Logic, Methodology and Philosophy of Science: Proceedings of the Thirteenth International Congress. C Glymour, W Wei, & D Westerstahl, eds. (College Publications: London) 2009. [Beijing, 2007]

Minimal Rationality vs Optimized Brain-Wiring:[1]

Christopher Cherniak
University of Maryland
cherniak@umd.edu

ABSTRACT. One critique of standard rationality idealizations is that they require the agent to be an Ideal Logician. – For instance, if finitely represented, this deductive competence would violate Church's Undecidability Theorem. Such an apercu motivates bounded-resource models of the agent, where rationality instead falls between nothing and perfection. In turn, this resource-realistic perspective focuses attention on bounded-resource models down at the neural hardware level – on the limited neuroanatomical connectivity available. However, deployment of some brain-wiring turns out to be optimized to the limits of detectability (e.g., layout of interconnected *C. elegans* ganglia, cat cerebral cortex areas, etc.). Furthermore, simple physical processes – without e.g., DNA involvement – seem to suffice to generate some of this complex structure. Prima facie, finely minimized brain-wiring phenomena appear in tension with the minimal rationality models that drove their discovery. This clash pertains to the Minimalist program in linguistics – the question whether language processing resembles optimal neural hardware.

1 Preface

The emphasis here is on combinatorial network optimization – that is, minimization of connection costs among interconnected components in a system. The picture will be that such wiring minimization can be observed at various levels of nervous systems, invertebrate and vertebrate, from placement of the entire brain in the body down to the subcellular level of neuron arbor geometry. In some cases, the minimization appears either perfect, or as good as can be detected with current methods – a predictive success story. In addition, these instances of optimized neuroanatomy include candidates for some of the most complex biological structures known to be derivable "for free, directly from physics" – i.e., purely from simple physical energy minimization processes. Such a "Physics suffices" picture for some biological self-organization directs attention to innate structure via non-genomic mechanisms, an underlying leitmotif of the Minimalist program in Chomskian linguistics.[20]

[1]The author is with the Committee for Philosophy and the Sciences. Department of Philosophy, University of Maryland. College Park, MD 20742 USA.

Real-World Concept	Formal Concept	Problem-Examples
⇐[C-T thesis]⇒		
"Absolutely Impossible" (?)	Absolutely uncomputable	Arithmetic Full pred calc
"Practically Unfeasible" (?)	ND-exponential time [$t = E(i)$]	WS1S Presburger arith Monadic pred calc
	NP-time / NP-complete	Steiner tree Component placement Travelling salesman Propositional calc
		. . . . .
"Practically Feasible" (?)	Polynomial time [$t = P(i)$]	Min spanning tree Nim Multiplication

Table I: A computational intractability hierarchy. (Conceptual / schematic cartoon.) Standard rationality models require an Ideal Logician, who can perform absolutely uncomputable inferences, as well as NP-complete ones

2 Rationality

Standard conceptions of rationality derive from models of the rational agent in microeconomic, game, and decision theory. The underlying idealization is that the agent, given its belief-desire system, optimizes its choices. To achieve this perfection of appropriate decisions in turn would require vast inferential insight.

While such logical omniscience might properly characterize a Supreme Being, prima facie it seems to clash with the most basic law of human psychology: We are finite entities. Among the more extreme departures from reality, for such ideal agents, portions of the deductive sciences would be trivial. (E.g., the role of the discovery of the set-theoretic paradoxes in the development of logic over the past century then ceases to be intelligible.) For a computational theory of mind, the agent's deductive competence must be represented as a finite algorithm; this ideal agent would in fact have to violate Church's Undecidability Theorem for first-order logic.[3]

Furthermore, formally correct and complete inference procedures are typically computationally complex, with surprisingly small-scale problem instances sometimes requiring cosmically unfeasible time and memory resources. To an extent, this practical intractability parallels classical absolute unsolvability. (Tab.I, Fig.1)

Since most of us on good days are not computationally hogtied or rationality-paralyzed, we must use instead quick and dirty heuristics – i.e., trade off al-

Test of truth-functional consistency [SAT] of Belief Set
by truth-table method, by ideal computer **I**:

I checks each line of truth-table in 1 "supercycle":
the time a lightray takes to traverse dia of proton.
[Speed of light = 299,726 km/sec, proton = 10^{-13} cm dia.]
So, supercycle = 2.9×10^{-23} sec.

Maximum number cycles available in history of Universe:
2×10^{10} years x 365 days x 24 hrs x 60 min x 60 sec
x 2.9×10^{23} cycles < 2×10^{41} cycles total.

What is largest Belief Set **I** could test?
Number lines in truth table = 2^n,
where n = number of logically independent props.

2^{138} atoms require more than 3×10^{41} lines.

So: at 1 truth-table line checked / supercycle,
I cannot evaluate the truth-table for even 138
independent propositions during interval from
BigBang to present.

Figure 1: Cosmic cost of simple truth-functional consistency tests. SAT is in-principle computable, but NP-complete. – The exponential growth of the cost-function is not humanly intuitive.[4, p.403]

gorithm correctness and / or completeness for speedup, etc. In turn, is this deployment itself anything like finely optimized, rather than just moderately well-managed? That does not seem likely, nor advisable. One can get trapped in paralyzing regresses if limited resources are expended on perfecting decisions in turn about how to use the resources, etc., etc. (E.g., Ryle[25].)

