Chapter 4

The Continuity of Levels of Nature

Contemporary Functionalism in the philosophy of mind began with a distinction between *role* and *occupant*. As we have seen, the seductive comparison of people (or their brains) to computing machines drew our attention to the contrast between a machine's program (abstractly viewed) and the particular stuff of which the machine happens to be physically made, that *realizes* the program. It is the former, not the latter, that interests us *vis-à-vis* the interpretation, explanation, prediction, and exploitation of the machine's "behavior"; people build computers to run programs, and use whatever physical materials will best lend themselves to that task.

The distinction between "program" and "realizing-stuff," or more familiarly "software" and "hardware," lent itself happily back to the philosophy of mind when Putnam and Fodor exposed the chauvinistic implications of the Identity Theory. What "c-fibers" and the like are doing could have been done—this role could have been performed—by some physiochemically different structure. And sure enough, if the same role were performed, the same functions realized, by silicon- instead of carbon-based neurochemistry, or if our individual neurons were replaced piecemeal by electronic prostheses that did the same jobs, then intuitively our mentality would remain unaffected. What matters is function, not functionary; program, not realizing-stuff; software, not hardware; role, not occupant. Thus the birth of Functionalism, and the distinction between "functional" and "structural" states or properties of an organism.

Functionalism is the only positive doctrine in all of philosophy that I am prepared (if not licensed) to kill for.[1] And I see the "role"/"occupant" distinction (some say obsessively) as fundamental to metaphysics. But I maintain that the *implementation* of that distinction in recent philosophy of mind is both wrong and pernicious. And my purpose in this chapter is to attack the dichotomies of "software"/"hardware," "function"/"structure" in their usual philosophical forms, and to exhibit some of the substantive confusions and correct some of the mistakes that have flowed from them.

1. *The Hierarchy*

Very generally put, my objection is that "software"/"hardware" talk encourages the idea of a bipartite Nature, divided into two levels, roughly the physiochemical and the (supervenient) "functional" or higher-organizational—as against reality, which is a multiple *hierarchy* of levels of nature, each level marked by nexus of nomic generalizations and supervenient on all those levels below it on the continuum.[2] See Nature as hierarchically organized in this way, and the "function"/"structure" distinction *goes relative:* something is a role as opposed to an occupant, a functional state as opposed to a realizer, or vice versa, only *modulo* a designated level of nature. Let me illustrate.

Physiology and microphysiology abound with examples: *Cells*—to take a rather conspicuously functional term(!)—are constituted of cooperating teams of smaller items including membrane, nucleus, mitochondria, and the like: these items are themselves *systems* of yet smaller, still cooperating constituents. For that matter, still lower levels of nature are numerous and markedly distinct: the chemical, the molecular, the atomic, the (traditional) subatomic, the microphysical. Levels are nexus of interesting lawlike generalizations, and are individuated according to the types of generalizations involved. But cells, to look back upward along the hierarchy, are grouped into tissues, which combine to form organs, which group themselves into organ systems, which cooperate—marvelously—to comprise whole organ*isms* such as human beings. Organisms, for that matter, collect themselves into organized (*organ-ized*) groups. And there is no clear difference of kind between what we ordinarily think of as single organisms and groups of organisms that function corporately in a markedly singleminded way—"group organisms" themselves, we might say.[3]

Corresponding to this bottom-up aggregative picture of the hierarchical organization of Nature is the familiar top-down explanatory strategy.[4] If we want to know how wastes and toxins are eliminated from the bodies of humans, we look for and find an *excretory system* interlocked with the digestive and circulatory systems. If we look at that system closely we find (not surprisingly) that it treats water-soluble and nonsoluble wastes differently. We find in particular a *kidney,* which works on soluble wastes in particular. If we probe the details further, proceeding downward through the hierarchy of levels, we find the kidney divided into renal cortex (a filter) and medulla (a collector). The cortex is composed mainly of nephrons. Each nephron has a glomerulus accessed by an afferent arteriole, and a contrac-

tile muscular cuff to control pressure (the pressure pushes water and solutes through the capillary walls into Bowman's Capsule, leaving blood cells and the larger blood proteins stuck behind). Reabsorption and so on are explained in cellular terms, e.g., by the special properties of the epithelial cells that line the nephron's long tubule; those special properties are in turn explained in terms of the physical chemistry of the cell membranes.

