CHRISTOPHER CHERNIAK

PHILOSOPHY AND COMPUTATIONAL NEUROANATOMY*

(Received 15 September 1993)

Cognitive science is sometimes extensionally identified, Venn diagram-style, as an intersection of allied fields of philosophy, linguistics, experimental psychology, and computer science. In recent years, neuroscience has been added to the cluster, and new labels – "cognitive neuroscience" or "mind-brain science". The discussion here focusses on the layering at the lens of intersection of mind-brain sciences, particularly on some of the philosophy and methodology of neuroanatomy. The objective is to argue for an obvious but recently relatively underexplored *via media* view of the interrelation between cognitive (i.e., intentional) and neural explanatory levels of mind-brain science, where each is neither irrelevant nor reducible to the other. I will advance the argument in terms of a case study, some neuroanatomical results of mine regarding applicability of combinatorial network optimization theory to brain structure. The experimental studies deal with neural component placement optimization: when anatomical positioning is treated like a microchip layout wire-minimization problem, a "best of all possible brains" hypothesis predicts actual placement of brains, their ganglia, and even their nerve cells.

1. PHILOSOPHY AND SCIENCE

We begin with the meta-philosophical issue of the interrelations between philosophy and the special sciences. Two simplest views mark the extremes: (1) Neither is at all relevant to the other. On the one hand, ordinary language philosophy of the 1950's epitomized this autonomist position, as on the other hand have most scientific fields in Anglo-America until recently. (An example from late in the ordinary language

Philosophical Studies **73**: 89–107, 1994.

era is Strawson's *Individuals*: "Metaphysics has a long and distinguished history, and it is consequently unlikely that there are any new truths to be discovered in descriptive metaphysics," nor, presumably, inadequacies such as antinomies; it is consequently not surprising to find no mention of the sciences in Strawson's "essay in descriptive metaphysics" as it purely describes, without recommendations for revision, the "massive central core of human thinking" which "changes not at all."[1]) (2) Philosophy just *is* science pursued by other means. Quine and later naturalizers of epistemology and metaphysics are current proponents of this reductionist position. (Quine: Philosophy "... is not to be distinguished in essential points of purpose and method from good and bad science." "Ontological questions ... are on a par with questions of natural science." "Epistemology, or something like it, simply falls into place as a chapter of psychology and hence of natural science."[2]).

The starting point here is that these extremes do not exhaust the options toward which one must gravitate. A third type of option to examine is that the complex relations between the two domains in a sense fall between these extremes: Philosophy cannot long be pursued as pure technique in isolation from contributions of the contemporary scientific worldview, any more than in isolation from the rest of our intellectual culture (e.g., its aesthetic enterprises); no field is an island, neither philosophy nor the sciences can be sustained without the other. Nonetheless, contrary to the naturalization paradigm, some problems and methods retain a distinctively philosophical character that can only be Procrusteanly reduced to those of the deductive and empirical sciences. A two-way street seems to run between philosophy and, for example, neuroscience. In the top-down direction, philosophy does not so much strictly entail predictions, but instead shifts and directs attention to particular internal scientific questions; we are thereby lead to look for the rabbit in the duck-picture. (Of course, this type of view has an affinity with Kant's account of the "regulative" role of the Ideas of Reason (as distinguished from the Concepts of the Understanding) in empirical science, as presented at the end of the *Prolegomena*.[3])

For, theory-blind data gathering can be a blunt instrument indeed. In the biological sciences one sometimes encounters a meta-methodology, more acquired reflex than self-consciously articulated position, that might be dubbed "Street Positivism," a rough, tacit empiricism manifested as profound suspicion of theory in general, whether philosophical or scientific. It is like a flip side of the anti-theory predilection of ordinary language philosophy of a later-Wittgensteinian "pictures mislead" bent.[4] In neuroanatomy, a caracature of the position might be, "To describe is to betray, to generalize is to murder." Each thing is what it is, and nothing else; there is only baffling diversity of structures, no commonalities or central tendencies. And, in fact, the very etymologies of neuroanatomical terms suggest a near-phantasmagoric, uncomprehending, free-associating subjectivism. For example, one finds among frequently used expressions, largely translating from Latin: seahorses, snails, shells, worms; almonds, olives, lentils; breasts, buttocks, teeth, tails, knees, horns; spiderwebs, nets, tufts; parasols, girdles, belts, ribbons, buttons, spurs; chandeliers, cushions, baskets, cups, funnels; chambers, roofs, gables, tents, bridges; stars, suns; hillocks, pyramids, wedges; fires, mosses, glue. And also a substantia innominata.[5]

Muscular and skeletal anatomy sometimes also get vivid, but not consistently so; I know of nothing comparable in the terminology of other fields – botany or geology, for instance. The fanciful vocabulary of neuroanatomy involuntarily summons an image that would put Salvador Dali to shame, a buzzing, blooming jumble of one damn thing following another. Words fail us in describing brain structures, we ransack the universe for comparisons. From the investigator's perspective, it indeed suggests (not entirely justly) an almost dream-like, irregular domain, lacking perceivable order. This is not a nomenclature to inspire confidence that there is a firm grip yet on the natural kinds of the domain, a grasp of the underlying causal order.

