numbers can occur in the spatiotemporal world is a clear manifestation of the inadequacy of a naturalistic construal of these arithmetic objects.

above intentions.[40] Unfortunately, in so far as our referential intentions are impregnated by some fallible beliefs about the intended referents, they can be never perfectly protected from such inconsistencies.[41] Nevertheless, if the beliefs in question are relatively fundamental, or the intentions under scrutiny are shared by many people, then it is highly unlikely that such inconsistencies remain concealed from most of us for such a long time.

For this reason, in this work, I shall largely neglect those proposals which attempt to save the received referentialist construal of truth in our paradigm *a priori* discourses by adopting a revisionist conception of the nature of the relevant subject matters. Thus, for instance, I shall not argue against those philosophers who deny that the subject matter of our beliefs and claims in pure logic and mathematics is abstract in character. The only versions of revisionism whose implications to our current topic I shall briefly examine are those (highly influential) structuralist construals which nevertheless maintain that the subject matter of our paradigm *a priori* discourses is abstract in character.[42]

[40] The historical example involving this term occurs in Evans (1973) as an illustration of the phenomenon of reference change. By invoking it in the current context, I intend to emphasise that the idea that referential intentions play an essential role in the determination of the nature and identity of referents is compatible with a broadly causal theory of this semantic relation. In chapters 5 and 7, I shall say more about how I think these conceptions can complement each other.

[41] Of course, not everything that we believe about a certain subject enters our referential intentions while we are thinking of this entity. One important difference between reference-fixing and other beliefs in the cognitive and linguistic practice of a given thinker is that the falsity of the former type of beliefs is not conceivable for this thinker.

[42] In chapter 4 and chapter 5, where I discuss the deflationist and anti-realist forms of referentialism about truth, I shall not argue separately against the structuralist versions of these doctrines. I believe that the arguments that I put forward in these chapters equally hold against the corresponding structuralist conceptions as well. In chapter 6, on the other hand, where I examine the viability of platonist (i.e. realist) forms of referentialism in the semantics of our paradigm *a priori* discourses, I shall be more explicit about why I think that the structuralist versions of these doctrines cannot save the adequacy of referentialism either.

Finally, it must be noted that reference in the suggested deflated sense will amount to a very fine-grained asymmetric "word-world" relation. In order to see this, let me briefly compare it with the relation that Frege denoted by this term. On the Fregean construal, two semantically complex symbols, such as the expressions 'the evening star' and 'the morning star', can actually refer to the same thing even if they are composed of semantically non-equivalent constituents. Unquestionably, what Frege conceptualised under the symbol 'reference' is a semantic relation that is exemplified in the actual world. In a coarse-grained sense of the term, the two complex symbols mentioned above are indeed about the same entity: they are both about the planet Venus.

The fact, however, that Frege's notion of reference is an adequate representation of a certain asymmetric "word-world" relation between our meaningful symbols, on the one hand, and what we intend to think of or talk about while applying these representations in particular contexts, on the other, does not mean that there are no other such relations in the actual world to be conceptualised. In the case of our previous example, for instance, there is clearly a sense in which the two expressions, which proved to be coreferential in the Fregean sense, are not exactly about the same aspects of the actual world. While the symbol 'the evening star' is semantically related to evenings, the symbol 'the morning star' to mornings on Earth. As is well known, Frege described this semantical contrast as a difference in sense. Note, however, that every particular exemplification of this fine-grained semantical contrast between two semantically complex meaningful expressions is a difference between the deflated referents of at least some components of the contrasted symbols under consideration. In other terms, what Frege attempted to explain by invoking a platonic semantic correlate of our meaningful expressions is nothing but a (deflated) referential

(i.e. asymmetric "word-world") link between our complex symbols and their respective fine-grained intended referents.[43]

Summing up, my main opponents in this work will be those philosophers who believe that an adequate theory of truth can state or imply that the truth conditions of our truth-apt mental and linguistic symbols are always to be understood in terms of those conditions or entities that these symbols refer to in the above deflated, essentially intentional and fine-grained sense of the term. Contrary to this overarching referentialist construal of truth, in the following chapters, I shall argue that in the semantics of our paradigm *a priori* discourses (and in the semantics of our presumably justified claims about causally inert entities in general) truth cannot be adequately defined as the actual existence or obtaining of the relevant fine-grained intended referents.