The brain is no exception to this hierarchical picture of the organism and its organs. *Neurons* are cells, comprised of *somata* containing a nucleus and protoplasm, and fibers attached to those somata, which fibers have rather dramatically isolable functions; and we are told even of smaller functional items such as the ionic pumps, which maintain high potassium concentration inside. Neurons themselves are grouped into nerve nets and other structures, such as columnar formations, which in turn combine to form larger, more clearly functional (though not so obviously modular) parts of the brain. The auditory system is a fair example. There is evidence that the auditory *cortex* displays two-dimensional columnar organization:[5] columns of variously specialized cells arranged along one axis respond selectively to frequencies indicated by incoming impulses from the auditory nerve, while columns roughly orthogonal to these somehow coordinate input from the one ear with input from the other. The particular sensitivities of the specialized cells is to be explained in turn by reference to ion transfer across cell membranes, and so on down. For its own part, the auditory cortex interacts with other higher-level agencies—the thalamus, the superior colliculus, and other cortical areas—which interactions are highly structured.

Thus do an aggregative ontology and a top-down epistemology of nature collaborate. The collaboration has been eloquently argued for the science of psychology in particular, by Attneave (1960), Fodor (1968b), and Dennett (1975). I shall develop the point at some length, following Lycan (1981a).

2. *Homuncular Functionalism*

Dennett takes his cue from the methodology of certain AI research projects:[6]

> The AI researcher *starts* with an intentionally characterized problem (e.g., how can I get a computer to *understand* questions of English?), breaks it down into sub-problems that are also intentionally characterized (e.g., how do I get the computer to *recognize* questions, *distinguish* subjects from predicates, *ignore*

> irrelevant parsings?) and then breaks these problems down still further until finally he reaches problem or task descriptions that are obviously mechanistic. (P. 80)

Dennett extrapolates this methodological passage to the case of human psychology, and I take it to suggest that we view a *person* as a corporate entity that corporately performs many immensely complex functions—functions of the sort usually called mental or psychological. A psychologist who adopts Fodor's and Dennett's AI-inspired methodology will describe this person by means of a flow chart, which depicts the person's immediately sub-personal agencies and their many and various routes of access to each other that enable them to cooperate in carrying out the purposes of the containing "institution" or organism that that person is. Each of the immediately sub-personal agencies, represented by a "black box" on the original flow chart, is in turn describable by its own flow chart, that breaks *it* into further, sub-sub-personal agencies that cooperate to fulfill *its* purposes, and so on. On this view, the psychological capacities of a person and the various administrative units of a corporate organization stand in functional hierarchies of just the same type and in just the same sense.

To characterize the psychologists' quest in the way I have is to see them as first noting some intentionally or otherwise psychologically characterized abilities of the human subject at the level of data or phenomena, and positing—as theoretical entities—the homunculi or sub-personal agencies that are needed to explain the subject's having those abilities. Then the psychologists posit further, smaller homunculi in order to explain the previously posited molar behavior of the original homunculi, etc., etc. It is this feature of the Attneave/Fodor/Dennett model that ingeniously blocks the standard Rylean infinite-regress objection to homuncular theories in psychology:[7] We explain the successful activity of one homunculus, not by idly positing a second homunculus within it that successfuly performs that activity, but by positing *a team* consisting of several smaller, individually less talented and more specialized homunculi—and detailing the ways in which the team members cooperate in order to produce their joint or corporate output.