It is therefore not surprising to sense, behind a procession of announced revolutions, booms, and breakthroughs through the history of neuroscience, the eternally nagging doubt, Do we today really understand how the brain works much better than chimpanzees taking apart a radio? Or, are we essentially clueless, still in the dark about some of the inventory of basic mechanisms? – For example, as Descartes in at least some respects had to be before Galvani's start upon uncovering the electrochemical character of nerve transmission (or action-potential electrophysiologists before discovery of graded potentials, or reticularist neuroanatomists before observation of the synapse, and so on). One downside opinion is that, given the unparalleled complexity of the human brain – it is the most complex physical structure we know of in the universe – anatomy is an effectively endless task, never completed but only passed on to the next scientific generation. The more optimistic starting point here identifies an external role for a philosophical framework in driving some of the research agenda of neuroscience. Some of the abstractive power of philosophical concepts is required to cope with the crushing complexity of brain anatomy.

In particular, I have been exploring how a bounded-resource philosophical framework I originally worked out for more realistic models of the rational agent ("minimal rationality"[6]) might be extended to a set of abstract constraints on neuroscientific models and, further, to generate a positive research program in computational neuroanatomy. The link between the philosophy and the science has turned out to be via some of "the formalisms of scarcity," combinatorial network optimization theory from computer science. The perspective that emerges is a kind of low-key Pythagoreanism:[7] there is order in the universe, in particular, rather simple mathematical form in Nature, even in the neural jungle. For, the brain is not magic meat. Even if Descartes were right that mind floats outside the domain of scientific explanation, we certainly would not want to go on to infer that the brain also does – any more than we would want to infer from a Cartesian thesis that mind is non-spatial that the brain also is.[8]

2. BOUNDED-RESOURCE MODELS

The philosophical framework: One of the most fundamental laws of psychology is that human agents are finite objects. Human minds confront a finitary predicament – there is some fixed upper limit upon available cognitive resources. A contrasting ideal agent view is well-expressed in rational theology, for example Aquinas in the *Summa Theologica* on God's omniscience:

> ...God sees all things together, and not successively ...whosoever proceeds from principles to conclusions does not consider both at once... to advance thus is to proceed from the known to the unknown.[9]

And God's knowledge is not incomplete or imperfect. Aquinas' account of God's IQ bears an uncomfortable resemblance to a central element of models of the agent in standard epistemic logic, in particular, the usual logical competence axiom that the agent's belief set is deductively closed: $(\mathrm{B}p \;\&\; (p \rightarrow q)) \Rightarrow \mathrm{B}q$.[10] Again, Aquinas:

... God knows contingent things not successively, as they are in their own being, as we do, but simultaneously ... all things that are in time are present to God from eternity ...[11]

Corresponding to such a picture of God understanding things *sub species aeternitatis*, without time-consuming inference, one can recognize a profound lack of realism regarding the human condition in conventional rationality idealizations that philosophy inherits from microeconomic, game, and decision theory. The ideal rationality models entail the triviality of large portions of the deductive sciences (among others), and thereby deny the fundamental reality that our inquiries in fact have a *history*, where a research community discovers one thing and then uses that to go on to discover other things. If in the long run we will all be dead, an ideal agent model that is a type of perpetual-motion machine seems interestingly inappropriate. No wonder a folk psychology pictured with such an ideal rationality framework at its core would provoke crypto-instrumentalist/anti-realist belittlement, or outright eliminativist rejection.

Laplace's vivid image of a perfect intelligence making predictions in classical mechanics updates the traditional conception of an omniscient God:

An intelligence knowing all the forces acting in nature at a given instant, as well as the momentary positions of all things in the universe, would be able to comprehend in a single formula the motions of the largest bodies as well as of the lightest atoms in the world, provided that its intellect were sufficiently powerful to subject all data to analysis; to it nothing would be uncertain, the future as well as the past would be present to its eyes.[12]

Again, one can note that the human span of apprehension, unlike that of this ideal Predictor, is limited.[13] For human beings there *is* an *ignorabimus* – here, things we do not, will not, and cannot know, simply because of "practical" limits on our computational resources.[14] A minimal rationality account fits naturally with the contemporary "limits to growth" *Zeitgeist*: just as on a macrocosmic scale it pays to recall that Nature's resources are not in fact unlimited, so also on a microcosmic scale for the individual agent's cognitive resources.