5. Natural vs. Abstract Domains

As mentioned before, one of the core tenets that I wish to argue for in the following chapters is that the truth conditions of our paradigm *a priori* claims (i.e. our claims about the abstract domains of pure logic and mathematics) obtain in the natural world, and therefore cannot be specified in terms of the relevant abstract intended referents. In order to ensure the proper understanding of this tenet, in this section, I wish to say a few words about the sense in which I will talk about natural and abstract domains.

Our concepts of nature have been central to philosophy from the beginnings of the discipline. With the development of modern natural sciences and the emergence of naturalism in philosophy, it has become popular to specify the content of our notion of natural domains by reference to the future ontology of

[43] In chapter 4, I shall show how this fine-grained construal enables us to neutralise the so-called "slingshot" arguments, lines of thoughts concluding that all true sentences refer to the same thing, and thus consistently maintain that we may have true representations of different aspects of the world.

our complete empirical sciences.[44] Those who believe that, in an ideal state, all scientific laws and domains will be reduced to those of fundamental physics might maintain that this ontology will be that of our complete physical theory. Others might hold that it will involve the ontologies of all complete empirical sciences. According to these construals, the atemporal domains of pure logic and mathematics, Kantian moral theory, or Fregean semantics, the private domains of a Cartesian theory of mind or phenomenology, and the fictive domains of literature must be equally regarded as non-natural.

One may raise several objections to these specifications of what being natural amounts to. For instance, one may notice that, for many people, the fact that a certain domain is mental or fictive does not imply that it cannot be natural in character. The question whether something is abstract or natural is, for these people, orthogonal to the question whether it is mental or physical, as well as to that whether it is real or fictive. In response to this challenge, the proponents of the previous construals may observe that the existence of alternative concepts of nature does not entail the inadequacy of their analyses, which are meant to be merely of one of these alternatives.

A stronger objection to this type of characterisation is to observe that our ideas of nature are meant to stand for certain domains independently of what our empirical sciences may ever tell us about the actual world. Accordingly, the question whether something is abstract or natural is meant to be independent of the ontology of our complete empirical sciences as well. If this is true, and we suppose that our referential intentions impose essential constraints upon the nature and identity of referents, then the previous characterisations of what it means to be natural cannot be adequate.[45] The moral to be drawn from this insight is

[44] Dewey (1925), Hatfield (1990).

[45] Of course, if the idea of completeness is defined by reference to everything that obtains in a natural domain, then the characterisations may become sound, though circular and thus uninformative.

that our concept of natural domains is better elucidated in non-epistemic terms.

The second way in which people often explicate their notion of natural domains, which clearly satisfies the previous requirement, is by invoking those properties that are supposed to be the essential features of natural entities. There are two characteristics that are often essentially associated with natural conditions or entities: spatiotemporality and causal efficacy. Accordingly, one may adopt two simple construals of what being natural amounts to. On the first construal, a certain condition, individual, property, or event is natural if and only if it is meant to obtain or exist at some spatiotemporal location or locations. On the second interpretation, a certain condition or entity is natural if and only if its obtaining or existence is meant to be causally significant.

The two characterisations are different, since they reduce naturality to different essential properties. Whether the resulting notions of nature differ with respect to their extensions is more controversial. Someone who maintains that causally efficient mental states, events or properties appear in time without possessing a determinate spatial location will certainly deny that the two properties delineate the same domain. Similarly, someone who believes in the existence of causally inert, but nonetheless spatiotemporally located moral or other normative characteristics will definitely reject the idea that the two properties specify the same domain.

In the following chapters, I shall apply the term 'natural' interchangeably with the term 'spatiotemporal'. In other terms, I shall suppose that the domain of natural entities is identical with the domain of conditions, individuals, properties, states and events obtaining, existing or occurring in space and time. Further, in accordance with our received cognitive and linguistic practice, I will also assume that spatiotemporality is a prerequisite for causal efficacy, so that only natural conditions and entities can enter causal relations. Finally, I shall suppose that mental states and events, together with their distinctive characteristics, occur in space as much as in time, and thus they may enter causal relations

with other conditions or entities obtaining or existing in the natural world.[46] With these assumptions in mind, in what follows, I shall argue that in so far as knowledge requires the obtaining of a suitable causal relation between the knowing mind, on the one hand, and the obtaining truth conditions of the relevant true beliefs, on the other, a proper construal of the latter conditions in the semantics of our paradigm *a priori* discourses must be naturalistic in character.

Alternative notions of abstractness have been central to philosophy since the earliest records of the discipline too. In the current literature, there are at least three notions that must be clearly distinguished from each other. The first can be contrasted with the notion of concreteness, the second with the idea of particularity, while the third with the concept of spatiotemporality.