Cognitive and perceptual psychologists have a reasonably good idea of the sorts of sub-personal agencies that will have to be assumed to be functioning within a human being in order for that human being to be able to perform the actions and other functions that it performs. Dennett (1978a, chapter 9) mentions, at the immediately sub-personal

level, a "print-out component" or speech center,[8] a "higher executive or *Control* component," a "short-term memory store or buffer memory," a "perceptual analysis component," and a "problem-solving component." And Dennett (chapter 11) examines, in some clinical detail, a multilevelled sub-personal structure that models the behavior that manifests human pain. "Behavior" here must be understood very richly, since Dennett scrupulously takes into account, not just the usual sorts of behavior that are common coin among philosophical Behaviorists and the apostles of commonsense psychology, but subtler phenomena as well: very small differences in our phenomenological descriptions of pain; infrequently remarked phenomena such as the felt time lag between our feeling that we have been burned and our feeling the deep pain of the burn; and (most interesting from the Homunctionalist point of view) the grandly varied effects of a number of different kinds of anesthetics and other drugs on a patient's live and retrospective reports concerning pain. Considerations of these various sorts serve the psychologists (and Dennett) as vivid pointers toward complexities in the relevant functional organization of the CNS, indicating the distinct black-box components at various levels of institutional organization that we must represent in our hierarchically arranged flow diagrams—the kinds of receptors, inhibitors, filters, damping mechanisms, triggers, and so on that we must posit—and the comparably various sorts of pathways that connect these components with each other and with the grosser functional components of their owners such as perceptual analyzers, information stores, and the speech center.

The homuncular approach, teleologically interpreted, has many advantages. I shall recount them when I have said a bit more about teleology. In the meantime, I put my cards on the table as regards the general form of a type-identification of the mental with the not-so-obviously mental: I propose to type-identify a mental state with the property of having such-and-such an institutionally characterized state of affairs obtaining in one (or more) of one's appropriate homunctional departments or subagencies. (The subagencies are those that would be depicted in the flow charts associated with their owners at various levels of institutional abstraction.) The same holds for mental events, processes, and properties. To be in pain of type *T*, we might say, is for one's sub- . . . sub-personal ϕ-er to be in a characteristic state $S_T(\phi)$, or for a characteristic activity $A_T(\phi)$ to be going on in one's ϕ-*er*.

3. *Homunculi and Teleology*

It may be protested that the characterization "ϕ-er" and "$S_T(\phi)$" are themselves only implicitly defined by a teleological map of the organism, and that explications of them in turn would contain ultimately ineliminable references to other teleologically characterized agencies and states of the organism. This is plausible, but relatively harmless. Our job as philosophers of mind was to explicate the mental in a reductive (and noncircular) way, and this I am doing, by reducing mental characterization to homuncular institutional ones, which are teleological characterizations at various levels of functional abstraction. I am not additionally required to reduce the institutional characterizations to "nicer," more structural ones; if there were a reduction of institutional types to, say, physiological types, then on Homunctionalism the identity theory would be true. Institutional *types* (at any given hierarchical level of abstraction) are irreducible, though I assume throughout that institutional *tokens* are reducible in the sense of strict identity, all the way down to the subatomic level.

In fact, the irreducibility of institutional types makes for a mark in favor of Homunctionalism as a philosophical theory of the mental. As Donald Davidson and Wilfrid Sellars have both observed, an adequate theory of mind must, among its other tasks, explain the existence of the mind-body problem itself; this would involve explaining why the mental *seems* so different from the physical as to occasion Cartesianism in the naive, why it has historically proved so difficult even for the sophisticated to formulate a plausible reduction of the mental to the physical, and why our mental concepts as a family seem to comprise a "seamless whole," conceptually quite unrelated to the physiological or the physical family.[9] Homunctionalism provides the rudiments of such explanations. The apparent irreducibility of the mental is the genuine irreducibility of institutional types to the less teleological.[10] The difficulty of outlining a tenable reduction of the mental even to the institutional is due to our ignorance of the organizational workings of the institution itself at a sufficiently low level of abstraction. Nor is the irreducibility of institutional types to more physiological types an embarrassment, so long as our system of institutional categories, our system of physiological categories, and our system of physical categories are just alternative groupings of the same tokens.