3. SAVE WIRE

Within a bounded-resource framework, the step from the abstract cognitive level of explanation down to the neural hardware level can be pictured as taking literally the slogan, "We do not have God's brain." My first observations were at the computational level of explanation in computer science, for connectionist models of massively parallel and interconnected computation[15] that were intended to be more neurally realistic than conventional von Neumann computational architecture; as for higher-level rationality models in philosophy, these computational models still tended drastically to overestimate available resources – here, actual connectivity in the brain. At least initial connectionist models often tacitly assumed neural connections were virtually infinitely thin wires, with effectively instantaneous impulse-transmission. In assembling the quantitative neuroanatomy necessary for evaluating neural feasibility of connectionist models, it then became evident that a weaker but still discernible trend toward overestimation of resources pervaded even the most concrete "hardware" level of neuroanatomy.[16] "Impossibility engines" of this type turned up at all levels of mind-brain science.

A resource-realistic philosophical critique of mind-brain science turned attention from *in abstracto* boxology to actual, physical neuron connections as a critically constrained neurocomputational resource. The working hypothesis thereby emerged that because connections in the brain, particularly long-range ones, are a stringently limited resource both in volume and in signal-propagation times, minimizing costs of required connections strongly drives nervous system anatomy. Combinatorial network optimization theory has developed formalisms for expressing and solving problems of "saving wire." And so a positive research program emerged from the methodological critique: If actual brain connections are in severely short supply, is their anatomy correspondingly optimized?

As mentioned, in the face of overwhelming neural intricacy, neuroanatomy over its hundred-year modern history has traditionally tended toward "descriptive geography" of the nervous system, that is, *ad hoc* characterization of individual structures and relatively low-level empirical generalization. The idea of the present research was that the top-down power of concepts from computation theory could aid in coping with the unmatched complexity of the brain. Network optimization theory might provide a source for a "generative grammar" of the nervous system, some general principles that compactly characterize aspects of neuroanatomy. An hypothesis that network optimization virtually unidimensionally drives some aspects of neuroanatomical connectivity – that it dominates the many other plausible orthogonal dimensions of optimization of the nervous system – is relatively simple, if not itself *prima facie* very plausible. However, this line of inquiry has yielded some evidence that, in particular, component placement optimization and local Steiner tree are in fact organizing principles of nervous system anatomy.

4. NETWORK OPTIMIZATION

Combinatorial network optimization theory came of age about twenty years ago with the emergence of the theory of NP-completeness.[17] The key formal concept of a computational problem being NP-complete ("non-deterministic polynomial-time complete") need not be defined here; it is strongly conjectured to be linked with a problem being intrinsically computationally intractable, that is, not generally solvable without exhaustive search of all possible solutions. Because the number of possibilities combinatorially explodes as the size of a problem-instance grows, such brute-force searches are extremely computationally costly. Many of the most important real-world network optimization problems (e.g., most notably, Travelling Salesman) have been proven to be NP-complete or worse in computational complexity. Component placement optimization and Steiner tree, problems that have been the focus of the present research, are of this type, having been proven to be "NP-hard," that is, at least as difficult as NP-complete.

We concentrate here on the example of component placement optimization. Component placement optimization has received the most attention in computer science recently in connection with design of very large scale integrated (VLSI) microcircuits.[18] The problem can be defined as: Given the interconnections among a set of components, find the spatial layout – the physical arrangement – of the components that minimizes total connection costs. The simplest cost-measure is length of connections (often represented as the sum of squares of the lengths); usually the possible positions for components are restricted to a matrix of "legal slots." As a simple example, Figures 1a and 1b diagram two of six possible configurations of components 1, 2, and 3 in slots A, B, and C; for the connections among the components, the Figure 1b placement requires the least total connection length, Figure 1a the most.

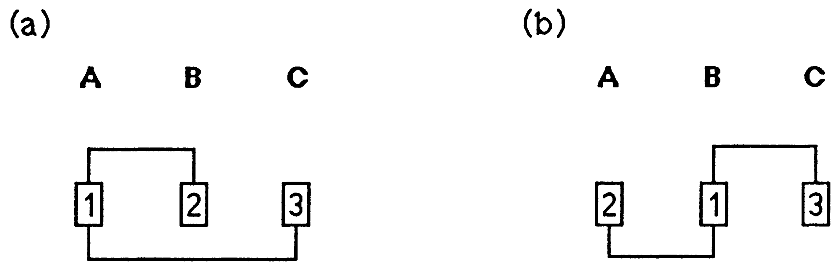


Figure 1. A three-component placement optimization problem. Two alternative placements of elements 1, 2, and 3 in fixed positions A, B, and C. For the given interconnections, placement (b) requires less total connection length than placement (a).

