Foundations of Cognitive Science, **Oxford, Clarendon Press, 2001**

INTRODUCTION

by João Branquinho

The present volume brings together thirteen new essays dealing with a wide variety of important topics in the foundations of the fascinating multi-disciplinary field of studies currently known as Cognitive Science. The purpose of this Introduction is to provide the reader with a synoptic view of the territory covered by the book, especially the main issues and problems addressed, and to give an outline of the central contributions made by each chapter to their discussion. It is expected that this will help readers, particularly those less familiar with the area, to be able to discern a background of shared theoretical concerns and general assumptions behind the diversity of topics and approaches displayed by the contributed chapters. Given the controversial nature of most issues in the foundations of cognitive science, it could hardly be expected from a description of the territory that it be theoretically neutral; however, we have tried as much as possible to stay close to a set of methodological claims that are very often seen as consensual.

The multi-disciplinary character of the area is clearly reflected in the volume, as witnessed by the fact that a large number of the academic disciplines usually regarded as engaged in the enterprise of cognitive science are represented herein. Indeed, even though some of the chapters (e.g. chapter 5) are somehow hybrid and could thus be seen as falling within more than one discipline, a natural way of sorting them out in that respect is as follows: neuroscience (chapters 4, 5, 6, and 7), linguistics (chapter 10), philosophy (chapters 1, 2, 8, 9, 11, and 13), and psychology (chapters 3 and 12). Of course, the relative predominance of philosophy on this list stems from the relative predominance of matters eminently foundational throughout the book, matters having to do with general questions concerning the relations that hold among the principal protagonists of the story about human cognition: mind, brain, language, world, and action.

It is worth mentioning at this point that, besides sharing a subject (broadly conceived), a common feature of the chapters in this volume is the fact that early versions of them were presented at the Lisbon International Conference on the Foundations of Cognitive Science at the End of the Century. Most people who had the chance to attend the Lisbon meeting in May 1998 would very likely refer to it as a memorable event, not only because of the friendliness of the environment and other traits of the same kind, but mostly on the basis of a feeling that something rather like a "meeting of minds" happened there on the occasion, among the several researchers contributing to the present collection. Indeed, the lively and cooperative discussion and exchange of views that marked most of the conference sessions, as well as the genuinely interdisciplinary dimension of the debates, very often generated real insight into some of the most interesting and difficult issues in the foundations of cognitive science. Although some of these aspects are extremely hard to capture in print, it is not unreasonable to think that what happened in the meeting has had some sort of benign bearing upon the final versions of the essays here included.

THE MIND AS A PROCESSOR OF INFORMATION

How should one characterize the task of cognitive science in a way that would enable us to obtain an integrated picture of the disparate contributions in this volume, in which they are all part of a single theoretical enterprise? A common and convenient way is to define cognitive science as the scientific study of the mind and of its role in the production of intelligent (purposeful, goal-oriented) behaviour.

Although there is some controversy surrounding the issue, one should perhaps note that the term ‘mind’ is usually taken in this context to refer not only to the human mind, but in general to the mind of any intelligent information-processing system. Hence, even machines and artefacts of certain kinds – not to speak of animals – should not be ruled out, at least in the outset, as lacking mental activity in the relevant sense. As long as they can be reasonably counted as intelligent information-processing systems, they are assumed to be endowed with minds. (Whether there really are, or could really be, machines that were capable of meeting that condition is in itself a moot issue in the foundations of cognitive science, and one that need not bother us here.)

The leading idea underlying the above identification of the subject matter of cognitive science is that minds are basically processors of information; or, given that cognition is just (in one sense) information processing, it is the

equivalent idea that minds are essentially cognitive devices. This idea seems to be rather pervasive in contemporary cognitive science. In fact, it even seems to constitute one of the few substantive foundational assumptions in the area that have been relatively immune to dispute so far. In particular, the assumption (or certain versions of it) is clearly endorsed both on classical or symbolic approaches to the mind, and on connectionist or non-symbolic approaches, these being the two main opposing methodological schools of thought currently available in cognitive science. Indeed, almost all those who work in this field would agree in identifying the processes to be studied in cognitive science as being those commonly involved in the information-processing activity of the mind (or of the brain), namely the processes of receiving, storing, retrieving, modifying, and transmitting information of various kinds. Accordingly, virtually everyone working in the area would be prepared to count as paradigmatic instances of information processing everyday phenomena such as visual perception and language comprehension, some important aspects of which are dealt with in chapters 12 and 7 of the present collection (respectively).

Thus conceived, the field of cognitive science covers not only cognition proper, the usual paradigms of which are mental events and processes such as conscious thoughts and inferences and intentional mental states like beliefs and desires, but also any other mental phenomena in which information processing happens to play a central role. Hence, in so far as they can be subsumed as instances of information-processing activity (which in many cases is clearly the case), a variety of mental states and events traditionally grouped under the heading of ‘experience’ and often contrasted *in limine* with cognitive states and events also fall within the scope of cognitive science. Among those states and events are notoriously sensations and perceptions, for instance visual experiences such as the experience of seeing a red object moving around in one’s visual field; and auditory experiences such as the experience of hearing a piano sonata. In other words, both propositional attitudes – as philosophers call psychological states like belief and desire – and sensory experiences are part of the subject matter of cognitive science (note that on some views cognitive scientists should be interested in propositional attitudes just to be able to eliminate them from proper scientific inquiry). Yet, only states of the former kind are strictly cognitive in the following sense. In order to be in one of these states a system must possess and employ an appropriate set of concepts or ways of categorizing things; for instance, a system or organism cannot be in the state of believing that a cow is in front of it without having the concept *cow*, but it certainly can see a cow in front of it (a visual experience) without having such a concept.

So, one way to form a unified picture of the plurality exhibited by the chapters in this volume is to see them all as engaged in the task of explaining some aspect or other of the workings of the mind as an information-processing device. (As we shall see in a moment, this description is only partially accurate and needs qualification; but it will do for immediate purposes.)

REPRESENTATION AND COMPUTATION

The widely held assumption that the mind is basically an information-processing device is normally supplemented by two other relatively uncontroversial foundational claims. These claims complement that assumption by providing further specifications of what is going on as part of the information-processing activity. In particular, they can be seen as providing answers to the following two questions. (a) What in general is processed by the cognising mind (or brain)? In other words, what is the nature of the information manipulated by the mind (or brain)? (b) What is the general form of the processing? What is it like and how does it proceed?

The first claim answers question (a) by specifying the objects over which the information-processing activity is defined as being essentially mental representations. Roughly, these are items in the mind or brain of a given system that in some sense "mirror", or are mapped onto, other items or sets of items; the latter are typically items external to the system: objects, events, situations, etc., in the world (typically, in the immediate environment of the system). Thus, it is assumed that cognition operates on mental representations, in the sense that these are basically what is being processed by the minds or brains of intelligent information-processing systems when they are performing given cognitive tasks. These tasks range from relatively simple ones, such as detecting the presence of a predator in the surroundings, to very complex ones, such as proving a mathematical theorem.

Note that mental representations might come in a wide variety of forms, there being no commitment in the claim itself to a specific kind of representation or to a particular sort of representational vehicle. According to taste, or theory, or purpose, mental representations might be thought of as images, schemas, symbols, models, icons, sentences, maps, and

so on. Their job is to provide systems with the information they need to control their behaviour and guide their interactions with the environment. Mental representations are thus supposed to depict, by means of any vehicle that turns out to be appropriate to do that job, not only aspects and states of the outer world, but also aspects and states of the inner world, namely internal states of the system.

The second claim answers question (b) above by identifying the nature of the processing, by describing in general what goes on inside the "black box" (to use a familiar metaphor). The processing is characterized as being essentially computational, as consisting of a series of computations executed by the system (or by its mind, or brain). Given the first claim, it follows that the computations involved in cognition are defined over a set of mental representations. Roughly, this means that the cognitive tasks performed by the mind or brain characteristically consist in generating, in an effective way, certain mental representations as outputs, on the basis of certain mental representations given as inputs. The operations performed by the visual system or by the language faculty in humans are clear cases of computations in this sense: in the former case they take 2D patterns (retinal images) as inputs, and yield representations of 3D scenes as outputs; in the latter case they take acoustic representations of utterances as inputs, and yield semantic representations as outputs.

Now it might turn out that even these two foundational claims, as well as the underlying conception of the cognising mind as a processor of information, come to be challenged at some point. Yet, as often emphasized in the literature in the area, they have so far enjoyed the special status of shared principles or background assumptions. In other words, they have been regarded as defining the whole enterprise of cognitive science. It comes then as no surprise that both claims are endorsed – even if tacitly in most cases – throughout the present collection, at least if one relies on appropriately generic formulations as the ones provided above. Of special interest in this context is the fact that the claims in question seem to cut across the methodological dispute between symbolic and non-symbolic models of cognition, a dispute that comes to the surface at several points along this volume.

An additional way of forming a unified picture of the variety of contributions and approaches contained in this book is to see them as somehow presupposing these two assumptions about the nature of the cognitive workings of the mind: the assumption that the mind represents and the assumption that the mind computes.

These foundational claims become controversial as soon as one tries to flesh them out, and be more specific both about the nature of the posited mental representations and about the nature of the computations defined over them.

Thus, one might want to hold the view that the mental representations involved in cognition are intrinsically symbolic or linguistic; that they are, to take a familiar proposal, sentences in a "language of thought". One might want to couple this with the sister idea that the computations involved in cognition are like the operations characteristically performed by digital computers, consisting in the application of purely formal rules or procedures to syntactically defined items (language-like mental representations). These views represent, however, specific methodological proposals, particular ways of fleshing out the above claims, and are thus far from being immune to criticism. Indeed, their joint endorsement partially defines the classical or orthodox approach to cognitive science, a view sometimes known as the Representational Theory of Mind. Despite still being very influential, the classical approach is now highly controversial. As far as the present volume is concerned, it is explicitly assumed in James Higginbotham’s chapter (chapter 10), with respect to language understanding, and endorsed in Zenon Pylyshyn’s contribution (chapter 12), with respect to visual representation and object tracking; but it is explicitly rejected in the chapter by Ilya Farber, Will Peterman, and Patricia Churchland (chapter 4) with respect to spatial representation and knowledge.