In the first sense of the term, something is abstract if and only if it can exist merely as an aspect of a concrete entity. Autonomous existence, some philosophers say, is the privilege of concrete entities. On this construal, properties of concrete individuals, such as the colour of this rose in front of me, count as abstract entities, while concrete objects that can exist only outside space and time, such as the number one, do not.

In the second sense of the term, something is abstract if and only if it can be fully present at various spatiotemporal locations. On this understanding, universal properties, such as the property of redness in general, are abstract entities, while the particular instances of this property or the particular objects of a non-spatiotemporal domain are not.

Finally, in the third sense of the term, something is abstract if and only if its existence is atemporal (i.e. it exists, if at all,

[46] Note that this construal of what being natural amounts to is compatible with the idea that there might be causally inert entities in the natural world. Moral, aesthetic or epistemological values attributed to spatiotemporal objects are often regarded as examples of this category. Further, the construal does not presuppose the reality of natural entities either. Little Red Riding Hood, for instance, is a fictive individual, but the world in which she is supposed to live her life is nevertheless spatiotemporal, and thus, in the suggested sense, natural too.

outside the natural world). Adopting this construal, atemporal objects and properties, such as numbers, *ante rem* structures, the content of Kantian categorical imperatives, moral values, Fregean senses, propositions, inferential relations and truth values, are abstract entities, while universal properties instantiated in space and time, or the particular instances themselves, are not.

In this work, I shall focus on the semantics and epistemology of discourses about abstract domains in the third sense of the term. This is simply because I intend to provide an account of *a priori* truth and knowledge, and the sense in which the subject matter of our paradigm *a priori* beliefs is meant to be abstract is exactly this third one. In line with this understanding, the semantical tenet repeated at the beginning of this section is meant to claim that the truth conditions of our pure logical and mathematical beliefs about atemporal objects, properties and states of affairs obtain in the spatiotemporal world, and consequently cannot be constituted by these atemporal subject matters.[47]

6. Representations

Having argued against the most influential referentialist understandings of the examined paradigms of *a priori* truth, in chapter 7, I shall put forward and defend a particular representationist form of non-referentialism concerning these entities, according to which the truth value of our paradigm *a priori* beliefs is determined by the obtaining or absence of some conditions in the domain of representations that we develop in our heads in the course of our cognitive engagement with our direct natural environment. The last conceptual element that may call for some elucidation from the advanced characterisation of

[47] In chapters 6 and 7, I shall explain how the conflation of ideas about abstract entities in the second and the third senses specified here may result in serious misconceptions of the role of mathematics in the empirical sciences, and of the way we can justify our beliefs about mathematical entities.

the dialectical purpose of this work is the intended notion of representation. In this section, I shall specify what the suggested representationist construal of *a priori* truth and knowledge presupposes concerning the existence, the nature, and the semantic content of representations.

Representations are commonly understood as syntactically identifiable objects (symbols) that have semantic properties in virtue of their relations to each other and some non-representing conditions or entities in the world. They are mostly supposed to have semantic contents, referential relations, conditions of correct applicability, analytic links, and (if they are truth-apt) also propositional contents, truth conditions, truth values, and various logical relations. This much is acknowledged by virtually everyone who is engaged in talk about representations. There is more disagreement about the metaphysical nature and actual existence of the entities invoked in this characterisation.

At one extreme, there are those who deny the existence of any kind of representation. One might think, for instance, that the world does not include objects with substantive semantic relations, and future science will explain human behaviour without any reference to the manipulation of meaningful symbols. Alternatively, one may argue that the idea of representation presupposes that there are representable entities in the world whose identity conditions are independent of the classificatory work of the mind. The world, however, runs this argument, has no such structure in itself. Thus, by the use of what others take to be mental symbols, in fact, we create, rather than merely represent, the intended aspects of the world.

Most people do not query the existence of representations. Nevertheless, they often disagree about the nature of these entities. Representations are usually regarded as physical or mental objects. Some philosophers, however, maintain that there are abstract (atemporal) representations as well. Physical representations are thought to include various physical symbols, such as pictures, sounds, letters, words, sentences, and patterns of neural activities in human brains. Mental representations are taken to include mental symbols, such as (syntactically

understood) mental images, concepts, thoughts and theories, that can be entertained with different attitudes (toward their semantic content) by human minds.[48] Abstract representations have been postulated as intersubjective counterparts of subjective mental representations or (somewhat confusingly) public reifications of semantic contents, such as (Fregean) senses, concepts, propositions and theories, which may stand in further semantic relations with other entities, such as (Fregean) referents. Clearly, one's belief in the existence of one of these kinds does not imply that one is also committed to the existence of the others.