Some philosophers might find the Homunctionalist "reduction" very cold comfort. Certainly it would bore anyone who antecedently understands teleological characterizations of things *in terms of* mental items such as desires or intentions. Of course, as the foregoing dis-

cussion implies, I do not understand teleological talk in that way; rather, I am taking mental types to form a small subclass of teleological types occurring for the most part at a high level of functional abstraction. But if so, then how *do* I understand the teleological?

On this general issue I have little of my own to contribute. I hope, and am inclined to believe, that the teleological characterizations that Homunctionalism requires can be independently explained in evolutionary terms. This hope is considerably encouraged by the work of Karl Popper, William Wimsatt, Larry Wright, Karen Neander, and other philosophers of biology;[11] I cannot improve on their technical discussions. However, I do want to make one theoretical point, and then offer one example to back it up.

The theoretical point is that the teleologicalness of characterizations is a matter of degree: some characterizations of a thing are more teleological than others. One and the same space-time slice may be occupied by a collection of molecules, a piece of very hard stuff, a metal strip with an articulated flange, a mover of tumblers, a key, an unlocker of doors, an allower of entry to hotel rooms, a facilitator of adulterous liaisons, a destroyer of souls. Thus, we cannot split our theory of nature neatly into a well-behaved, purely mechanistic part and dubious, messy vitalistic part better ignored or done away with. And for this reason we cannot maintain that a reduction of the mental to the teleological is no gain in ontological tractability; highly teleological characterizations, unlike naive and explicated mental characterizations, have the virtue of shading off fairly smoothly into (more) brutely physical ones.[12]

Let me give one illustration pertinent to psychology. Consider an organism capable of *recognizing faces* (to take one of Dennett's nice examples of a programmable psychological capacity). There is plenty of point to the question of *how* the organism does its job; the creature might accomplish its face-recognizing by being built according to any number of entirely dissimilar functional plans. Suppose the particular plan it does use is as follows: It will accept the command to identify only when it is given as input a front view, right profile, or left profile. The executive routine will direct a *viewpoint locator* to look over the perceptual display, and the viewpoint locator will sort the input into one of the three possible orientation categories. The display will then be shown to the appropriate *analyzer,* which will produce as output a coding of the display's content. A *librarian* will check this coded formula against the stock of similarly coded visual reports already stored in the organism's memory; if it finds a match, it will look at the identification tag attached to the matching code formula and show the tag to the organism's *public relations officer,* who will give phonological

instructions to the *motor subroutines* that will result in the organism's publicly and loudly pronouncing a name.

Knowing that this is the way in which our particular face-recognizer performs its job, we may want to ask for further details. We may want to know how the viewpoint locator works (is it a simple template?), or how the PR office is organized, or what kinds of subcomponents the analyzer employs. Suppose the analyzer is found to consist of a *projector,* which imposes a grid on the visual display, and a scanner, which runs through the grid a square at a time and produces a binary code number. We may go on to ask how the scanner works, and be told that it consists mainly of a light meter that registers a certain degree of darkness at a square and reports "0" or "1" accordingly; we may ask how the light meter works and be told some things about photosensitive chemicals, etc., etc. Now at what point in this descent through the institutional hierarchy (from *recognizer* to *scanner* to *light meter* to *photosensitive substance,* and as much further down as one might care to go) does our characterization stop being teleological, period, and start being purely mechanical, period? I think it is clear that there is no such point, but rather a finely grained continuum connecting the abstract and highly teleological to the grittily concrete and only barely teleological. And this is why the mental can *seem* totally distinct and cut off from the physiochemical without *being*, ontologically, any such thing.[13]

A final word about my reliance on barely explicated teleology: I do not claim that barely explicated teleology is good or desirable. I do not like it at all, myself. My point is only that the mystery of the mental is *no greater than* the mystery of the heart, the kidney, the carburetor or the pocket calculator. And as an ontological point it is a very comforting one.[14]