Another well-known way of fleshing out those claims is given in the rival connectionist picture of the workings of the mind. On this view there is still room for mental representations, as well as for computations performed on them. But connectionist representations and computations are of a radically different ilk. For one thing, mental representations are here definitely non-symbolic and non-linguistic: they do not represent in the way that words and sentences represent. According to usual versions of connectionism, the mind’s information-processing activity should be regarded as being intrinsically linked to, and impossible to divorce from, its physical medium: the brain. Thus, the computational processes involved in cognition are highly sensitive to the physical system in which they are implemented and cannot be assimilated to the purely formal operations performed by digital computers. Rather, connectionist computations proceed through the simultaneous operation of a large number of basic processing devices

– the so-called units, nodes, or artificial neurons (the counterparts of the brain's neurons) – that interact with one another and form a complex multi-layered network. In contrast with standard classical or symbolic computations, which are both local and sequential, connectionist computations are distributed and parallel (at least on some branches of connectionism); semantic content and representational properties might accordingly be assigned simultaneously to several nodes, and so connectionist mental representations are also distributed and not local. Cognition is modelled on the basis of artificial neural networks, structures that mimic in a simplified way the complex operations and processes of the brain. In spite of being highly controversial too, this connectionist picture has been embraced or explored by many cognitive scientists, especially those working in neuroscience and cognitive neurobiology. As far as the present volume is concerned, the approach is explicitly endorsed in the chapter by Farber, Peterman, and Patricia Churchland (chapter 4) and in Paul Churchland´s contribution (chapter 5), where cognitive skills, like those involved in spatial representation and moral discrimination (respectively), are modelled by means of neural networks and given a connectionist treatment.

BEYOND COGNITION?

Notwithstanding what has been said so far, recent developments seem to indicate that cognition is not in fact all there is to the mind that is of interest to cognitive science. This seems to be the case even if one adopts a liberal view about the scope of cognition, a view according to which cognition includes not only the realm of strictly cognitive or "conceptual" events and states, events and states endowed with propositional content, but also the realm of experience (sensation and perception).

The above remark has a certain air of paradox, since the modifier 'cognitive' employed in the term 'cognitive science' suggests a restriction to cognition, to the study of information-processing activity as it occurs in perception, memory, knowledge, language understanding, reasoning, problem solving, and so forth; that is to say, it suggests a restriction to the cognitive side of the mind. Nevertheless, the fact is that a gradually increasing focus has lately been placed within cognitive science upon what prima facie are non-cognitive mental processes, states and events, especially those belonging in the so-called volitional and emotional departments of the mind. (We employ here for convenience the classical tripartite picture of mental life, a picture on which the mind is seen as divided into the separate areas of cognition, conation, and affect.) The paradox is only apparent, of course. For either the mental phenomena in question are in the end correctly describable as genuine instances of information-processing activity and turn out to be at bottom cognitive in nature; or an adequate redescription of the province of cognitive science is needed, a redescription which should inter alia include a revision of the conception of the mind as a processor of information.

At any rate, a considerable amount of research has been going on in cognitive science into the nature of a range of mental phenomena which cannot straightforwardly be reduced to, or do not centrally involve, cognition or information processing.

Emotion is a case in point. Emotional processes have in recent times been at the centre of a great deal of work in cognitive science, work carried out not just in those disciplines where there is a traditional concern with such phenomena – disciplines such as psychology and philosophy – but also in other disciplines, most notably cognitive neurobiology. One of the hot issues here is precisely the extent to which cognition is involved in emotion. Roughly, the debate in the area opposes cognitivist accounts, on which emotion is at bottom a species of cognition, and non-cognitivist accounts, on which emotion cannot be reduced to cognition. That cognition is present in emotion seems to be beyond doubt. As shown by Antonio Damasio in chapter 6, emotional processes invariably contain a clearly cognitive ingredient in the form of an evaluative judgement made by the subject. The problem is rather whether or not cognition, in the sense of information processing, is the essence (so as to speak) of emotion. One should note that the converse issue, the issue about the extent to which emotion is involved in cognition, is also a key issue in the area; it is extensively discussed by Antonio Damasio in his book *Descartes's Error* (A. R. Damasio 1994).

Consciousness, as represented by a range of higher-level mental states traditionally viewed as mysterious in their irreducibility to objective or scientific description and explanation, is also a case in point. Of special interest in this context is the form of consciousness Ned Block calls *phenomenal* consciousness (chapter 1), and in particular the set of sometimes associated properties known as *qualia*. These are the subjective qualities of sensory experience, vivid examples being the allegedly unique phenomenological aspects of the experience of having an orgasm, and the

allegedly unique phenomenological aspects of the experience of listening to the sound of Galician *gaitas* (bagpipes). The extent to which consciousness in general, and phenomenal consciousness in particular, can be given a satisfactory account in terms of the information-processing model is also a matter of intense dispute.

This broadening of the scope of cognitive science, making room for mental phenomena that are not clearly cognitive in nature (even if they turn out to be so in the end), is well-represented in the present collection. Emotion and feeling are addressed by Antonio Damasio in chapter 6, consciousness by Ned Block in chapter 1, and morality in chapter 5 by Paul Churchland. Furthermore, if the main line of reasoning advanced by John Searle in chapter 13 is sound, then modelling human action in terms of information processing, in the sense of construing it as systematically arising out of some appropriate set of desires and beliefs, gives us a wrong picture of the relation between mind (volition) and action.

FOUNDATIONAL ISSUES

Apart from foundational work on methodological issues, including the discussion of background assumptions and guiding principles such as those considered earlier on, there are in the foundations of cognitive science – just as in the foundations of other scientific subjects – two main interrelated kinds of research. The first is concerned with the clarification and elucidation of the most basic or central concepts employed in the discipline. For example, the task of explaining such concepts as *information*, *knowledge*, *concept*, *consciousness*, *cognition*, and so on, belongs in this segment of foundational work in cognitive science. The second involves the investigation of a set of highly general and speculative issues about the nature of mind and cognition, issues that are motivated by developments occurring in the several disciplines contributing to cognitive science. These questions are predominantly philosophical in character and some of them, for example the so-called mind-body problem, have been discussed by philosophers for centuries. Thus, the age-old inquiry into the nature of consciousness, the question of whether it is possible to simulate mechanically intelligent behaviour, and the reflection on whether cognition and thought are possible without language, all belong in this segment of foundational work in cognitive science.

Foundational work of the latter kind ranges from research that is straightforwardly philosophical in nature, in which general questions are explicitly addressed, to research that, although not dealing explicitly with such problems, generates results that are highly relevant to their discussion. A related remark is that, by analogy with foundational work in other scientific disciplines, it is plausible to regard foundational work in cognitive science as being in most cases just a natural continuation of non-foundational work, the differences between the latter and the former being essentially differences in degree of generality and scope and not differences in kind. Accordingly, even though for obvious reasons philosophy should be seen as playing a pivotal role in the foundations of cognitive science, it is by no means accurate to view it as having the monopoly over research in the area. Indeed, it becomes apparent from the previous observations that important foundational work of any of the kinds mentioned above can be (and has been) carried out in any one of the other academic disciplines involved in cognitive science. A clear illustration can be found in the present collection. One can find conceptual work in the foundations of cognitive science (foundational work of the first kind) not only in Block's important distinction between two concepts of consciousness, *phenomenal* consciousness and *access*-consciousness (chapter 1), but also in Antonio Damasio's important clarification of the notions of emotion and feeling (chapter 6); likewise, one can find "philosophical" work in the foundations of cognitive science (foundational work of the second kind) not only in Donald Davidson's defence of a broadly linguistic approach to thought and cognition (chapter 8), but also in Susan Carey's discussion of certain forms of cognition in infants and non-human primates and in her conclusion that they are radically non-symbolic and non-linguistic (chapter 3).

Three research themes have acquired a salient position in foundational work of the second kind (of course, there are natural intersections with foundational work of the first kind and with foundational work on methodological issues). They all relate to the connections between the mind, especially the human mind, and other intervening elements in any adequate account of its workings as a cognitive device. The first theme concerns the relations between the mind and the brain – or, to be precise, the relations between the mind and the central nervous system. A second theme concentrates on the relations between mind and language, by 'language' meaning either a natural language like English, or any other symbolic system of representation. Finally, a third group of issues concerns the relations between the mind and the world, including aspects about the evolutionary history of both and the connections between the mind and intentional behaviour and action upon the world.

These three research themes provide us with a convenient way of sorting out the principal contributions made by each author in the present collection to the foundations of cognitive science.

MIND AND BRAIN

The first group of essays – chapters 1, 4, 5, 6, and 7 – addresses mostly issues concerning the relations between mind and brain.

It is known that mental activity has some sort of physical implementation in the brain. Indeed, there is good evidence that mental states, events, and processes are directly correlated with certain states, events, and processes in the brain – or, in general, in the central nervous system – in the sense that they co-vary with the latter in regular and predictable ways. The search for neural correlates (as they have been called) for a wide range of mental phenomena has always been accorded a key position within cognitive science, especially within those disciplines – the neurosciences – that are particularly suited to the task. More recently, the focus has somehow been shifting from a concern with cognitive phenomena that are more or less familiar in that respect, like language processing or memory, to a concern with prima facie non-cognitive and much less familiar aspects of mental life, like consciousness and emotion.

The inquiry into the neural basis of mental states, events, and processes is regarded not only as being intrinsically valuable but also, on some non-symbolic views of mentality and cognition, as disclosing the nature or essence of these states, events, and processes themselves. These views assume, roughly, that the physical implementation of mental phenomena in certain systems in the brain is something that defines them (at least partially); they contrast thus with symbolic views, on which the neural realization of mental phenomena – the "hardware" on which they are actually "run" – is in general seen as largely irrelevant to their identity.

In chapter 1 Ned Block discusses the issue of the neural basis of consciousness. He begins by noting that, as employed by philosophers and cognitive scientists, the term ‘consciousness’ is ambiguous, as it is used to express different concepts of consciousness: phenomenal consciousness and acess-consciousness. The former is identified with experience: one is phenomenally conscious if one is having an experience; sensations are typical cases of conscious states of this kind. The latter is defined thus: "a representation is access-conscious if it is actively poised for direct control of reasoning, reporting, and action" (p. 3); thoughts and propositional attitudes are typical examples of conscious states of this kind. The case is then made for the distinctness of the concepts by showing that it would be possible for access-consciousness to be instantiated without phenomenal consciousness being thereby instantiated and also, more controversially, that the latter might occur without the former. Yet, as one can see from pairs of concepts such as *water* and H_2O, distinct concepts may nevertheless be co-referential; so nothing would prevent phenomenal consciousness and access-consciousness from being actually correlated with the same system in the brain.