Computation costs for exact solution of component placement optimization problems are of a magnitude not encountered in most scientific computing. For n components, the number of alternative possible placements is $n!$ (Size of the search space is unaffected by whether permissible component positions are located in 3, 2 or 1 dimensions.) For instance, solving a mere 20-component problem can require dramatic resources, since $n! = 2.4 \times 10^{18}$ layouts, more than the total number of seconds in the 20 billion year history of the Universe since the Big

Bang. Consequently, the most distinctive experimental technique of the studies here involved very large-scale computer searches. Quick and dirty "heuristic" procedures that yield approximately optimal solutions can be much more feasible, but their performance (e.g., how close to optimality are they likely to come) is not well understood.[19]

5. NEURAL COMPONENT PLACEMENT

We studied the "brain as ultimate VLSI chip" hypothesis of component placement optimization in the nervous system at multiple hierarchical levels, from gross to microscopic anatomy.[20]

(a) Explaining "why the brain is in the head." At the highest level, positioning of the entire brain in the body constitutes a one-component placement problem (see Figure 2). The simplest connection cost-measure is just total length of individual fibers in all sensory and motor tracts to and from the brain. Physical siting of all sensors and effectors is treated as a given, fixed "edge-constraint." A portent of the "scale-blind," deeply non-quantitative character of recent neuroanatomy is that the low-tech information required for solution of this problem – in particular, for obtaining numbers of fibers in all nerve tracts – has been published only for human and nematode worm nervous systems.[21] For both creatures, number of nerve fibers to and from locations anterior to the brain (or predominant concentration of the nervous system) exceeds number of fibers to and from locations posterior to the brain. It immediately follows that the "wire-minimizing" placement of the brain will be as far forward as possible. And actual positioning of human and nematode brains is in fact consistent with this wire-minimization prediction. (In general, inspection of gross anatomy drawings suggests that whenever anterior connections exceed posterior ones, as in the case of all vertebrates and most invertebrates, the brain is correspondingly placed as far forward on the body axis as possible.)

(b) Layout of functional areas of cerebral cortex. The placement optimization hypothesis for these fifty components is that they are positioned on the 2-dimensional cortical sheet to minimize total length of their interconnections. As mentioned above, search of all possible

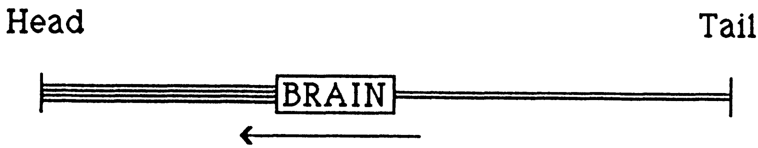


Figure 2. A biological one-component placement optimization problem. The brain has more sensor/motor connections to head than to tail. Hence, to minimize total length of peripheral nerves, brain should be positioned as far forward as possible.

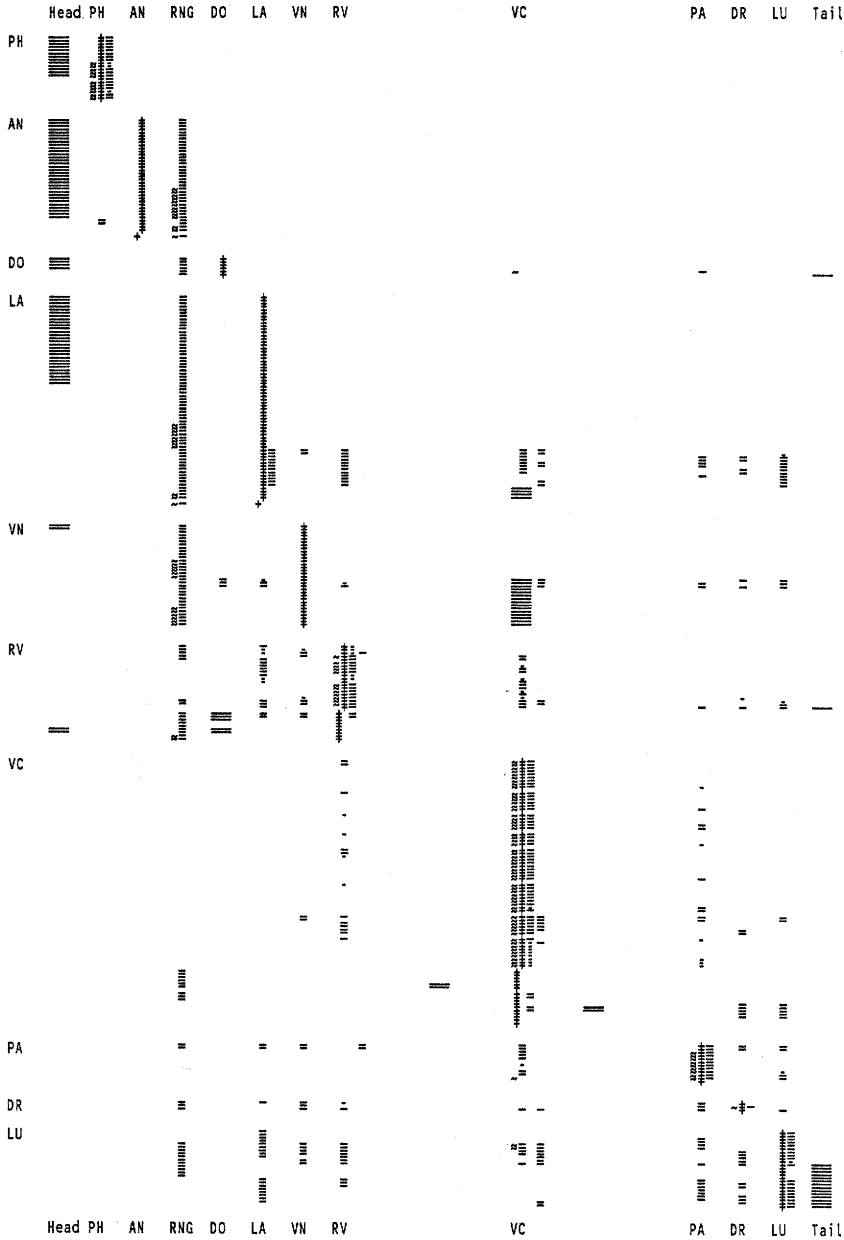
alternative layouts of even 20 components to verify optimization would require resources of cosmic scale. However, if cortical components are placed to minimize interconnection lengths, one would expect to find quite tractable statistical confirmation of the "adjacency rule": If components *a* and *b* are interconnected, then they are positioned contiguous to each other, other things equal. And in fact, our connectivity and contiguity databases compiled for macaque monkey, cat, and rat show that each cortical layout strongly departs from random placement in favor of the adjacency rule ($p < 0.0001$). The rule is a powerful predictor of the anatomy, a kind of "plate tectonics of the cortex."