In this work, I shall suppose that there are both mental and physical representations in the spatiotemporal world.[49] In fact, I shall explain various cognitive and linguistic phenomena by assuming that we can develop within our heads more or less correct mental and physical representations of a number of different domains. This assumption is apparently incompatible with an eliminativist attitude to the standard ontology of folk psychology. According to this view, the generalisations of folk psychology are strictly speaking false, since the mental states and representations posited by this theory do not exist in the actual world. The internal determinants of human behaviour are different from these representational states, and will be revealed by our mature neuroscience. The fate of mental representations, accordingly, will be similar to that of any other entities, like the caloric fluid or the crystalline heavenly spheres, that were posited by the advocates of explanatorily inferior theoretical frameworks:

[48] According to the so-called Representational Theory of Mind, intentional mental states, such as thoughts, beliefs, desires, perceptions and imaginations, are relations of subjects to mental representations, and the intentionality of these states is best understood in terms of the semantic properties of the latter entities. In this work, I shall follow this terminology, and regard the previous kinds of mental states as various attitudes to mental representations. The only divergence that I wish to emphasise here is that by applying the term 'thought' I shall mostly refer to a truth-apt mental representation, rather than to the specific attitude of "entertaining" such an entity as the advocates of RTM tend to do.

[49] In chapter 6, I shall explain why I think that we can never reasonably believe in the existence of platonic entities, and thus in the existence of platonic representations either.

they will be eliminated from the ontology of a more adequate characterisation of the field.[50]

There are two major points that I intend to make in response to the previous observation. First, I admit that the formulations adopted in this work indicate a commitment to the existence of mental representations in the actual world. I chose these formulations, because I believe (with many others) that the considerations advanced in support of eliminativism rely on a false premise concerning the nature of folk psychological concepts, and these concepts (or some others of this kind) will be present in our daily cognitive practice independently of whether or not they are needed in an exact scientific explanation of human behaviour. The premise whose adequacy I question is that our ideas of mental entities are similar to our notion of electron or notion of caloric fluid, which represent unobservable theoretical entities and are part of our conceptual scheme only in so far as they are needed in our best scientific explanation and prediction of some phenomena. In contrast to this picture, I maintain that our notions of mental entities represent properties, states and events that are observable from a first-personal perspective, which means that our belief in the existence of these entities is based on evidence that is independent of our best scientific explanation of overt behaviour.[51]

Note, however, and this is the second major point to be made, that if mature neuroscience retained the idea of correct and incorrect neural representations, and preserved the distinction

[50] The most influential exposition of this view is Churchland (1981) and Churchland (1988/1992).

[51] In chapter 7, I shall argue that our pre-scientific concept of water has to be distinguished from its scientific counterpart that was introduced by the scientific reduction of water to the chemical compound H_2O. On the account that I shall defend there, today we have (at least) two mental representations of by and large the same stuff in the actual world: the new scientific and the old pre-scientific notions of water. If this account is correct, then our pre-scientific concept of water provides an excellent illustration (and thus amounts to a much more plausible paradigm than the eliminated concepts mentioned by Churchland) of what I think we should expect to happen with the vocabulary of folk psychology after the emergence of a mature neuroscientific explanation of overt behaviour.

between those neural mechanisms that underlie the emergence of correct representations within *a priori* discourses, on the one hand, and within empirical discourses, on the other, then an eliminativist could reconstruct the central topic of this work within her own theoretical framework as well. She could understand that my primary aim here is to provide a rough (i.e. scientifically further specifiable) account of the nature of what is currently regarded as correct *a priori* belief formation, something that she could construe within her own framework as correct *a priori* neural representation formation. Although the representations that she would quantify over would differ from those posited by a functionalist, reductive materialist, or a property dualist advocate of folk psychology as the neural correlates of representational mental states in the actual world, her construal of the *a priori* / *a posteriori* distinction could nevertheless still coincide with mine.[52] The coincidence would occur if she maintained that the essential feature that