Armed with this distinction Block goes on to examine two views, which he diagnoses as guilty of conflating the two notions of consciousness. One is John Searle’s account of consciousness as not being necessarily involved in habitual, routine, and memorized activities, such as for example driving a car. The other is Francis Crick and Christof Kock’s claim that the area in the brain associated with vision known as V1 is not part of the neural correlate of consciousness. Block contends that their argument faces the following dilemma. If ‘consciousness’ means access-consciousness, then the argument is trivial; if phenomenal consciousness is meant, then it is unsound. Either way, the argument does not succeed in establishing the desired substantive conclusion. However, there are two positive claims made by Block. One is the claim that Crick and Koch’s contention that V1 is not part of the neural correlate of consciousness may in the end be right, but on the basis of independent considerations that are nevertheless implicitly contained in their own work; the other is the claim, also extractable from considerations presented by Crick and Koch, that the two concepts of consciousness may not be co-referential as they may have different (though overlapping) neural correlates.

The nature of the processes involved in emotion and feeling and their neurobiological basis are discussed by Antonio Damasio in chapter 6. Damasio is in general sympathetic to William James’s approach to emotion, especially to James’s idea that the essence of the emotional process is a "sense of the body", an emotion being basically a perception of a set of physiological changes that take place in the body in response to a given situation or event. Nevertheless, James’s account should be subjected to two important kinds of amendment. The first concerns the role of cognition in emotion. According to Antonio Damasio, James generally underestimates that role. In particular, in James’s account no

place is assigned to the clearly cognitive and non-automatic evaluative judgements that precede most cases of emotional response. Damasio argues that a mental appraisal of the significance of the stimuli that cause an emotion is in many cases a crucial component of the emotional process; it is notably present in most emotions we experience as adults. The second correction to James's account concerns the role of the body in emotion. Damasio's claim is that such a role is in a sense overestimated by James. In particular, there is no room in James's account for a neural mechanism that would be able to generate, without the intervention of the body (or else in a supplementary way), that state of awareness of body changes that is the distinctive mark of an emotion. Damasio argues that evidence clearly shows that such a mechanism is available, that the brain has in itself the resources to form an "as-if-body" state (p. 106), a state that depicts the body as if the body were being activated and modified in the way characteristic to emotion. Hence, contrary to James's idea, the interposition of the body in the emotional process is not necessary (even though it is a fact in many cases).

Within the highly complex phenomenon of emotion Damasio makes a subtle distinction between (a) emotion proper, which he identifies as the expressive element, the public manifestation of the phenomenon, and (b) feeling, which he identifies as the internal, subjective experience. The core of an emotion is given in a set of dispositions to respond that induce a collection of changes both in the state of the body and in the state of the brain, these emotional responses being preceded by the mental appraisal process mentioned above. The subject's representation of those body and brain states, which assumes the form of a complex set of mental images, is what constitutes the experiential component – the feeling. According to Damasio, feelings are thus a privileged means the brain has of providing us with knowledge of our body, not in the everyday sense of the term, but in the sense of "cognition of our internal milieu, visceral, and musculoskeletal states"; feelings are nicely described by him as "the first step in letting us mind the body" (p. 107). The presence of both non-cognitive (or automatic) and cognitive elements in the emotional process is paralleled at the level of the brain systems that constitute the neural basis of emotion. Thus, the traditional picture of emotion as being strictly connected with the limbic system and older brain structures, especially the amygdala and the anterior cingulate cortex, is replaced in Damasio's account with a richer picture that also makes room for brain systems that support the *cognitive* aspects of the emotional process; in this picture modern brain structures, in particular the neocortex, are seen as playing a key role.

Damasio's view is that emotion is not an excrescence, a redundant feature of mentality; it has a salient regulatory function and contributes in important ways to the overall success of the interactions between organisms and their environments. Another feature of the mind that is surely no less crucial in that respect, having a conspicuous "survival value", is spatial representation and reasoning. Indeed, a great deal of successful behaviour, behaviour that satisfies our needs and wants, is highly dependent upon our ability to represent space, to perceive the relative positions of independently existing things with respect to our body (or head) and to one another. This ability is involved in the execution of both very modest tasks, such as reaching for an object and grabbing it, and rather sophisticated ones, such as solving a geometry problem. As one can see from simple cases where the ability is exercised, forming an internal representation of space requires in general the joint use of sensory and motor information, as well as the integration of various sensory modalities (sight, touch, etc.). The nature and structure of spatial representation and reasoning, and how it relates to other forms of mental representation, have become a topic of great interest in the foundations of cognitive science.

In chapter 4 Ilya Farber, Will Peterman, and Patricia Churchland focus on the issue of how brains and nervous systems, especially mammalian brains and nervous systems, are able to represent space and carry out spatial reasoning. Their central thesis is that spatial representation, as it occurs in humans and other mammals, is fundamentally non-symbolic. This means that the way in which it happens to be physically implemented in the brain and nervous system, the specific configuration of the underlying neural systems, is a constitutive feature of our ability to represent space, something that determines the intrinsic nature of spatial representation. Given *inter alia* the important position this form of representation occupies in human cognition (for example, it is clearly involved in forms of self-representation such as the representation of one's own body), this thesis is intended to be part of a broader picture of mental representation as nonsymbolic in the above sense. Accordingly, Farber, Peterman, and Churchland mount a vigorous attack on symbolic or linguistic approaches to cognition in general. Their main target is the familiar brand of symbolic approach known as functionalism. Functionalists assert that the nature of mental states can and should be described without any reference to the physical structures in which they are implemented; they are to be individuated rather in terms of their function within a given network of mental states.

In support of their nonsymbolic approach to spatial representation Farber, Peterman, and Churchland draw on three kinds of empirical evidence: (i) behavioural data from animal psychology; (ii) data from neuroscience about the neural basis of spatial representation in mammals; and (iii) data from lesion studies, particularly studies of impairments of spatial representation and reasoning that result from damage to the parietal lobe. They argue that the symbolic view is immediately undermined when confronted with data coming from any of these sources. As to data of kind (i), Farber, Peterman, and Churchland take such data as indicating that spatial reasoning based on representations of object-centred space is carried out in organisms in which it seems absurd to postulate any relevant representational systems that are language-like or symbolic. Baboons, nutcrackers, bears, ravens, and rats are able to solve a host of problems – the hiding problem, the trapeze problem, the problem of finding or retrieving food, etc. – that require complex spatial reasoning, reasoning surely not different in kind from that which is carried out by humans. In particular, in order to solve such problems animals must be able to represent accurately the relation of their own bodies to various objects and other bodies in space. The proponent of the symbolic view is then confronted with the following dilemma: either she claims that there is indeed a difference in kind between human and animal representation of space, which is implausible; or she rules out spatial representation in general as being nonsymbolic, in which case a way would have to be found to discriminate between it and other cognitive skills of the same level of complexity. Data of kind (ii) are taken to show that the mammalian brain and nervous system is able to build representations of object-centred space in accordance with a radically nonsymbolic model. The region of the brain that is responsible for producing "objective" representations of space is the posterior parietal cortex, and the crucial explanatory hypothesis invoked by Farber, Peterman, and Churchland is the Pouget-Sejnowsky Hypothesis. According to this hypothesis, the posterior parietal cortex generates *basis functions*, which compute such representations; a basis function is the product of two kinds of processing units (neurons), eye position units and retinal position units, such a product being computed by hidden units (interneurons in area 7 of the posterior parietal). According to Farber, Peterman, and Churchland, these results from neuroscience provide us with strong evidence against symbolic models of spatial representation. Finally, as to data of kind (iii), the authors claim that impairments in spatial reasoning, such as hemineglect, the tendency to ignore or neglect objects in particular regions of space, are hardly accountable in terms of symbolic approaches to spatial representation (whereas nonsymbolic approaches have no difficulty in explaining them).

Language processing, the way the human mind is able to produce and understand spoken or written language, is a research area in cognitive science that has many significant implications for foundational inquiry into the nature of the cognising mind and its relation to the brain. The neural basis of some central aspects of language processing is the issue addressed by Hanna Damasio in chapter 7. From the perspective of neuroscience, and working at the level of large-scale systems in the human brain, she deals with two sorts of cognitive tasks that are crucially involved in language production and comprehension: word retrieval and concept retrieval. The former consists in the ability competent speakers of English (say) normally exercise when, given a perceptually presented item (a particular person, an animal of a certain kind, an artefact of a certain kind), they recognize it by coming up with a correct word for it or for the kind it belongs to (a proper name like ‘John Smith’, a common noun like ‘cat’, or a common noun like ‘hammer’). The latter consists in the ability we exercise when, given a perceptually presented item (e.g. a cat), we recognize it by describing it in an appropriate way, as having such and such salient features and properties (e.g. ‘small feline that people keep at home…’). As Hanna Damasio points out, the fact that the former ability has not been exercised for some reason (one sometimes forgets the words) is not evidence that concept retrieval has thereby failed: one might have a concept for the presented item, as incorporated in some adequate description one might produce, without being able to produce the name. This is paralleled at the level of the corresponding neural systems: an impairment of the area of the brain that supports the retrieval of names for persons (the left temporal pole) may leave intact the concept-retrieval ability (see p. 110). Of course, word retrieval is taken as being in general sufficient for concept retrieval in the above sense.

Hanna Damasio investigates the neural systems underlying word retrieval and concept retrieval. She draws heavily on studies of two kinds, lesion studies and PET (Positron Emission Tomography) studies, conducted by her, in which recent sophisticated techniques for correlating cognition and brain are applied. These techniques are: functional imaging, by means of which brain activity is indexed when a subject is performing a certain cognitive task; and a modern version of the lesion method, by means of which hypotheses about the neural basis of cognitive processes are tested by considering damaged areas of the brain. It turns out that the results obtained in the lesion studies about the neural systems involved in word and concept retrieval largely coincide with the results obtained in the PET studies.