(c) Layout of ganglia of *C. elegans*. For only one animal is there now approximately complete neuroanatomy, down to synapse level – the millimeter-long roundworm *Caenorhabditis elegans*. During two decades, the Cambridge University *C. elegans* group has published about a thousand pages of anatomy drawings on the 302 neurons of the worm nervous system,[22] a measure of the daunting intricacy of even so simple a brain. From these diagrams, we compiled a hundred-page database giving for each neuron its location and all known connections; from the database, a ten-page connectivity matrix was then computed. Figure 3 is a one-page representation of the matrix, giving a synoptic view of all connections of all neurons ("PH," "AN," ... "LU" designate the ganglion-level components). This appears to be the first complete depiction of a nervous system, at neuron-level detail, in a single image. In this way, a philosophically-driven Gestalt switch leads one simply to put the trees together and look at the forest. One immediately perceives a clustering of connections along the diagonal from top left to lower right, which in fact signals that the above adjacency rule is again strongly confirmed – here, for positioning of the worm's ganglia.

However, a more striking finding can be obtained. The ganglion-level optimization problem has 11 movable components, with $11! = 39{,}916{,}800$ possible alternative orderings. A dozen microcomputers running in parallel for a week were able to search exhaustively all of these placements, yielding the result that the actual is the ideal, or optimal: The worm's actual ganglion layout in fact requires less total length of connections than any of the other millions of possible layouts. In terms of methods, searches of this scale are unprecedented in computational anatomy.[23] The unfamiliar scale of such a one in a million search problem can be difficult to absorb; if each layout were described in a single line, just listing them all would require about one "mega-page" – a million pages.

Or again, suppose instead that 2,000 alternate layouts had turned out to require less "wire" than the actual layout. If each of the $\sim 40,000,000$ possible layouts surpassed by the actual one represented an increment of a millimeter in a "possible-worm race," then the actual layout would still have covered all but the last *two meters* of the total Darwinian racetrack of 40,000,000 mm – i.e., of 40 km. One natural interpretation of such a two-thousandth-place finish would be to round off, that is, to consider the possibility that, after beating the rest of the 40,000,000 alternative layouts, failure of the actual layout to beat the last 2,000 was merely apparent, e.g., plausibly suspected to arise from some type of small-scale "noise" or measurement-error. In effect, this search constitutes a kind of crude simulation of the maximal possible history on Earth of the evolution of the worm nervous system.

(d) Individual neuron placement. Finally, there is also evidence that placement optimization is so sensitive that it fine-tunes even the positioning of individual neuron cell bodies in *C. elegans*. (i) The following modification of the above "wire-saving" adjacency rule can be tested: If neurons *a* and *b* are interconnected, then they are placed near each other, in particular, clustered in the same ganglion, other things equal. Again, statistical confirmation is strong. (ii) Even positioning of neuron cell bodies within ganglia conforms to a component placement optimization prediction. Namely, there is a highly significant trend


Head PH AN RNG DO LA VN RV VC PA DR LU Tail
PH
AN
DO
LA
VN
RV
VC
PA
DR
LU
Head PH AN RNG DO LA VN RV VC PA DR LU Tail

for somata making exclusively anterior extra-ganglionic connections to be located in the front half of a ganglion, while somata with external connections only to sites posterior to the ganglion tend to be placed in the rear half of the ganglion. Even at the individual cell level, layout optimization powerfully predicts anatomy.