4. *Advantages of the Teleological Approach*

The reader will not have failed to notice that I take *function* very seriously and literally: as honest-to-goodness natural teleology.[15] The policy of taking "function" teleologically has some key virtues: (i) As we have seen, a teleological understanding of "function" helps to account for the perceived *seamlessness* of the mental, the interlocking of mental notions in a way that has nothing visibly to do with chemical and physical concepts.[16] (ii) By imposing a teleological requirement on the notion of functional realization, we avoid all of chapter 3's counterexamples to Machine Functionalism, and, I would claim, to any other version of Functionalism; see below. (iii) A teleological functionalism also helps us to understand the nature of biological and

psychological *laws*, particularly in the face of Davidsonian skepticism about the latter (Lycan, 1981c; Cummins, 1983). (iv) If teleological characterizations are themselves explicated in evolutionary terms, then our capacities for mental states themselves become more readily explicable by final cause; it is more obvious why we have pains, beliefs, desires, and so on.[17] (v) The teleological view affords the beginnings of an account of *intentionality* that avoids the standard difficulties for other naturalistic accounts and in particular allows brain states and events to have *false* intentional content. Causal and nomological theories of intentionality tend to falter on this last task (cf. chapter 6, and see Lycan, forthcoming).

I have argued above that we need a notion of teleology that comes in degrees, or at least allows for degrees of teleologicalness of characterization, and that we already have such a notion, hard as it may be to explicate—recall the examples of the face-recognizer and the key. Philosophers may differ among themselves as to the correct analysis of this degree notion of teleology—for my own part, I tend to see the degrees as determined by amenability to explanation by final cause, where explanation "by final cause" is reconstrued in turn as a sort of evolutionary explanation (though some details of this remain to be worked out). But two main points are already clear: (i) At least for single organisms, degrees of teleologicalness of characterization correspond rather nicely to levels of nature.[18] And (ii) there is no single spot *either* on the continuum of teleologicalness or amid the various levels of nature where it is plainly natural to drive a decisive wedge, where descriptions of nature can be split neatly into a well-behaved, purely "structural," purely mechanistic mode and a more abstract and more dubious, intentional, and perhaps vitalistic mode—certainly not any spot that also corresponds to any intutive distinction between the psychological and the merely chemical, for there is too much and too various biology in between.

My own panpsychist or at least panteleologic tendencies are showing now. Many tougher-minded philosophers will find them fanciful at best, and of course (in my lucid moments) I am prepared to admit that it is hard to see any use in regarding, say, *atomic*-level description as teleological to any degree;[19] certainly explanation-by-final-cause does not persist all the way down. *But:* Unmistakably teleological characterization (description that is obviously teleological to some however small degree) persists *as far* down as could possibly be relevant to psychology (well below neuroanatomy, for example). And the *role/occupant* distinction extends much further down still. Thus the vaunted "function"/"structure" distinction as ordinarily conceived by philosophers fails to get a grip on human psychology where it lives.

For that matter, ironically, the "function"/"structure" distinction applies in no unproblematic way to computers themselves. Just as the good old "analog"/"digital" distinction has been seen in recent years to be vexed at best,[20] even the "software"/"hardware" distinction *as it is literally applied within computer science* is philosophically unclear. (There is a nice paper yet to be written on this issue, entirely disregarding the philosophy of mind or Artificial Intelligence.) Note first that "software"/"hardware" does not (even *prima facie*) coincide with "*program*"/"hardware."[21] According to one current usage as I understand it, "software" is what is electronically alterable, paradigmatically *packaged* input such as is loaded into memory from a disk drive (or perhaps entered from the keyboard), while "hardware" is whatever is hard-*wired* or fixed in such a way that it can be altered only by physical snipping and resoldering inside the machine. *This* distinction (by whatever name) obviously does not coincide with the "program"/"realizing-materials" distinction, for what is intuitively and universally designated a *program* may be either loaded from without (as a way of structuring the previously "blank" memory) or entirely hard-wired *ab initio*. A non-"programmable" pocket calculator, e.g., uncontroversially *has* a program that computes arithmetical functions—calculators of different brands have different programs—but all these programs are hard-wired and unalterable from the keyboard; they are not "software" in the computer scientists' sense. Similarly, some computers have hard-wired programs corresponding to other brands' applications software: a dedicated machine might have Wordstar or some other word-processing program built in unalterably rather than loaded from a disk in the more customary way.[22]