The general conclusions Hanna Damasio draws from these studies are as follows. First, the traditional anatomical map of language areas, or neural correlates of language processing, comes out as rather incomplete. According to such a map, language correlates in the brain are mostly restricted to two well-known areas (both located in the left hemisphere), Broca's area and Wernicke's area. Hanna Damasio argues that this picture is too simple and should be replaced by a more complex one, a picture in which many other regions of the brain (some of them in the right hemisphere) are involved in language production and comprehension; these regions are connected by bi-directional pathways and form a complex system. For example, the retrieval of names for particular persons seems to correlate with the left temporal pole, an area that is distant from both Broca's area and Wernicke's area. Furthermore, the neural correlates for the retrieval of concepts for persons, on the one hand, and for the retrieval of concepts for tools, on the other, seem to be situated in different hemispheres. Second, the neural correlates for word retrieval with respect to a given item are separable from the neural correlates for concept retrieval with respect to the item in question; word retrieval seems to correlate with regions in higher-order cortices of the left temporal lobe, whereas concept retrieval seems to correlate with regions in "higher-order cortices in right temporal polar and mesial occipital/ventral temporal regions, and in lateral occipital-temporal-parietal regions" (p. 116-17). Third, if we focus our attention just on word retrieval, we see that neural correlates seem to vary according to the conceptual category to which the presented item belongs. The systems in the brain that support word retrieval with respect to persons, word retrieval with respect to animals, and word retrieval with respect to tools seem to be partially segregated from one another; the first systems appear to be located in the left temporal pole, the second systems in the left infero-temporal cortex, and the third systems in the posterolateral infero-temporal cortex. Similar results apply to concept retrieval, which also seems to vary according to the conceptual category of the presented object.

In chapter 5, Paul Churchland deals with the question of how the human brain and nervous system is able to produce such things as, for example, moral virtues (as well as vices, of course), and how the brain is able to generate higher-level mental functions such as those associated with moral representation and knowledge. Approaching the issue from a connectionist standpoint, Churchland's central contention, for which he provides detailed argument and evidence, is that a vast number of significant phenomena involving morality can be given an integrated explanation in terms of neural-network theory. The basic idea is that the concepts and methods of connectionist cognitive modelling, which are employed to account for our performance of a host of natural cognitive tasks (such as visual perception or language processing), can be successfully applied to the moral realm too. Connectionist cognitive modelling proceeds by using simulated networks of simple processing units to explain human cognition. Such units resemble in a number of relevant respects individual neurons in the human brain, so that we thus obtain a model of the cognitive workings of the brain. Of course, the model is oversimplified: for one thing, neurons in living brains clearly outnumber processing units in artificial networks. But it gives us an overall correct picture of cognitive activity as consisting in complex interactions among large numbers of simple processing units.

Paul Churchland focuses on issues in the branch of moral theory known as metaethics. This is *inter alia* the study of moral cognition, of the nature of moral judgement and moral knowledge, and it includes questions about how these are acquired and exercised. A cardinal assumption he adopts, an assumption needed to warrant the extension of neural-network theory to the moral realm, is the view that moral cognition is essentially a set of skills, a set of complex perceptual, reflective, and behavioural skills that a morally knowledgeable adult possesses (p. 79). As such it differs only in detail from any other form of human cognition, most notably scientific knowledge. (One should note at this point that the parallel between moral cognition and progress, and scientific cognition and progress, plays an important role in the main argument of Paul Churchland's paper.) Hence, the formation and deployment of the set of skills embodying moral knowledge and representation should be, in principle, as capable of being modelled by means of neural networks as those skills embodying any other form of knowledge and representation.

Paul Churchland begins by highlighting the merits of a prospective research program based upon a systematic collaboration between the moral discipline of metaethics, on the one hand, and cognitive neurobiology, on the other hand. Such collaboration assumes the form of a bilateral interaction, with results from research undertaken at the micro structural level contemplated by neural-network theory simultaneously feeding and being fed by high-level reflection in the domain of moral knowledge. The program is motivated by an analogy established with the interaction existing between cognitive neurobiology and other philosophical disciplines, especially the philosophy of science, an interaction Paul Churchland sees as theoretically fruitful and insightful. He then goes on to examine in detail, from the point of view of neural-network theory, virtually all of the topics and issues commonly addressed in metaethics: moral

knowledge, moral learning, moral perception, moral ambiguity, moral conflict, moral argument, moral virtues, moral character, moral pathology, moral correction, moral diversity, moral progress, moral realism, and moral unification. Moral knowledge, for example, is described (p. 81) in neural-network terms as being embodied in an intricate configuration of weighted synaptic connections; such connections partition an abstract space of possible activation patterns of neurons or processing units into a hierarchical set of prototypical moral categories ("morally significant action"/"morally nonsignificant action", "morally good action"/"morally bad action", etc.).

Paul Churchland sees the general account that emerges from such an application of connectionist cognitive modelling to the above moral phenomena as belonging in a particular pre-existent tradition or school of thought in metaethics, a tradition known as virtue ethics. Virtue ethics is a family of views about the nature of morality and moral judgement that go back to Aristotle and have been recently endorsed by ethicists such as Mark Johnson, Owen Flanagan, and Alasdair MacIntyre. The leading idea of virtue ethics is that morality is to be fundamentally explained, not in terms of some fixed set of ultimate principles or rules that would govern moral judgement, but in terms of inner characteristics or virtues that individuals gradually acquire on the basis of complex dealings with the social environment. According to Paul Churchland, the account of morality provided by cognitive neurobiology fits nicely with virtue ethics. However, in spite of recognizing a large area of agreement, he closes his chapter with a critical examination of the treatment given to the issue of moral progress by virtue ethics theorists Flanagan and MacIntyre. Both these writers endorse, for different reasons, a frankly sceptical view with respect to moral progress, claiming that there are no real advances in the sphere of human moral consciousness. Paul Churchland argues that such a scepticism is misplaced by drawing on the already noted parallel between moral cognition and scientific cognition, a parallel supported by the findings of neural-network theory; since moral knowledge does no differ in kind from scientific knowledge, to the extent that there is genuine progress in the domain of science, to that very extent there is also genuine progress in the moral realm.

MIND AND LANGUAGE

A second group of contributions – chapters 3, 8, and 10 – can be seen as dealing mainly with issues concerning the relations between mind, especially cognition and thought, and language.

Questions about the mind/language connection, for example questions about the extent to which human thought and cognition are constitutively determined by their usual verbal expressions in a natural language, are nowadays best addressed from a multidisciplinary perspective. They have always been of interest to philosophers, occupying a prominent position in current debates in philosophy of language and in philosophy of mind. Recently they have also attracted the attention of many researchers working in other branches of cognitive science, particularly linguistics, psychology (cognitive and developmental), anthropology, and computer science; a number of significant results obtained in these disciplines have helped shed new light onto a number of traditional problems about the nature of the relation between human cognition and thought, on the one hand, and their linguistic clothing, on the other.

The mind/language connection is very often thought of as bi-directional, as consisting in some sort of mutual dependence between its terms. Yet, this way of looking at the connection may naturally give rise to conflicting views. Taken in the mind-language direction, the connection seems straightforward. Language processing, the ability to learn, use, and understand a natural language, is in some sense a mental activity. The key concept here is meaning, a concept under investigation in many areas of cognitive science. Indeed, the meaningfulness of otherwise meaningless sequences of sounds or marks is something that the mind somehow imposes upon language. But the specific role assigned to the mind in that respect, as well as the nature of the mental and cognitive machinery involved in language production and comprehension, are matters of intense dispute. On some accounts cognition and thought are claimed to be prior to language either metaphysically or in the order of theoretical explanation. These are different, independent priority claims. The priority of thought in the latter sense usually means that linguistic meaning should be explained ultimately in terms of mental states, complex beliefs and intentions in the minds of speakers and hearers; in the former sense, it usually means that there could be thought without language. Taken in the language-mind direction, the connection is sometimes captured in the claim that cognition and thought are shaped by language in the sense that the availability of thoughts of certain kinds is dependent upon the availability of certain linguistic resources adequate to express them; in other words, the claim is that there are thoughts we are able to apprehend or express but would not be able to apprehend or express if we lacked a language endowed with such resources. Likewise, on some accounts

language is claimed to be prior to cognition and thought either metaphysically or in the order of explanation. The priority of language in the latter sense usually means that our capacity for entertaining thoughts, especially our capacity for having propositional attitudes such as beliefs and wishes, should be ultimately explained in terms of certain relations holding between us and linguistic items such as sentences of a natural language; in the former sense, it usually means that thought is not possible without language.

The question of whether language is prior to thought in the metaphysical sense receives an explicit positive answer in Donald Davidson's contribution to this collection (chapter 8) and an implicit negative answer in Susan Carey's contribution (chapter 3). Of course, a genuine theoretical conflict would arise here only if roughly the same thing were meant by the term 'thought' on both views, which is far from being clearly the case.

Carey takes as her starting point a certain set of cognitive primitives, concepts that are basic ingredients in our conceptual scheme or system of representation of the world. She considers two key kinds of basic representational resources: concepts of object and concepts of number. The former include the concept referred to in the psychological literature as the *object* concept, in the sense of our general concept of a *bounded physical object* or *concrete particular*; and also representations of numerical identity and distinctness between objects – one and the same object versus two distinct but physically similar objects – as given in basic quantifiers such as *one*, *another*. The latter include our representations of the series of natural numbers, especially the first integers 1, 2, 3, and the associated counting system. These concepts are expressed in several ways in the lexicon or syntax of any natural language and play an important role in our overall representation of the world.

The crucial foundational questions Carey raises in her paper about the above set of cognitive primitives are as follows. Are these concepts language-dependent, in the sense of presupposing a prior representational system with the structure and resources of a natural language? Are the associated cognitive abilities constitutively linked to the language capacity? Are they necessary parts of the language acquisition device? These questions are addressed from the broader perspective of a general research program in comparative cognition, a program that has been carried out by Carey and Marc Hauser and whose results are reported in this chapter. In such studies the presence or absence in infant human cognition of given representational resources and abilities is compared with their presence or absence in nonhuman primate cognition. The underlying idea is that by identifying the cognitive features in question both from an ontogenetic and from an evolutionary standpoint one obtains a clearer grasp of their nature and of their connection with the language capacity. Thus, two questions have to be answered with a view to arriving at a satisfactory answer to the above questions. First, do prelinguistic human infants have concepts of object and number of the sorts under consideration? Second, are these representational resources available to nonhuman primates? Carey argues that evidence arising from research in comparative cognition indicates that both questions should be given affirmative answers, and hence that our concepts of object and number should be viewed as parts of a core of innate knowledge; such knowledge is claimed to be ontogenetically and evolutionarily prior to the system of symbolic representation embodied in our use and understanding of a natural language.