Thus, from first principles, from the most abstract reaches of philosophy, via the formalisms of scarcity, one obtains predictions at the most concrete hardware level of neuroanatomy. An Occam's Razor of the nervous system, the simple *logos* "Save wire" invokes a significant portion of the vast neuro-wiring diagram. For instance, if the finding is confirmed that the actual worm ganglion layout is the unique first-place winner requiring minimum total length of connections, it constitutes one of the predictive success stories of recent quantitative anatomy. Using another network optimization concept, Steiner tree, we also have obtained some similar results for local optimization of the dendritic and axonic arbors of a wide variety of nerve cells.[24]

6. PEACEFUL COEXISTENCE

We began by looking at the relation between philosophy and science. Correspondingly, we now turn to the relation within mind-brain science

---

←

Figure 3. Total ganglion-level connectivity map for *Caenorhabditis elegans* nervous system. Each partialoly superimposed micro-line represents one of the 302 neurons: +, soma; –, asymmetrical (chemical) synapse; -, symmetrical (gap) synapse; ~, muscle connection; —, sensor. (Non-ganglionic somata appear below and one space to left of somata of nearest ganglion; connections to non-ganglionic neurons appear in the column one space to right of column for connections to neurons of the nearest ganglion.) "PH," "AN," etc. are codes for the ganglia. Compiled from published anatomy, this appears to be the first complete depiction in a single image of a nervous system at the individual neuron level. An immediately evident trend of distribution of connections around the diagonal from upper left corner to lower right corner indicates that the layout conforms to an "adjacency rule" that tends to minimize total connection costs. Horizontal scaling, approximately 100x. A hand magnifier and a transparent straightedge will reveal further detail. (Adapted from Cherniak, "Component Placement Optimization in the Brain"; see Note 20.)

between explanations at the abstract cognitive/intentional and computational levels and at the neuroscientific hardware level. Again, tendencies toward two familiar exclusivist extremes appear: History threatens to repeat itself in counterparts, respectively, to a Cartesian "autonomy of the mental" view and to an "empty organism" *redivivus* reductionism.

(a) The autonomist, cognitivist position. There is a science of internal mental representations and processes. But information about hardware is *irrelevant* to cognitive/computational explanations.[25] (One practitioner recently remarked informally that the brain is a junk yard; that neuroscience is as germane to cognitive psychology as ornithology is to aeronautical engineering; that so far as psychology is concerned, one could just as well be a dualist and a creationist; that cognitive psychology would not be worse if all the neuroscience books were burned.) One etiology of the view begins with the Chomskian thesis of the psychological, and neural, non-reality of properly abstract language-acquisition "competence" models.[26] Another origin may stem from the functionalist credo: Cognitive psychology is computationalist; the software level of explanation is hardware-independent; "therefore," hardware-level information, including neuroscience, is irrelevant for cognitive psychology.[27]

(b) The quasi-reductionist, eliminativist position. In the mind-brain domain, neuroscience is the only real science. It must eventually entirely supplant cognitive explanation (e.g., via rationality concepts) as informal, pre-scientific folk chat or slang.[28] In this sense, eliminativism tends *de facto*, if not deliberately, toward reviving a type of "empty-organism" position. The connectionist revival has renewed the force of this view. (Perhaps such accounts also gain force in part from the longstanding tendency throughout the human sciences toward a conception of the natural sciences as the only model for more problematic fields.)

And again, an obvious *via media* compatibilist position emerges: mutual coexistence of the two explanatory levels, cooperation instead of zero-sum. To begin with, one can promote a third option at least to the extent of disputing "the only games in town" arguments for either of the two traditional options. Going further, one can draw upon the computational neuroanatomy example to illustrate some modes of interrelation of the levels: As described above, an idea at one level can drive the research program at the other, even by merely suggesting a different way of looking within the other level. Also, the different levels can reinforce each other by contributing converging, consilient support to one another; e.g., bounded-resource anatomical models can lend the plausibility of a measure of systematicity to the original bounded-resource rationality models.

Another scheme of weak interrelation of the levels can be pictured in terms of the total multidimensional space containing possible explanations of mind-brain at every level. A resource-constraint at a given level is like a plane passed through this explanation-space, delimiting the set of feasible explanations on one side. So, one modest type of progress in mind-brain science can consist just of narrowing the space of possible explanations for a given degree of idealization by successively locating more such constraint-planes in the explanation-space. For instance, it is relevant to identifying the degree of realism of a connectionist model that it requires a brain the size of a bathtub; or again, it is a useful consistency test in either direction to estimate crudely the size of a normal human cognitive system (e.g., the belief set) and to find that there are in fact easily enough synaptic resources for its neural representation.[19]