[52] I suppose that the construal defended in this work is compatible with a functionalist, a reductive materialist and a property dualist position as well in philosophy of mind. The only assumption in it that may impose some constraint upon one's background theory of the nature of mind is the idea that mental properties, states and events occur in space and time and they can be causally related with each other as well as with some non-mental aspects of the spatiotemporal world. By adopting this assumption, one can no longer subscribe to an epiphenomenalist form of property dualism, and one runs counter to Davidson's influential anomalous monism in philosophy of mind too. As is well known, the latter position is based on the idea that mental properties, states and events are ascribed to a subject by others from a third-person perspective and with an interpretative stance towards the subject's overt behaviour, and the ascriptions in question are constrained by the rationality principles of correct interpretation, rather than by some evidence of the hidden causal determinants of the observed behaviour. Putting it otherwise, folk psychological ascriptions are meant to serve the rational, rather than the causal explanation, of overt behaviour. Clearly, this view about the content and function of mental concepts is just as antagonistic with the earlier endorsed belief in the first-person observability of mental properties, states and events as the eliminativist premise of the theoretical character of these entities. If that earlier belief is correct, then, having acquired the content of her mental concepts, each person becomes capable of recognising her own mental states without relying on observations of her own overt behaviour and the principles of rationality. Davidson (1970).

distinguishes our *a priori* neural representations from the empirical ones is that the correct declarative use conditions of the former obtain (if they do) in the domain of neural representations that we develop in our heads in the course of our cognitive interaction with our direct natural environment. In view of this possible agreement, one may conclude that (with some reformulations) the representationist construal of apriority suggested in this work can be reconciled with an eliminativist stance towards mental representations as well.

The opponent may object that eliminativism about the ontology of folk psychology involves the denial of the existence of mental representations, the primary bearers of truth and falsity, and consequently it is hard to see how an eliminativist could reasonably accept any particular account of the nature of *a priori* truth and knowledge. Note, however, that the contrast between *a priori* and *a posteriori* knowledge, justification and evidence is compatible with various conceptions of truth and the nature of primary truthbearers. A physicalistic construal of representations in human (and other) heads does not undermine the idea that the essential feature of *a priori* representations is that the correct applicability of these entities is a function of the obtaining or absence of some representational conditions in the relevant heads.

Beyond assuming the existence of mental and physical representations in the actual world, in what follows, I shall suppose also that the semantics of these representations (i.e. the semantics of thought and language) is compositional. In other terms, I shall suppose that the semantic content of complex representations is determined by that of their constituents and the way they are composed with each other. One may think that this assumption is incompatible with a connectionist conception of the computational architecture of the mind. According to this conception, mental representations are realised by patterns of activation in (natural or artificial) neural networks, and mental processes are the spreadings of such activation patterns through the networks in which they appear. Connectionism is often regarded as the major alternative of the classical view of the

subject, which took mental representations as quasi-linguistic symbolic structures consisting of semantically contentful constituents, while mental processes as rule-governed manipulations of these symbolic structures. The conflict between connectionism and the idea of a compositional semantics emerges, because the computational units in a connectionist network have no separable semantic contents, so their organisation into semantically contentful patterns cannot be compositional.[53]

As some connectionists observed, however, the compositionality of meaning does not imply anything about the semantic properties of the computational units of a representational system. What the adequacy of this principle implies is that the semantic content of a semantically complex representation realised in the system is determined by that of the semantically simple constituents of this representation. If some higher-level patterns of activation in a connectionist network can be identified with the semantically simple, while some others with the semantically complex, representational elements of a classical computational system, and the relation of the semantic contents of these elements corresponds to what is required by compositionality, then the attribution of a compositional semantics to our thought and language proves to be compatible with a connectionist view of the computational architecture of the mind as well.[54]

Speaking about semantic content, a further important tenet that I shall endorse in this work is that the semantic content of a symbol consists in the sum of those conditions whose obtaining or absence in various situative and representational contexts determine whether or not the symbol in those contexts can be

[53] Major advocates of the classical view include Turing (1950), Fodor (1975), Newell and Simon (1976) and Marr (1982), while the chief representatives of the connectionist tradition include McCulloch and Pitts (1943), Rumelhart and McClelland (1986), Smolensky (1988) and Rumelhart (1989). For an apt overview of the debates among classicists and connectionists, see MacDonald and MacDonald (1995) and Millican and Clark (1996).