Carey takes the *object* concept as a sortal concept, a concept that provides us with criteria for individuating and counting the entities falling under it. The individuation principles associated with our concept of a physical object include: (a) principles about object permanence, particularly the idea that material objects are expected to continue to exist independently of us as time goes by and when perceptual contact with them is lost; (b) spatiotemporal principles, especially the idea that two material objects cannot occupy the same portion of space at the same time and the idea that there must be a continuous spatiotemporal path linking successive appearances of one and the same material object. These and other principles are said to be constitutive of our *object* concept. Drawing on empirical studies where the technique known as the *looking time* method is applied both to preverbal human infants and to nonhuman primates, Carey contends that creatures of both kinds can be credited with the concept of a physical object as defined by these features.

Contrary to Jean Piaget's account, on which the *object* concept is intrinsically linked to the language capacity, Carey takes research carried out by Elisabeth Spelke, Karen Wynn, and others to show that the *object* concept, as well as the basic quantificational concepts *one*, *another*, are already available to prelinguistic human infants – babies with ages ranging from 2 and a half months to 3 months. Infants are described as knowing that material objects continue to exist when behind barriers and as making use of spatiotemporal principles such as the ones above for individuating material

objects. A sharp contrast is drawn between the case of the sortal *object* and other object concepts, on the one hand, and the case of specific sortals, concepts such as *book*, *bottle*, and *doll*, on the other; indeed, Carey argues that human infants can only employ the latter range of concepts when the language device is already in place. Furthermore, in a series of studies undertaken by Carey and Hauser the looking time technique was also applied to nonhuman primates, first to wild rhesus monkeys and then to cotton-top tamarins, creatures that are even more distant from us in evolutionary terms. The results obtained turn out to be analogous to those relative to infant cognition: nonhuman primates of both species are described as possessing the sortal *object* concept and the concepts of numerical identity and distinctness associated with the quantifiers *one*, *another*.

As to concepts of number, Carey discusses three proposals concerning the nature of the system used by human infants to represent the first three integers 1, 2, 3: the Numeron List Proposal, the Accumulator Proposal, and the Object File Proposal. She interprets the literature in the area as providing us with evidence in support of the Object File Proposal, an account that distinguishes itself from the other two accounts in virtue of being definitely nonsymbolic; unlike them, the Object File Proposal is not based on the idea that a distinct symbol should be assigned to represent each integer. Carey concludes that "there is no evidence for a prelinguistic representational system of the same structure of natural language count sequences, such as '1, 2, 3…'" (p. 43). However, she is inclined to favour the Accumulator model with respect to the system employed by nonhuman primates to represent integers.

The question of whether speechless creatures can or could have concepts and think is a highly controversial issue in the foundations of cognitive science. As we have seen, from Carey's viewpoint there is a range of basic concepts that are available to preverbal human infants and to nonhuman primates; it seems that such speechless creatures can be credited on that basis with thought and cognition.

A different and *prima facie* opposed approach can be found in Davidson's essay. Davidson investigates the requirements of thought: the conditions that an object – a system, a device, a creature – must in general satisfy in order to be correctly identified or recognized as a cognising thing, in the sense of something that has concepts and is able to employ them in thought. He famously argues, here and elsewhere, that only systems or creatures that are very similar to us (mature human beings) can be legitimately credited with concepts and thoughts. One of the aspects that is taken by Davidson as crucial in establishing such similarity is precisely speech, mastery of a natural language; hence, on his view, speechless creatures in general and animals in particular do not literally have any concepts and are not literally capable of entertaining any thoughts. Of course, we may feel tempted to describe them as cognising things or thinkers; but such ascriptions are only projections that we make on the basis of our own case, they should not be taken literally.

The gist of Davidson's main argument towards that conclusion is as follows. In order for a system or creature to possess concepts, and exercise them in thinking, it must have propositional attitudes: beliefs, desires, intentions, and so on. Propositional attitudes are something we ascribe by essentially using 'that'-clauses to specify propositional contents, for example when we classify a creature as believing *that there is food in the fridge* or as wishing *that the rain would stop*. Now, notwithstanding appearances to the contrary, animals and other speechless creatures are arguably incapable of having any propositional attitudes: no such attitude report can be literally true of them. Therefore, they lack concepts and thought.

In fact, Davidson invites us to equate having concepts and thinking with having propositional attitudes. Concepts are ways of sorting out items in the world, and to have a concept is to be able to, or to be disposed to, class (or not) a given item under the concept. But to be capable of recognizing an item as instantiating (or not) a concept is to be capable of judging or believing that it falls (or does not fall) under the concept; in other words, it is to be capable of having propositional attitudes. A difference should be discerned between recognizing an *F* simpliciter – in the sense of being able to, or being disposed to, discriminate *F*s from non-*F*s – and recognizing an *F* as such (as an *F*); the latter ability, but not the former, requires a creature to be able to, or to be disposed to, judge or believe that the item in question is an *F*. Davidson dismisses in this way the temptation we often have to ascribe concepts and thoughts to animals on the basis of their sometimes rather strong powers of discrimination, on the basis of their apparent ability to distinguish between items that do and items that do not fall under a concept; as he puts it, this ability consists of nothing other than mindless dispositions to respond in specific ways to items that *we describe* as instantiating the concept. Although it can surely be said of animals and other speechless creatures that they are able for example to *see* food in the fridge, it cannot be said of them that they are able to *see that* there is food in the fridge.

Davidson introduces three different but related sorts of requirement for having concepts or propositional attitudes; it turns out that each of those conditions can only be met by creatures endowed with a representational system with the structure and properties of a natural language. The first requirement is based on Davidson's holism about meaning and concept ascription. Concepts are individuated, not only in terms of certain relations they bear to items in the world, but also in terms of certain relations they bear to other concepts; some of the latter relations are entailments, others are relations of evidential support. Thus, one cannot possess a concept or have a belief without possessing a host of other concepts in a network of interrelated concepts or having a host of other beliefs in a web of beliefs. In order to be correctly classified as a thinker a creature must therefore be in possession of a relatively sophisticated conceptual repertoire, a system very much like the one embodied in a natural language. The second requirement is that in order to have concepts a creature must not only be in a position to make occasional mistakes in applying a given concept, for instance judging that a rat is crossing the street while watching a squirrel doing it, but also be in a position to recognize that the creature itself has made a mistake, which is clearly a higher-order propositional attitude. As Davidson puts it, "a creature that cannot entertain the thought that it may be wrong has no concepts, no thoughts" (p. 126); again, this involves crediting thinkers with sophisticated concepts such as the concepts of *error* and *objective truth*. Finally, a third set of requirements is imposed by concept possession upon the structure of thought and upon the structure of any language adequate to express it. Let us mention the most basic of such conditions: (a) the creative property of concepts and thoughts, i.e. the fact that concepts can apply to an endless number of items and the fact that there are infinitely many thoughts to be entertained, requires that such a language contain demonstrative terms and truth-functional connectives; (b) assuming that one cannot have thoughts without a framework for objectual reference and without possessing the general concept of an object, the language should also contain resources adequate to play the role of variables and quantifiers in the usual symbolism of quantification theory (note the contrast between such a concept of an object and the *object* concept envisaged by Carey in chapter 3).

In chapter 10 James Higginbotham's essay (chapter 10) focuses on semantics, the study of linguistic meaning, and its relation to the enterprise of cognitive science. Meaning is a key concept in the study of the mind/language connection. Indeed, an important feature of language processing, our ability to use and understand a natural language, consists in assigning meanings to spoken or written words and sentences of the language on particular occasions; and this is in some sense a mental activity, something that our minds do. How should word and sentence meaning in general be explained? How should one characterize understanding, knowledge of meaning? Aspects of the twin topics of meaning and understanding are the subject matter of Higginbotham's chapter.

Higginbotham approaches the topic of meaning from the standpoint of referential or truth-conditional semantics. According to this view, meaning is to be centrally explained in terms of reference and other related extensional notions, such as satisfaction and truth; all these are language-world relations, holding between linguistic expressions and items or sets of items in the world. Aspects of language use that apparently cannot be accounted for solely in terms of referential properties, for instance racial epithets and euphemism, are nevertheless only understood against the background of reference. According to referential semantics, the meaning of a declarative sentence is given in terms of its truth-conditions, as these are compositionally determined on the basis of referential properties of lexical elements occurring in it and its syntax – the specific mode of combination of such elements in the sentence. The nuclear part of an account of meaning for a natural language takes the familiar Tarskian shape of an axiomatized body of statements, a *theory of meaning* for the language. The axioms of the theory specify referential properties for the primitives of the language, as well as the semantic significance of the available modes of combination; the theorems of the theory state truth-conditions for arbitrary sentences in the language according to their structure and the referential properties of their constituents as specified by the axioms.

The issue Higginbotham then addresses concerns the place that should be assigned to referential semantics within cognitive science. He assumes a familiar view of cognitive science, on which the mental states studied therein are computational in the following sense. They are individuated in part by their content or representational properties, and in part by their causal powers (i.e. their interactions with other mental states and behaviour), the computations involved in the latter being essentially symbolic, i.e. defined over a purely formal basis. The connection between referential semantics and cognitive science so construed is captured in the following two claims. First, understanding or knowledge of meaning consists in genuine propositional knowledge, although mostly tacit, possessed by articulate speakers and hearers of a natural language; it is *knowledge that* as contrasted with *knowledge how*. In other words, there is such a thing as semantic *competence* in Noam Chomsky's sense of the term: a system of internalised

representations of semantic rules and principles. Secondly, the objects of such knowledge are conditions on reference as given by the statements of referential semantics; for instance, able speakers are said to have tacit knowledge of axioms stating the denotation of singular terms of the language. Knowledge of meaning is thus modelled as a complex computational mental state, the inputs to which are the statements of referential semantics as embodied in a meaning theory as outlined above and the outputs of which are "behaviour and adjustments of states that go to exemplify our rationality in the use of language" (p. 147).

These two claims play the role of major premisses in the account of meaning and understanding proposed by Higginbotham. He argues indirectly in their support by examining a set of three alternative proposals that have been influential in semantic theory and by identifying their shortcomings when compared to referential semantics.