Much more to the point, there is not, even in a particular computer, a single program that is "*its*" program; there is no one level of programming. We constantly hear talk, especially from philosophers of mind discussing functionalism pro or con, about a computer's program as opposed to the hardware that is realizing "it," but this is a misconception; in computer science—as in botany and zoology—there is a continuum or hierarchy of levels of organization rather than a two-levelled structure. Flipflops are grouped into banks and registers. In an 8-bit machine there are 2^8 possible settings in each unit, each of which can be expressed in binary machine code or alternatively in hex notation. A level higher, assembly language collects individual machine-code operations into often-used sequences and allows the defining of subroutines and the giving of function names; a *macro*-assembler introduces variables, affording a *library* of generically characterized subroutines (without knowing, e.g., the exact register locations that will be specified in machine code). A standard

programming language such as BASIC or PASCAL or C can then be similarly constructed out of assembly language; PASCAL commands call sequences of machine-code instructions. Further programs are written "in" PASCAL or the like by the same aggregative process. And there are special-purpose and/or still higher-level languages (as they are happily called), including self-compilers and the like, based on the simpler and more general languages.[23] A programmer can program at whatever level suits the purposes of the moment. (I program, in an infantile sort of way, in BASIC, but there are people who work primarily in assembly language, and it is entirely possible though pointless and self-punishing to program directly in machine code; professional programmers, I am told, are likely to start a tiny seed in machine code or in assembly language and then bootstrap crazily through permutations of compilers and self-compilers.) But—to get to the point—which level of description of the machine's operation counts as "the" program, as opposed to the mechanical stuff that realizes "the" program, is entirely observer- and interest-relative. The question, "What program is the machine now running?" has more than one answer: "Do you mean in assembly language, in BASIC, in C, or in [say] PILOT?"—and the *preferred* answer will vary in context according to interests and purposes. My moral is that the absolute "function"/"structure" distinction, borrowed from automata theory by philosophers and then misapplied to living organisms, does not even apply to computers in the real world; there too, the distinction (though real enough) is *relative* to level of organization, though due to human artifice computers do not exhibit the same degree of physical modularity that organisms do.

Incidentally, there is an interesting terminological point[24] to be made about the coordinately contrastive expressions "function" and "structure" (which usage dates, I believe, from Putnam's "Minds and Machines," 1960): "Structure" is (when you think about it) a surprisingly organizational, I would say almost explicitly teleological, term; a *structure* is an organized collection of elements, somehow *held in place* and/or serving to hold other things in place for some purpose or other. It does not contrast markedly with "function," even though it is not synonymous with it (and even though "*a* structure" normally *serves* a function). How might we better express the notion of brute, primitive realizing-stuff that does or is supposed to contrast with the functional? We might try "functional"-as-opposed-to *purely mechanical*. "Mechanical"? Hardly—*mechanisms* are functional items *par excellence*. Purely . . . what? We are in search of prime matter here, or else perhaps Sartre's yucky grey dead matter. And that stuff, if there is any, can be characterized in either of only two ways: by contrast with

the functional, at some chosen level of abstraction, as in "the stuff that cells are made of," or by direct reference to a specific kind of level-bound entity, such as molecules or atoms or quarks and gluons. In neither way do we succeed in isolating the desired *general* mode of purely nonfunctional characterization, the vernacular of *pure occupancy.* There may be "pure occupants" or prime matter, ultimate unrealized realizers, even *necessarily* fundamental particles—presumably there are, despite my own tendency to think that it is functions all the way down—but further descent is always *epistemically* possible for us, and so we have no ordinary word for pure occupancy. "Role"/"occupant" remains a level-relative distinction; all we can mean by "pure occupant" is *stuff at a level L that realizes entities of level L + 1 but is not in fact realized at any lower level.*

Everything I have said so far may seem dull and obvious. I hope it does. I am trying to call attention to what I consider a home truth about the structure of the physical world, because I think neglect of this truth, inattention to the hierarchical nature of Nature, has led to significant errors and confusions about consciousness and "qualia." In the next chapter I shall review some of these and try to correct them.