[54] For a defence of this compatibility, see Smolensky (1989).

correctly applied. In what follows, I shall focus merely on that component of semantic content which is relevant to the correct *declarative* application of representations. After all, my primary aim here is to provide an adequate construal of *a priori* truth and knowledge, which phenomena emerge with the declarative application of our mental and physical symbols. With this specific interest in mind, the assumption just made boils down to the idea that the above component of a symbol's semantic content consists in the correct declarative use conditions of the symbol.[55]

In the theory of meaning, use theories are often contrasted with truth conditional accounts. Further, the contrast is often construed as a clash between anti-realist and realist conceptions of the subject. In chapter 4, I shall argue that these assumptions are false. First, the idea that semantic contents can be identified with conditions of correct use is compatible with the idea that (in the case of non-embedded truth-apt representations) they can be identified with conditions of truth. Second, the idea that semantic contents can be identified with conditions of correct use is compatible with the idea that the latter conditions have to be construed along realist lines. In this work, I shall explore both compatibilities. On the one hand, I shall stipulate that the truth conditions of a certain truth-apt representation are identical with the declarative use conditions of this symbol in "zero representational context" (i.e. in a non-embedded state). On the other hand, to account for the objectivity of truth and correct use in general, I shall also assume that the declarative use conditions of our meaningful symbols obtain (if they do) in the real world (i.e. independently of anyone's actual thought or knowledge of this particular circumstance).[56]

Another crucial assumption concerning the semantic content of representations that I shall rely on in this work is that the declarative use conditions of a mental or physical symbol are

[55] In chapters 5 and 7, I shall explain how I think our mental and physical symbols acquire their relations to their correct declarative use conditions.

[56] In chapter 4, I shall say more about the explanatory considerations underlying this realist construal of the declarative use conditions of meaningful symbols.

in some truth-apt representational contexts not referential. In other terms, in some contexts, they are not identical with those conditions that the symbol in question purports to be about. The general semantical picture underlying this assumption is the following. Each meaningful symbol may have two kinds of semantic relation to its environment. On the one hand, it may have referential links to some conditions that may or may not be real, and if real, that may or may not obtain in the actual world. On the other hand, it may have analytic links to other contentful symbols within the domain of representations that it belongs to. Both kinds of relations are essential to the constitution of the symbol's semantic content. While, in most cases, referential relations associate a symbol with its declarative use conditions in synthetic representational contexts, the analytic links between it and its symbolic environment constitute its declarative use conditions in analytic representational contexts.[57]

Now, there is a much discussed issue in philosophy of mind, which must be addressed by anyone who maintains, as I do, that (at least some part of) the semantic content of mental symbols is constituted of conditions that are supposed to obtain outside the

[57] Earlier in this chapter, I elucidated the sense in which I intend to speak about reference in this work. The concept is meant to stand for a fine-grained semantic relation obtaining between symbols and some entities or conditions that the user of these symbols intends to think of or speak about, independently of whether the latter *relata* are fictive or real, and whether they exist (obtain) or not in the actual world. The concept of analytic relations will be elucidated in chapter 7, where I shall provide an account of the emergence of these relations among some representations in our head. Here and now, it suffices to say that by an analytic link I shall understand a relation between two arbitrary symbols in space and time that ensures that in every synthetic representational context the correct declarative use conditions of one of these symbols are constitutive of the correct declarative use conditions of the other. The analytic link between our concept of bachelor and our concept of man, for instance, ensures that in each synthetic representational context the applicability of one of these concepts implies the applicability of the other. Since analytic links are supposed to obtain independently of anyone's actual thought or knowledge of this circumstance, they can constitute a factual ground for realistically interpreted analytic truths, and substantiate the traditional semantical distinction between analytic and synthetic beliefs.

subject's psychic realm. As some philosophers rightly observed, there is a manifest conflict between this externalist view of the semantic content of mental representations, on the one hand, and the relatively common conviction that psychological explanations are causal and content-based, where the causal factors invoked operate inside the subject's psychic realm, on the other.[58] The standard answer to this problem is to observe that psychological explanations can do without invoking the external components of the semantic content of mental representations, exploring merely the so-called narrow content of these symbols.[59] The assessment of this answer, of course, depends on how we construe, and therewith how we draw the line between, the external and the internal components of content. The looming problem is clearly that psychological explanations seem to invoke not merely the analytic (or inferential) relations, but also the subject matter (the least questionably external part of the semantic content) of their *explanans*.