The first proposal derives from deflationary or minimalist accounts of reference, as well as of the other semantic notions (satisfaction, truth) that are central to referential semantics. A consequence of such views is that the notion of reference is inadequate to play any explanatory role in an account of linguistic meaning. At least in the version considered by Higginbotham, which he attributes to Paul Horwich, word and sentence meaning are explained rather in terms of concepts and propositions expressed, which are essentially extracted from the way we use words and sentences. According to Higginbotham, the proposal founders because it is incapable of providing us with a satisfactory account of the significance of even such a simple sentential mode of combination as the predication schema; he argues that the proposal is unable to explain how we are able to know the meaning of a monadic predication on the basis of knowledge of the meanings of its components and appreciation of its syntactic structure.

The second proposal rejected by Higginbotham is given in representationalist accounts of meaning and knowledge of meaning, on which the objects of tacit knowledge are taken to be mental representations; he claims that such views turn out to be inconsistent with the nature of linguistic competence and even with the practice of linguistics itself. The third proposal is encapsulated in Jerry Fodor’s semantic views, especially in his thesis that, strictly speaking, natural languages have no semantics; only mental representations, words and sentences in the language of thought, have a semantics (in fact, a referential semantics). Higginbotham discerns three independent ingredients in Fodor’s views: (a) the thesis that to understand a sentence of a natural language consists in mapping it onto a thought, the thought it conveys; (b) the thesis that such a mapping consists in translating the sentence into a sentence of the language of thought, a sentence which expresses the thought in question; and (c) the thesis that learning a natural language is learning how to map its sentences onto thoughts. Higginbotham argues that claims (a) and (b) are in the end consistent with referential semantics construed as a theory of knowledge of meaning; it is rather thesis (c), and the underlying idea that knowing a language is a practical ability, that should be rejected.

MIND AND WORLD

The dominant theme for a third and last group of chapters is the connection between mind and world, a connection of lasting theoretical interest to researchers working in cognitive science and its foundations. The mind/world connection is multifarious and can be unravelled into a number of different relations. It is convenient to use the familiar notion of a direction of fit to introduce some of these relations. One of them is knowledge, under which are subsumed several other central mind/world relations such as perception and memory; knowledge, as well as each of its species, can be described as having the mind-to-world direction of fit, in the sense that it is the mind that has to adapt to the world in order to produce knowledge. Aspects of knowledge are the subject matter of Zenon Pylyshyn’s chapter (chapter 12), in which issues involving visual perception are investigated, and of chapter 11, in which Christopher Peacocke addresses issues about the integration of epistemology and metaphysics within a general theory of concepts. Another central relation between mind and world is action. Action can be described as having the world-to-mind direction of fit, in the sense that it is the world that adapts to the mind when it is acted upon by us and other creatures; issues involving the explanation of action and intelligent behaviour are discussed in John Searle’s chapter (chapter 13).

At a more general level than knowledge, a mind/world relation that has always been at the heart of philosophical reflection, especially in philosophy of mind, is the relation known as Intentionality or Aboutness. (Action is also intentional, but in a different and more usual sense). Many mental states and events are said to be intentional in the sense of being about, or being directed upon, non-mental things; these are typically things in the world: specific objects, events, or situations. Take the mental state someone might be in when she believes that London is pretty; such

a state is said to be intentional both in the sense of having an intentional object, the city of London itself, and in the sense of having an intentional content, that London is pretty. The same goes for a host of other mental states and events such as desires, fears, thoughts, regrets, and so on. Intentionality cannot obviously be defined in terms of direction of it: knowledge is an example of an intentional state with the mind-to-world direction of fit, and desire an example of an intentional state with the world-to-mind direction of it.

Aspects of the problem of intentionality or meaning (in one sense of 'meaning') – the problem of explaining how there can be things in the mind/brain, presumably worldly things, that are about other worldly things – are explicitly addressed in Margaret Boden's chapter (chapter 2) and in Daniel Dennett's chapter (chapter 9).

Dennett launches a sustained criticism of the general picture of intentionality defined by two methodological claims, which he regards as having been omnipresent within the foundations of cognitive science. He labels these claims the *"content capture" assumption* and the *"isolated vehicles" assumption*. The first is a claim about content and (roughly) states that intentional content should be captured in terms of propositions or intensions; although Dennett has in mind mostly linguistic or symbolic representational resources, the claim is taken by him in a broader sense to mean that intentional contents should be directly specified in terms of data structures, a term he employs to cover not only linguistic items like sentences in the language of thought, but also representational devices like images, icons, and maps. The second claim is about the vehicles of intentional content and (roughly) states that it is theoretically possible, as well as desirable, to isolate such vehicles from the "outside" world; in particular, with respect to creatures like us (i.e. embodied nervous systems), there is an important level at which the bearers of intentionality – the *things-about-things* (as Dennett calls them) – can and should be taken in complete abstraction from their connections with the body.

According to Dennett, these two assumptions about intentionality have been rather pervasive in cognitive science. They owe their pervasiveness in part to their original status as useful idealisations or oversimplifications; like any other science at its initial stages of development, cognitive science needs some such oversimplified assumptions in order to get off the ground. However, these assumptions have paved the way for a set of foundational questions about intentionality that are, on Dennett's view, the wrong questions to ask at the current stage of inquiry. One such question is the ontological issue about what sorts of things, what kinds of items in the mind or in the brain, are the bearers of intentionality, the things-about-things. Are they mental representations, for instance sentences in the language of thought, or else non-symbolic items like icons or maps, or are they just complex abilities generated by structures (neural networks) in the brain? From the standpoint of a broadly evolutionary account of cognition and intentionality, an account on which intentional properties of mind or brain components are basically explained in terms of their evolutionary functions, Dennett argues towards the conclusion that these two enabling assumptions should be put aside as they give us a distorted picture of the intentionality present in embodied nervous systems.

Dennett claims that the "content capture" assumption invites us to ask the wrong kind of questions and leads us in the wrong direction. In a nutshell, his argument is that describing the intentional contents of given pieces of cognitive machinery in the mind/brain by specifying certain propositions or certain concepts, the propositions believed or the concepts entertained, does not provide us with any idea of how these mind/brain components are able to carry out their intentional role; it does not give us any idea of how they are able to indicate, or contain information about, particular objects, properties, or events in the environment. In other words, content descriptions of that kind, the model of which is given in explicit expressions of intentional contents, do not provide us with an adequate basis for an explanation of intentional properties; they are bound to be incomplete in this respect. Intentional content should rather be approached in terms of functional descriptions, descriptions of the functions realised by the things-about-things present in embodied nervous systems as such functions are determined by the evolutionary history of the systems. Content specifications given on the basis of the propositional model are utterly unable to do the work; at best, they give us only mnemonic labels for functional structures, labels that merely allude to intentional roles performed by things-about-things.

Appealing as it may be, the "isolated vehicles" assumption also conveys a misguided picture of the mind and its intentional properties. Dennett describes the cardinal idea behind the assumption by means of a suggestive image, an image in which the mind is depicted as a control system and the body as the system controlled by it. Just as the control systems of many familiar machines can be isolated from the controlled systems, so that the latter systems can continue to operate even upon complete replacements of the former systems (for instance remote controls of TV sets can be

replaced without any loss of function); so the mind, although as material as the body, can be isolated from the body in the sense that its central activity as a processor of information is insensitive to the physical medium in which the information is transmitted, processed, and stored (in particular, such activity is not affected by the physics and chemistry of the body). Of course, there must be points of contact between control system and controlled system, channels of information that link the mind to the body and to the "outside" world. They are of two restricted kinds, transducer or input nodes, and effector or output nodes, and their physical composition is surely relevant. Yet, according to the "isolated vehicles" assumption, it is possible to segregate transducers and effectors by regarding them as parts of the controlled system, the body, not of the control system, the mind.

Dennett thinks this is in many respects a seductive picture. Nevertheless, the envisaged segregation, as well as the underlying conception of the mind as an isolated and "clean" control system, fails when applied to creatures like us and other embodied nervous systems. First, unlike the control systems available in many human artefacts, it turns out that minds, in other words the control systems for animals, are shaped in crucial ways by evolution and natural selection; in particular, minds are bound to interact and collaborate with control systems that happen to be much older than them in terms of evolutionary history, such as for instance hormonal systems. According to Dennett, facts of this ilk strongly militate against the envisaged segregation of bodily systems. Secondly, as Dennett aptly puts it, "evolution embodies information in every part of every organism" (p. 139). There are aspects of the interaction between mind and body that suggest that the picture of the mind as the control system of a body could even in a sense be reversed. Indeed, there is evidence that very old bodily systems sometimes guide the central nervous system and play a crucial role in controlling cognitive processes. An interesting case Dennett introduces and discusses is that of the role played by the emotional states we might loosely call boredom and interest in controlling the performance of cognitive tasks by young children; these states are crucial in guarding the cognitive systems of the children against debilitating mismatches. Dennett’s supposition is that there is a very old bodily system, associated with visceral reactions to frustration, which underlies a capacity for boredom or interest and accomplishes the control in question.

Intentionality (or meaning) and evolution are also discussed in Margaret Boden’s chapter, but in the different context of an examination of the connections holding between cognition (or intelligence) and life. This is an important and hot topic in the foundations of cognitive science, naturally related to issues and developments in the disciplines of AI (artificial intelligence), on the one hand, and A-Life (artificial life), on the other. Boden addresses two interrelated questions in her chapter. (1) What is the relation between cognition - as exemplified in perception, knowledge, language, and reasoning - and life? More specifically, is life presupposed by cognition, is it a distinctive mark of cognition, so that only living things could be cognising things? (Note that these are intended as modal questions in the sense that even if, as a matter of fact, there are no non-living things that are capable of cognition, this would not yield affirmative answers to them. The same holds with respect to the questions below.) (2) Could human artefacts, for instance computers or other sophisticated systems, have life or cognition? Actually there are two different questions here: (2a) could there be artefacts that were living things? (2b) Could there be artefacts that were capable of cognition?

On the one hand, Boden argues that whereas question (2a) should be answered negatively, thus contradicting the thesis defended by a number of people working in A-Life, question (2b) should be answered positively, thus vindicating the thesis defended by a number of people working in AI. A consequence of the conjunction of these two claims is that question (1) should receive a negative answer as well. On the other hand, Boden discusses in detail an argument adduced with a view to giving a positive answer to question (1), an argument that depends on the view that meaning and intentionality presuppose evolution.