A final example of how higher-level framework can impinge upon lower-level theory concerns the issue of intelligible form in nature. Suppose that, in the pattern of synaptic connections of the nematode nervous system (see Figure 3), one could discern the Lord's Prayer encoded in some conventional binary-based notation. This would be an odd but securely uninteresting coincidence (except perhaps for "Ripley's Believe it or Not"); it seems epiphenomenal, like the resemblance of horses and seahorses. It is scientifically unimportant because we cannot plausibly envisage an account of how such a pattern would regularly arise from the basic causal order of the Universe. Similarly, finding in the structure of worm neural connections the one in a million minimal wire-cost layout is also a remarkable "coincidence." However, the latter pattern can bear significance when interpreted in terms of a bounded-resource framework, because we can then propose an explanation of why the anatomy assumes this structure. However tentatively, we thereby go beyond "and-then-a-miracle-happens" numerology.[29]

The computational neuroanatomy example has illustrated some particularly long-range modes of interrelation between higher and lower levels of mind-brain science. To conclude by just broaching some further interlevel queries: Finding very good network optimization of neuroanatomy immediately raises questions concerning the mechanisms by which the optimization is actually accomplished. "Nature, the blind watchmaker" executing simple brute-force exhaustive search for a mere 50 component problem would require, even at quite unrealistically high speed and parallelism, more than the age of the Universe.[30] Exact solutions, as opposed to "quick but dirty" approximate/probabilistic ones, would computationally hogtie the entire cosmos. We thereby turn from consideration of bounded-resource models of the individual human agent to constructing them for the Universe itself, and back to the familiar issue in philosophy of evolutionary biology of optimization in Nature – in this case, whether *Deus sive Natura* can build the best of all possible brains without supernatural or magical powers. Widespread phenomena of refinement of connectivity optimization also raise a prior, if inchoate, question, Why should saving wire be so distinctively important, despite the many other crucial desiderata in engineering a nervous system? Perhaps the singularly special value of the bounded resource of connections gives yet another hint about how brains must function.

NOTES

* Invited paper, Pacific Division Meetings, American Philosophical Association, San Francisco, March 1993. The paper benefitted from the valuable contributions of Dan Lloyd and Lisa Lloyd for this session; I am also grateful to session chair Carol Cleland and to Karen Bell for ideas and help. The neuroanatomical studies reported were supported in part by the University of Maryland Institute for Advanced Computer Studies, and by National Institute of Mental Health grant MH49867.

[1] *Individuals* (New York, Doubleday: 1959), pp. xiii–xv. On the conservative presumption that there is no room for improvement in core ordinary-language concepts, hence again presumably no possible philosophical role for scientific discoveries, see also G. Warnock, *English Philosophy Since 1900* (London, Oxford Univ. Press: 1958), e.g., "... it is at the very least unlikely that [language] should contain either much more, or much less, than [its] purposes require. ... It is at the same time very unlikely that any invented ... terminology will be an improvement," p. 150. Once more, in Wittgenstein's words, philosophy "leaves everything as it is."

[2] Respectively: *Word and Object* (Cambridge, Mass., MIT Press: 1960), pp. 3–4; "Two Dogmas of Empiricism," in *From a Logical Point of View* (Cambridge, Mass., Harvard Univ. Press: 1961), p. 45; "Epistemology Naturalized," in *Ontological Relativity* (New York, Columbia Univ. Press: 1969), p. 82. A salient instance of latter-day naturalizing along Quine lines is the explicit psychologism of G. Harman, *Thought* (Princeton, N.J., Princeton Univ. Press: 1973), e.g., pp. 15–19.

[3] I. Kant, *Prolegomena to Any Future Metaphysics*, L. Beck, tr. (New York, Bobbs-Merrill: 1950), secs. 40–44, 56–57. I develop this "distinct but interconnected" view of the philosophy-science relation in ch. 6 of my *Minimal Rationality* (Cambridge, Mass, MIT Press: 1986), and in my "The Division of Intellectual Labor," in preparation.

[4] See, for instance, the first sections of *Philosophical Investigations* (New York, Macmillan: 1958).

[5] Respectively: hippocampus, cochlea, putamen, vermis; amygdala, olive, lentiform nucleus; mamillary body, nates, dentate nucleus, caudate nucleus, geniculate bodies, ventral horn; arachnoid, reticular formation, flocculus; cingulum, limbic formation, lemniscus (and ribbon synapse), synaptic boutons, calcarine fissure; chandelier cell, pulvinar, basket cell, cupula, infundibulum; thalamus, tectum, fastigium, tentorium, pons; stellate neuron (and astrocyte), solar plexus; colliculus, pyramidal neuron, cuneate nucleus; pyriform cortex, mossy fiber, glial cell.

[6] See *Minimal Rationality*, op. cit.

[7] E.g., in the tradition of D. Thompson, *On Growth and Form* (New York, Cambridge Univ. Press: 1960).

[8] On some signs of the latter tendency in neuroscience, see my "The Bounded Brain: Toward Quantitative Neuroanatomy," *Journal of Cognitive Neuroscience* 2 (1990): 58–68.