So, the crucial question to be answered here is whether we can find a coherent characterisation of mental content, which is compatible with the following three observations: first, that most of our mental representations are about conditions that are supposed to obtain outside our psychic realms; second, that the referential aspects of our mental representations (i.e. the subject matters of our beliefs and desires) are highly significant to the psychological explanation of our behaviour; and third, that the obtaining or absence of the intended extra-psychic conditions is entirely irrelevant to the same explanations.[60]

Obviously, the three tenets cannot be reconciled unless there is a difference between the referential aspects of our mental

[58] Stich (1983) has argued that the conflict is best resolved by abandoning the idea that psychological explanations must be content based. Instead, he proposed a syntactic theory of the mind, according to which the semantic features of mental states have no role in the explanations of a scientific psychology.

[59] Classic construals of narrow content include Putnam (1975b), Fodor (1981), Fodor (1987), Fodor (1994), and Block (1986).

[60] It may be relevant to the explanation of the success or failure of our behaviour, but not to the explanation of what we actually do.

symbols invoked in psychological explanations and the extra-psychic conditions of which these symbols purport to be about. After all, the former are supposed to obtain (if they do) within, and the latter outside, our psychic realms. This invites the acknowledgement of the existence of narrow semantic contents, which include the former referential aspects of our mental symbols as well. Note that the existence of these narrow contents may provide us with explanatory resources for a proper account of what happens in reality when a certain subject alternates between two mental states, the entertaining of a thought about a certain condition and then another about another, without thereby bringing about any change in her extra-psychic environment of which these thoughts are supposed to be about. But how should we think about these narrow components of semantic contents?

The first thing that may jump into our mind is that these components can be conceptualised as ideas of those external conditions that the mental symbols in question purport to be about. The problem with this proposal is that it fails to register the difference between narrow contents and mental symbols syntactically understood. One may argue that, in contrast to their various public expressions, mental symbols have no syntactic identity conditions at all. Rather, they are all identified by reference to their narrow semantic contents, which in turn stand in a one-to-one correspondence with those (mostly external) conditions that the symbols in question purport to be about.

Some observations may support this construal of mental symbols. First, in contrast to the case of physical symbols, we never come across a mental representation without knowing something about its semantic content. Second, while a physical symbol can be associated with different semantic contents, the meaning of a mental representation seems, at least on one natural construal, to be devoid of any ambiguity.[61] Note, however, that a

[61] Of course, if mental symbols are identified by reference to the physical symbols that are used for their public expression (so that, for instance, the concept of water is supposed to be whatever happens to govern the use of our term 'water'),

certain semantic content can be associated with more than one mental symbol, as it happens in the case of definitions or the analysis of a concept in terms of some others. If this is true, then the identity conditions of mental symbols cannot be merely semantical in character.

Beyond illuminating the fruitfulness of a syntactic construal of mental representations, the last insight may help us articulate a suitable notion of narrow semantic content as well. Consider again the case when a certain subject alternates between the entertainings of two mental symbols, but now suppose that the symbols in questions have the same semantic content, like a defined concept (e.g. bachelor) and its *definiens* (e.g. unmarried man). Earlier we saw that by invoking narrow semantic contents we can explain how a subject can alternate among the entertainings of mental symbols about different external conditions without inducing any change in the intended environment. Now I wish to observe that the same explanatory perspective can illuminate not merely the difference between narrow semantic contents and the mental symbols that they are associated with, but also the way in which we can identify the former semantic entities as correlates of our mental representations syntactically understood. In the case of the current example, we can assume that whenever a subject alternates among the entertainings of various mental representations of the same external conditions, the narrow semantic content of the symbols entertained is also identical. The internal difference among the contrasted mental states in this case lies in the distinct syntactic characteristics of the chosen mental representations.

Before concluding this section, I wish to raise one more issue whose proper treatment should inform our idea of the nature of representations. In daily cognitive and linguistic practice, we usually distinguish between the possession and the actual entertainment of some representations. For instance, it is

then these contrasts between mental and physical representations will no longer obtain.

commonly held that a subject may possess a certain concept or have a certain belief (an attitude to a certain truth-apt representation) without actually entertaining the symbol in her conscious mind. If I have acquired the notion of unicorn or developed the belief that there are no unicorns in the actual world, then (in some sense) I possess this representation in my head even if I am actually not thinking of unicorns. Now, one may wonder how a certain representation could possibly qualify as mental if it is meant to be able to exist (and be possessed by a subject) even while it is not entertained in its possessor's conscious mind.

A brief reconstruction of this challenge runs as follows. First, it is assumed that concepts and thoughts are mental representations. Second, it is stipulated that mental representations are constituents of conscious mental states. Third, it is supposed that concepts and (various attitudes to) thoughts can be possessed by a subject even when they are not consciously entertained in the subject's mind. Finally, it is observed that the three premises just mentioned cannot be true at the same time, which means that at least one of them has to be eliminated from any acceptable conception of the field.