Boden begins by adopting a characterisation of the notion of life whose essentials go back to Aristotle. According to this notion, the following features define life in the sense that a thing or system must possess each of them in order to be a living thing or system: self-organisation, emergence, autonomy, growth, development, reproduction, adaptation, responsiveness, and metabolism. She then contends that, among these constitutive features of life, self-organisation should be seen as central as it entails any of the other features on the list; indeed, each of them can apparently be extracted from the definition of self-organisation and can thus be seen as a variety of self-organisation.

The connection between life and evolution, particularly the issue of whether evolution should be added to the above list of features, is considered in the context of the already mentioned argument towards the conclusion that life is essential to cognition. The premisses of the argument are the following claims (crudely stated): (a) meaning and

intentionality are necessarily involved in cognition; (b) evolution is necessarily involved in meaning and intentionality; (c) evolution is a criterion of life (note that reading this claim as a sufficient condition and taking evolution as sufficient for life would be enough to run the argument). Boden accepts claim (b) on the basis of a discussion of several accounts of meaning and intentionality. Claiming that naturalistic accounts of meaning should be in general preferred to non-naturalistic ones, she goes on to consider two such accounts: informational semantics, on which meaning and intentional properties possessed by a system or thing are explained in terms of reliable causal correlations between worldly items and internal representations in the system; and evolutionary semantics, on which they are explained in terms of traits of the evolutionary history of the species of the system or thing. She argues that evolutionary semantics, especially the version that has been proposed by Ruth Millikan, is the best naturalistic account of meaning available. It follows from such an account that, as stated in claim (b), only systems or things that have evolved are capable of having meaning or intentional properties. Hence, taking premiss (a) for granted and endorsing (c), evolutionary semantics is committed to the thesis that only living systems or things are capable of cognition.

As a result of her discussion of questions (2a) and (2b), Boden comes to reject this thesis. As to question (2b), she dismisses an argument towards an affirmative answer, an argument against the possibility of ascribing cognition to human artefacts. The argument is inspired in Searle's attack against strong AI: artefactual systems, for instance AI-programs, are incapable of cognition because they necessarily lack genuine intentionality or meaning and meaning is presupposed by cognition (this is claim (a) above). According to Searle, such systems lack meaning because they lack the appropriate kind of interaction with the environment. Boden's reply is that it is not clear that this argument applies to classical robots, and it is even less clear that it applies to evolved, embodied artefacts; in both cases the systems are capable of entering into real causal interaction with the environment. On the other hand, she also rejects a different argument to the same effect, the argument that artefacts cannot have cognition because they are unable to evolve and evolution is essential to cognition. Boden rejects this argument by rejecting its first premiss. Indeed, she argues that human artefacts can be genuinely evolved and mentions a few actual examples: certain creatures that populate virtual worlds, like the inhabitants of the software "Tierra", and certain material robots, some of which are not only evolved but also embodied. As to question (2a), she argues towards the conclusion that human artefacts cannot be alive. Of course, this conclusion is not drawn on the basis of the claim that they cannot evolve; for, as we have seen, she does not accept this claim. Rather, Boden's argument is that human artefacts, even in the case of both embodied and evolved robots, necessarily lack one of the constitutive features of life, namely metabolism.

In his contribution (chapter 13) John Searle deals with issues involving the nature of practical reason, the relationship between rationality and deliberation, and the explanation of human intelligent behaviour. On the one hand, Searle's chapter contains an impressive attack against what he calls the Classical Model of rationality in action, a general conception of rationality and its connection with decision-making that has been largely influent in the philosophical tradition. On the other, the constructive part of his argument consists of a general account of desire-independent reasons for action and their role in practical reason.

According to the Classical Model, human action is invariably explained by means of the belief-desire model, that is to say, in terms of an appropriate set of inter-related beliefs and desires which the agent has on the occasion; such beliefs and desires not only cause the agent's course of action but also explain it – are reasons for it. Thus, there is a straightforward connection between the explanations favoured by the Classical Model and the explanations that are familiar from so-called Folk Psychology, explanations whose theoretical status and value have been widely discussed in the foundations of cognitive science. Here is an example of an explanation of the sort. I suddenly decide to stand up and go to the kitchen *because* I believe there are bottles of mineral water there, I want to drink mineral water, and I believe that I will not get mineral water unless I go to the kitchen; these beliefs and desires, as well as a number of other beliefs and desires that would need to be invoked in a complete description, are taken as reasons for my purposeful behaviour.

Desire occupies a central position in the account of rationality in action embodied in the Classical Model. On the one hand, the presence of desires - construed broadly so as to include wishes, inclinations, preferences, etc. - is needed to generate action (belief alone is insufficient); on the other, rational deliberation consists in selecting the course of action that one recognises as best satisfying our desires. From Searle's perspective, essential to the Classical Model is not (or not only) the claim that desire-dependent reasons can be reasons for action, which merely states a sufficient condition; it is rather (or also) the claim that only desire-dependent reasons can be reasons for action, which states a necessary

condition. Searle does not dispute the sufficiency claim: it seems obvious that very often our actions can be adequately explained in terms of our needs and wants. It is the necessity claim, the idea that desires must be invoked in every adequate explanation of human action, which he challenges.

Searle claims not only that there are counter-examples to the Classical Model, i.e. cases of desire-independent reasons for action, but also that such cases are in fact abundant and can be found in many segments of our social lives. Searle describes the general mechanism operating in these cases, by means of which desire-independent reasons for action are generated, as follows. First, desire-independent reasons are in fact created by free rational agents when they perform certain actions; second, these desire-independent reasons are brought about in virtue of the fact that the agents act with the intention that their actions should create such reasons. Searle introduces two kinds of cases in which desire-independent reasons for action are created in that way: speech acts, like making statements and making promises; and everyday situations where practical reason is involved.

Statement-making gives us the simplest illustration of desire-independent reasons. When one makes a statement one is thereby committed to speaking truthfully. This commitment is not an extra element, additional to the speech act; it is something that is internal to the institution of statement making: a statement is by definition a commitment to the truth of the proposition expressed. Lying, which involves believing correctly that the proposition expressed is false, and being wrong, which involves believing incorrectly that the proposition expressed is true, both presuppose such an internal link between making a statement and being committed to the truth of what it expresses. Therefore, when one makes a statement one has thereby created a reason for oneself to tell the truth. Such a reason does not depend on any pre-existing desires of the agent; in particular, it is independent of any desire to tell the truth, or of any moral preference for the truth or inclination to speak truthfully.

Similar remarks hold with respect to promising, which is discussed at length by Searle. When one makes a promise one is thereby under the obligation to keep it. The obligation is not an extra element, additional to the speech act; it is something that is internal to the institution of promising: promises are by definition creations of obligations. Making insincere promises, promises made with no intention to keep, presupposes such an internal link between making a promise and being under the obligation to keep it. Therefore, given that obligations are reasons, when one makes a promise one has thereby created a reason for oneself to keep the promise. Such a reason does not depend on any desires one might have; in particular, it is independent of any desire one might have to keep the promise, or of any endorsement of a moral principle to the effect that one ought to keep one's promises. Searle discusses three common mistaken ideas about the source of the obligation to keep a promise: the idea that the obligation is prudential, the idea that it derives from our acceptance of moral rules such as the above principle, and the idea that if the obligation derives from the institution of promising, then nothing would prevent any other social institutions, including morally condemnable institutions like slavery, from creating similar obligations. These ideas stem from the Classical Model and from the divorce it is forced to postulate between the act of promising and the obligation to keep a promise. Searle rebuts the third idea by rejecting the premiss that the obligation to keep a promise derives from the institution of promising. Rather, the obligation is entirely created by the agent: when she performs an act of promising she freely and voluntarily creates a reason for herself to act in a certain way (i.e. to keep the promise made). Freedom of the will, which is absent in the case of slavery, is thus essential to the creation of desire-independent reasons for action.

Searle extends these results to cases of practical reason, especially cases where the performance of an act on a given occasion rationally grounds the performance of an act on a later occasion. The mechanism operating in such cases is essentially the one mentioned above. By performing now an act with the intention to create a desire-independent reason to perform a certain act in the future, a free rational agent now creates a desire-independent reason for herself to perform such an act in the future; hence, one can say that the agent now has a desire-independent reason to perform the act in question in the future. For instance, by ordering a beer now and drinking it one has intentionally created a commitment or obligation, and hence a reason, to pay for the beer later on; and such a reason does not depend on any desire one might have on the later occasion to pay for the beer. Searle offers a general theory of the role of desire-independent reasons in practical rationality, a theory on which the following five elements are combined: temporality, freedom, the first-person point of view, language and other institutional structures, and rationality. Temporality comes in because, in practical reason, we organise time in the sense of taking actions we perform now as causally responsible for actions we are supposed to perform in the future, regardless of what we really want to do then. Language and other institutional structures are essential because, as in the case of making a promise, or ordering a beer, there are crucial

features of such structures that enable us to use them as tools in the creation of desire-independent reasons for future action. Freedom and the first-person perspective are essential because we freely and intentionally create reasons for future action when we act in a certain way in the present, on the one hand, and because the reasons so created are reasons for ourselves and not for anyone else, reasons we see as binding our own will in the future, on the other hand. Finally, rationality is essential in virtue of the requirements of logical consistency it imposes; in particular, "the agent should recognise the reason created as binding on his subsequent behaviour" (p. 210).

We move now from action to another central relation between mind and world, knowledge, and more specifically to knowledge acquired by means of visual perception. In chapter 12 Zenon Pylyshyn employs the tools of cognitive psychology to address a set of important foundational issues in the theory of vision. These concern a crucial but often neglected aspect of the connection between our visual representations and the world, namely the way in which particular elements in our visual representations relate to particular objects in the world. The central problem Pylyshyn discusses is an interesting special case of the general problem of intentionality: how the mind is able to "reach out" and represent items in the world.