[9] T. Aquinas, *Summa Theologica*, I, Q. 14, Art. 7; in A. Pegis, ed., *Basic Writings of Saint Thomas Aquinas* (New York, Random House: 1945), vol. I.

[10] E.g., as it appeared in J. Hintikka, *Knowledge and Belief* (Ithaca, N.Y., Cornell Univ. Press: 1962).

[11] Op. cit., I, Q. 14, Art. 13.

[12] P. Laplace, *Philosophical Essay on Probabilities* (New York, Dover: 1951), p. 4.

[13] See ch. 3 of *Minimal Rationality*, op. cit.

[14] (Assertions of human-deity equivalence sometimes emerge explicitly, as when Hamlet declaims, "What a piece of work is a man! ... how infinite in faculty! ... in apprehension how like a god!" (Act II, Sc. 2))

[15] Cf., e.g., D. Rumelhart and J. McClelland, eds., *Parallel Distributed Processing*, vols. I & II (Cambridge, Mass., MIT Press: 1986); and *Cognitive Science* 9 (1985).

[16] See "The Bounded Brain: Toward Quantitative Neuroanatomy," op. cit.

[17] M. Garey and D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (San Francisco, W. H. Freeman: 1979). The best nontechnical introduction to the field remains H. Lewis and C. Papadimitriou, "The Efficiency of Algorithms," *Scientific American* 238 (1978): 96–109; with L. Stockmeyer and

A. Chandra, "Intrinsically Difficult Problems," *Scientific American* 240 (1979): 140–159.

[18] For a review, see E. Kuh and T. Ohtsuki, "Recent Advances in VLSI Layout," *Proceedings of the IEEE* 78 (1990): 237–263.

[19] See my "Undebuggability and Cognitive Science," *Communications of the Association for Computing Machinery* 31 (1988): 402–412.

[20] For technical description of the following results, see C. Cherniak, "Component Placement Optimization in the Brain," *University of Maryland Institute for Advanced Computer Studies Technical Report* (1991) No. 91–98; and "Component Placement Optimization in the Brain," *Journal of Neuroscience* (1994), in press. (Historical note: A quite similar result to the adjacency rule of (b) below for monkey cortex can be found in M. Young, "Objective Analysis of the Topological Organization of the Primate Cortical Visual System," *Nature* 358 (1992): 152–155. However, in a November 19, 1991 letter, *Nature* had evaluated a report of the results summarized here as not of sufficiently immediate interest to a general readership; Young's piece appeared in *Nature* with a December 6, 1991 "received" date.)

[21] See, respectively: S. Blinkov and I. Glezer, *The Human Brain in Figures and Tables: A Quantitative Handbook* (New York, Plenum: 1968); and W. Wood, ed., *The Nematode* Caenorhabditis Elegans (Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory: 1988).

[22] The canonical summing up of the project is Wood, op. cit.

[23] Indeed, my very capable programmers at first worried the machines might physically melt down during such long runs. – The mathematicians Gregory and David Chudnovsky, emigres from the Soviet Union, reportedly have been computing the value of pi to new limits using much huger kluges of ordinary microcomputers as their supercomputer (Richard Preston, "The Mountains of Pi," *The New Yorker* (March 2, 1992): 36–67).

[24] C. Cherniak, "Local Optimization of Neuron Arbors," *Biological Cybernetics* 66 (1992): 503–510; and "Global Optimization of Neuron Arbors," in preparation.

[25] E.g., Z. Pylyshin, "Computation and Cognition: Issues in the Foundation of Cognitive Science," *Behavioral and Brain Sciences* 3 (1980): 111–132.

[26] N. Chomsky, *Aspects of the Theory of Syntax* (Cambridge, Mass., MIT Press: 1965), ch. 1.

[27] Cf. "The Bounded Brain: Toward Quantitative Neuroanatomy," op. cit., p. 65.

[28] Some affinities with this type of view seem present in: Paul Churchland, *Matter and Consciousness* (Cambridge, Mass., MIT Press: 1984); Patricia Churchland, *Neurophilosophy* (Cambridge, Mass. MIT Press: 1986), part II; Patricia Churchland, "The Significance of Neuroscience for Philosophy," *Trends in Neurosciences* 11 (1988): 304–307.

[29] There is some parallel with the less extreme case of J. Balmer's empirical formula for deriving positions of lines of the atomic hydrogen spectrum and his non-model to account for the good fit of his equation, versus later explanations of the Balmer equa-

tion in terms of basic properties of the hydrogen atom. (See G. Holton and S. Brush, *Introduction to Concepts and Theories in Physical Science*, 2nd ed. (Reading, Mass., Addison-Wesley: 1973), pp. 475–478.)

[30] See the last sections of "Component Placement Optimization in the Brain," op. cit., and "Local Optimization of Neuron Arbors," op. cit.

Committee on History and Philosophy of Science
Department of Philosophy
University of Maryland
College Park, MD 20742
USA