In what follows, I shall assume that in the actual world people can possess concepts and beliefs without entertaining them in their conscious mind. Whether or not a person possesses a concept or a belief with a certain semantic content will be regarded as a real fact, which obtains independently of anyone's actual thought or awareness of this circumstance.[62] The assumption will be significant, since the construal of *a priori* truth that I shall defend in this work suggests that the truth conditions of our *a priori* beliefs obtain (if they do) within the domain of the representations that we possess in our heads. Beyond this assumption, I shall leave it open whether or not a concept or a belief that is merely possessed by a subject should be classified as

[62] Since narrow contents are supposed to obtain within the subject's psychic realm, I suppose that the possession of a concept or a belief is always a possession of a semantically contentful representation.

a *mental* representation.[63] When it is actually entertained by its possessor, I shall regard it as mental.

Further, I shall also suppose that the entertainment of a concept or a belief is the successful realisation of a (natural) epistemic access to a particular representation or representational state that exists in the head of (i.e. is possessed by) the subject independently of the occurrence of this epistemic event.[64] In this regard, entertaining a mental symbol is similar to perceiving an external object or property. Both activities involve the occurrence of conscious states and presuppose the obtaining of some non-epistemic conditions in the actual world. While in the case of perception the object perceived must exist independently of the occurrence of these epistemic events, in the case of thinking some representations in the subject's head must so exist. An important semantical difference between the contrasted types, however, whose appreciation will play a crucial role in this work, is that, in contrast to the case of perception, the entities (i.e. the possessed representations) accessed by a subject in the course of her conscious thinking need not be identical with the subject matter of the resulting epistemic states. By developing an epistemic access to our concept of zebras, we mostly formulate ideas of the animals, not the accessed representation thereof.

With these observations, I conclude the elucidation of the last central concept of this work. Before proceeding to the next chapter, let me provide a brief summary of what has been done so far.

[63] In order to indicate this neutrality, I shall keep the phrase "the domain of representations in our heads" (as opposed to the phrases "the domain of representations in our brains" or "the domain of representations in our minds") to talk about the location of *a priori* truth conditions in the natural world.

[64] By realising this access, we come to know the semantic content (i.e. the correct declarative use conditions in various representational and situative contexts) of the symbols entertained. Note, however, that our knowledge of this content may be just as partial as our knowledge of other parts of the real world. Discovering facts about the correct declarative application conditions of our concepts in various representational and situative contexts may be just as difficult as discovering facts about the intended referents of these representations.

Summary

In this chapter, I attempted to elucidate the key notions of the construal of *a priori* truth and knowledge that I am to defend in this work.

As a point of departure, I put forward the three main tenets that I will argue for in the chapters to come. The first two concerned the nature of truth within our paradigm *a priori* discourses about abstract domains (i.e. pure logic and mathematics), while the third the nature *a priori* truth, knowledge, justification and evidence in general. The first tenet was that the truth conditions of our paradigm *a priori* beliefs obtain (if they do) within the domains of representations that we develop in our heads in the course of our cognitive engagement with our direct natural environment. The second ascertained that, accordingly, the truth conditions of our paradigm *a priori* beliefs cannot be adequately specified in terms of those conditions or entities that these representations purport to be about. Finally, the third laid down that by reference to the first of the previous two characteristics (viz. representationality), we can provide a useful and adequate real definition of *a priori* truth, knowledge, justification and evidence as well.

After presenting these tenets, in the six consecutive sections of this chapter, I reviewed seven crucial constituents of the given formulations, and elucidated what I shall mean by these terms in the remaining chapters of this work. In particular, I advanced the most important assumptions that I shall adopt, sometimes on explicitly stated considerations and sometimes without further reasoning, of apriority, truth, realism, reference, naturalness and abstractness, and, finally, representation.

I am aware that many of these assumptions are substantial and in no way uncontroversial. Unfortunately, a detailed discussion and defence of these premises would intolerably exceed the confines of this work. The most I could do here is to present them in a clear form. So far, what I was aiming at was not so much persuasion as the elimination of misunderstanding. If the reader of this chapter has managed to acquire a clear sense of

why she might not like the philosophical position that I wish to advocate in this work, then the efforts invested were arguably not in vain.

Having finished this preliminary study on the conceptual fronts of my enterprise, in the following chapter, I shall turn to the articulation of the major methodological assumptions that will inform my argumentation in this work.