The problem is this. Consider a visual representation of a scene consisting of a number of particular objects, for example lines, or squares. On the plausible assumption that visual percepts are incrementally constructed by the visual system, two twin questions have then to be addressed. How can a particular part or token of the visual representation - as taken at a certain time or stage of construction - refer to, or pick out, a specific object in the visual scene? How can a token of the representation - as taken at a later time or stage - refer to, or pick out, the very same specific object? More crudely stated, the questions are these. How can a visual representation identify a specific object among a number of objects in the visual field? How can that very object be tracked or re-identified by the visual system as time passes and the representation is being built up? Such questions must be answered because the building up of a representation of a visual scene invariably involves the detection of new properties of the scene; so we must know to which part of the existing visual representation the freshly acquired information should be attached, and that depends on our being able to identify the specific object the information is about. Suppose that information of the form *P(a)*, where *P* is a property (e.g. Redness) and *a* a specific object in the visual field (e.g. a particular square), has been already stored as part of a visual representation of a scene. Suppose we want to update such information with respect to a newly detected property *Q* (e.g. being right-angled). Then we must be able to distinguish between cases in which the updated information has the form of a property conjunction, *P(a)* Ù *Q(a)*, where *Q* is predicated of the very same object, and cases in which it has the form *P(a)* Ù *Q(b)*, where *Q* is predicated of a different object *b* in the visual field. In any case, the visual system must be able to identify or re-identify a specific object as the object the information encoded in a visual representation is about.

Pylyshyn contrasts two lines of explanation of how a reference to a specific object could be secured with respect to a given visual representation. (These lines of explanation have clear counterparts in accounts of linguistic and mental singular reference that are familiar from recent discussions in the areas of philosophy of language and philosophy of mind.) One is the descriptive model, a model according to which visual representations pick out specific objects by means of purely conceptual representations they supposedly provide, i.e. by means of sets of properties the objects uniquely exemplify. The vehicles or symbols that are appropriate to carry out such reference are *descriptors*, mental items that work like singular definite descriptions work in a natural language; thus, a visual representation picks out a specific object as long as this object happens to be the unique satisfier of the set of properties encoded in some descriptor. The other is the demonstrative model, a model according to which visual representations pick out specific objects in a direct manner, unmediated by any purely descriptive or conceptual representations. The vehicles or symbols that are appropriate to carry out such reference are what Pylyshyn calls *visual demonstratives* or *visual indexes*, mental items that work like names (or labels) and demonstratives work in a natural language; thus, a visual representation picks out a specific object as long as this object is demonstratively identified by some visual demonstrative, or is assigned to some visual index.

Pylyshyn's central thesis is that the reference mechanism at work in the case of visual representation is basically non-descriptive and demonstrative; it is the so-called FINST mechanism supplied by his theory of visual indexing (FINSTs, in other words 'fingers of instantiation', are visual indexes). This thesis is argued for not only on the basis of theoretical considerations but also on the basis of empirical evidence provided by a series of MOT (Multiple Object Tracking) studies conducted in Pylyshyn's laboratory.

A number of arguments are adduced against the descriptive model, the most important of which is that descriptive identification is hardly harmonisable with the incremental nature of visual representation. Indeed, if object identification by the visual system always proceeded via the employment of some descriptor, then the information-updating process with respect to a given object would consist in a continual updating of descriptors; it would thus involve a long series of retrievals of individuating descriptions and of the properties encoded therein. Hence, the descriptive model would demand amazing capacities of storage and retrieval, capacities we do not seem to exercise in very simple cases of continued reference to elements in the visual field. In contrast, the demonstrative model has the resources to account in a straightforward manner for the incremental elaboration of visual representations.

On Pylyshyn's view MOT studies provide evidence in support of the FINST theory of visual indexing. In MOT experiments subjects are asked to track a number of independently and unpredictably moving objects in a visual scene consisting of a set of qualitatively identical objects (e.g. circles); the experiments show that subjects are capable of tracking up to five such objects. Since the only way the moving objects are discernible from one another is via their locations, and since evidence shows that successful tracking does not proceed by storing locations and using descriptors based on spatial properties, it follows that visual elements are picked out and tracked directly, independent of any properties they have, including locations. According to Pylyshyn, these studies give credence to the hypothesis that there is in the early (pre-attentive) visual system a primitive mechanism that enables visual representations to pick out and track a small number of objects in a visual scene. This reference mechanism is primitive in the sense of pre-conceptual; in particular, it is prior to focal attention and property detection. Pylyshyn makes two important points concerning the question of how specific objects in a visual scene are linked to visual indexes (note that this is a one-to-one correlation, so that the number of visual indexes active in a visual representation is rather limited: it cannot be greater than five). On the one hand, specific objects are segregated from the background, separated from the rest of the visual field, in a primitive way, a way that is strongly similar to the familiar gestalt separation of figure from ground. On the other hand, visual indexes are attached to specific objects in the visual scene by virtue of data-driven causal chains that extend from these objects to certain symbol structures via primitive mechanisms of early vision; such causal connections constitute the element that ultimately assures that visual representations are able to pick out and track specific objects in the visual field.

The framework provided by the theory of visual indexing enables Pylyshyn to introduce a special category of objects, that of primitive visible objects. Primitive visible objects are by definition those objects that attract visual indexes and allow multiple object tracking. This notion of an object is contrasted with more usual notions, especially those according to which objects are individuated by means of sortal concepts. Moreover, unlike most objects in the usual sense, primitive visible objects are mind-dependent in the sense that they are identifiable in terms of certain features of our early visual system. Pylyshyn also notes some interesting ontological implications of his notion of a primitive visible object; the notion suggests an ontology where objects are assigned a primitive status and places a derivative one, where a conception of places as properties of objects replaces the usual conception of objects as properties of places.

The interrelations holding between epistemological questions, questions about the nature of our knowledge, metaphysical questions, questions about how the world is, and questions about the nature of our concepts and thoughts, are examined by Christopher Peacocke in chapter 11.

Concepts are the constituent units of intentional contents, the propositional contents of thoughts and other intentional mental states and of statements and other speech acts. Substantive theories of particular concepts are philosophical accounts of the nature and role of concepts employed in statements or thoughts belonging to specific domains, for instance mathematical thoughts or statements and modal thoughts or statements. Peacocke's central thesis is that there is a highly general task that substantive theories of particular concepts must carry out: in order to be adequate any such theory must satisfy a general condition that Peacocke calls the Integration Challenge. What a theory of particular concepts must integrate in order to meet the challenge are its metaphysical side, i.e. the account it gives of the truth-conditions of thoughts or statements in the particular domain covered by the theory, and its epistemological side, i.e. the account it gives of the methods by means of which such thoughts or statements can be known by us. Integrating the two aspects means providing an epistemology and a metaphysics for a given area of human thought that are not only separately acceptable but also coherent with one another in a certain specifiable way.

Besides containing a number of considerations meant to support the claim that the Integration Challenge exists and must be addressed by theories of particular concepts, Peacocke offers a detailed proposal concerning how the envisaged integration of metaphysics and epistemology should be in general attained and about what a substantive theory of particular concepts that meets the Integration Challenge would be like. Moreover, the way in which the challenge bears upon psychological theories of concepts is also discussed at length.

Why is the Integration Challenge a challenge? Why isn't the intended integration of metaphysics and epistemology a trivial task? The reason is that there happen to be various domains of thought and discourse with respect to which the tension between metaphysics and epistemology is conspicuous. It has been notoriously hard to reconcile an account of the truth-conditions of thoughts or statements in those areas with an account of their knowledge. Peacocke gives three paradigmatic examples of domains where the gap between metaphysics and epistemology seems difficult to close: modality, knowledge of the intentional contents of our own mental states, and the past. In some cases, for instance modality, the metaphysics is clear, at least if we assume – as we should on Peacocke's view – a realist account of the truth-conditions of modal statements; but the epistemology remains unclear for the truth of modal statements risks to become, on that assumption, epistemically inaccessible. In other areas we have a clear epistemology but an unclear metaphysics; and there are still other areas of thought where neither is clear.

Peacocke goes on to devise a type of strategy that theories of particular concepts might follow with a view to meeting the Integration Challenge. His starting point is a substantive claim about concept individuation, a claim he calls the Linking Thesis. This thesis is in fact a conjunction of two claims: (a) there is a category of concepts, the epistemically individuated concepts, which are individuated by their role in the acquisition of knowledge, i.e. in terms of certain conditions under which intentional contents containing those concepts are known; (2) the category in question is fundamental in the sense that every concept either is itself an epistemically individuated concept or turns out to be partially individuated in terms of certain relations it bears to epistemically individuated concepts. The Linking Thesis establishes thus a tight connection between concepts and knowledge, between the identity of a concept and its epistemology, the way in which intentional contents containing it come to be known.

Peacocke mentions two kinds of epistemically individuated concepts: observational concepts like the concept of a red thing or surface; and logical concepts like the concept of conjunction. He also offers a general characterization of the notion along the following lines: a sufficient condition for a given concept to be an epistemically individuated concept is for that concept to be individuated in terms of its role in judgement, that is to say, in terms of certain conditions under which intentional contents containing the concept in question are judged or accepted by a thinker. This claim is supported by a detailed argument that is then used to unfold a way in which the Linking Thesis (together with a background supposition to the effect that a theory of particular concepts must determine an assignment of truth-conditions to contents containing the concepts) yields a general strategy to address the Integration Challenge. The strategy is applicable to a specific range of cases, namely those cases in which the following two conditions are satisfied: (a) a solution to the integration problem in the area relies fundamentally on truth-conditions; (b) the relevant concepts in the area are epistemically individuated concepts. The leading idea is roughly that it should be possible in such cases, via the Linking Thesis, to extract a determination of truth-conditions for intentional contents containing a target concept from an account of the role of the concept in judgement and hence in knowledge. Peacocke outlines two substantially different styles of solution to the integration problem, two ways of implementing the above strategy. One is the model of *constitutively causally sensitive conceptions*, which is shown to be adequate to account for the case of the past tense. The other is the model of *implicitly known principles*, which is shown to be adequate to account for the case of metaphysical necessity.

Peacocke closes with an extensive discussion of the psychological aspects of the Integration Challenge. He introduces two ways in which a psychological theory of understanding, in the sense of a characterization of a set of mental states and capacities required for the employment of concepts in a specific domain of thought, may fall short of providing a solution to the Integration Challenge. The first consists in what Peacoke designates as the insufficiency of the psychological theory and, on his view, is illustrated by Philip Johnson-Laird's theory about the nature of modal understanding. The second consists in the non-necessity of the psychological theory and is taken to be illustrated by Josep Perner and Ted Ruffman's theory of episodic memory. Peacocke argues that the difficulties afflicting these theories can be overcome only by devising more substantive theories of concepts, theories following the models mentioned above.