One response to problems with the traditional idealizations is a via media strategy. Acknowledging that nothing could count as an agent or person that conformed to no rationality constraints, one pauses to ask: Must one leap to a conclusion that the agent has to be ideally rational? Is rationality all-or-nothing, or is there some golden mean ranging between unattainable perfect unity of mind and utter disintegration of personhood? The normative and empirical rationality models of Simon [26] are among the earliest of this less stringent type: The central principle is that, rather than optimizing or maximizing, the agent only "satisfices" its expected utility. Similarly for agent-constitutive rationality [2]: "No rationality, no agent."

Such a resource-realistic rationality framework possesses systematicity, in that it links together independent research programs in cognitive psychology and in computer science. The psychological studies focussed on people's strikingly ubiquitous use of reasoning procedures that are formally incorrect and/or incomplete heuristics (e.g., Kahneman et al.[23]). Another field that developed contemporaneously with, but separately from, non-ideal rationality accounts and experimental studies of reasoning heuristics is computational complexity theory in computer science[22]: As mentioned above, the basic insight is that formally correct and complete inference procedures appear to be intrinsically intractable. So, if formally adequate procedures entail practical paralysis, then the above heuristics may constitute the ultimate tradeoff of perfection for speed and usability.

A bounded-resource paradigm for rationality also of course meshes with the basic scientific worldview of humans' place in Nature. Like other creatures in the natural order, we are only finite.[1]

3 From Rationality to Neuroanatomy

The philosophical framework here extrapolates from the above rationality critique: Tacit unbounded-resource assumptions, that human cognitive resources are virtually without limit, pervade all levels of mind / brain science. (Fig.2) – Such "impossibility engines" turn up even at the most concrete hardware level, neuroanatomy. For example, some anatomical descriptions of cortex connectivity resources are not even quantitatively consistent.[5] It is as if neural connections were virtually infinitely thin wires, with effectively instantaneous signal transmission.

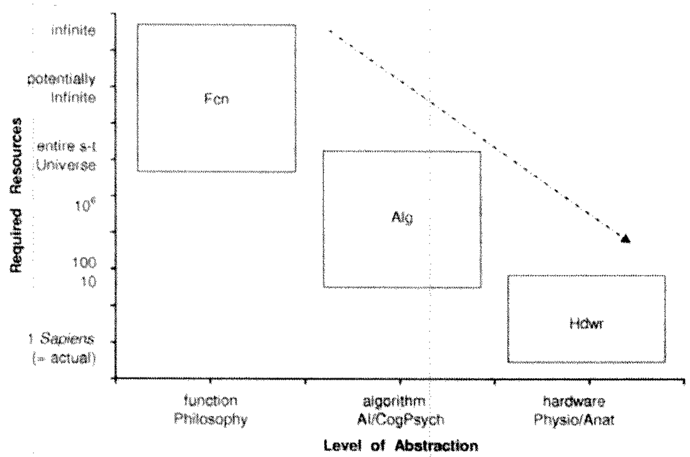

Figure 2: Required resources as a function of level of abstraction of model in cognitive neuroscience. (Conceptual cartoon.) "1 Sapiens"≈ resources of 1 *H. sapiens*. The less abstract the model, the less overestimation of resources presupposed; but, even at the most concrete level, overestimation persists.[6, p.239]

For instance, for connectionist models of massively parallel and interconnected computation in the brain, how much brain wiring in fact is available? And, how well is it employed; how well do neurons optimize even just local connectivity?

Long-range connections in the brain are a critically constrained resource, hence there seems strong selective pressure to optimize finely their deployment. The "formalism of scarcity" of interconnections is network optimization theory, which characterizes efficient use of limited connection resources. The field matured in the 1970's for microcircuit design, typically to minimize the total length of wire needed to make a given set of connections among components. When this simple "Save wire" idea is treated as a generative principle for nervous system organization, it turns out to have applicability: To an extent, "Instant brain structure – just add wire-minimization." The main caveat is that in general network optimization problems are easy to state, but enormously computationally costly to solve exactly. The ones reviewed here are "NP-hard", each conjectured to require computation time on the order of brute-force search of all possible solutions, hence often intractable. The discussion here focuses upon component placement optimization.

4 Component Placement Optimization

A key problem in microcircuit design is component placement optimization: Given a system of interconnected components, find the positioning of the components on a 2-d surface that minimizes total connection cost (e.g., wirelength). This concept seems to account for aspects of neuroanatomy at multiple hierarchical levels.

"Why the brain is in the head" is a 1-component placement problem. That is, given the positions of receptors and muscles, positioning the brain as far forward in the body axis as possible minimizes total nerve connection costs to and from the brain, because more sensory and motor connections go to the anterior than to the posterior of the body. This seems to hold for the vertebrate series (e.g., humans), and also for invertebrates with sufficient cephalization to possess a main nervous system concentration (e.g., nematodes).[8, 10]

Multiple-component problems generally require exponentially exploding costs for exact solutions; for an n-component system, all $n!$ alternative layouts must be searched. One neural wiring optimization result is for placement of the 11 ganglionic components of the nervous system of the roundworm *Caenorhabditis elegans*, with about 1,000 interconnections. (See Fig.3.) This nervous system is the first to be completely mapped, which enables fair approximation of wirelengths of connections. (See Fig.4.) When all ~ 40 million alternative possible ganglion layouts are generated, the actual layout turns out in fact to be the minimum wirelength one.[8] Some optimization mechanisms provide convergent support for this finding: A simple genetic algorithm, with wirecost as fitness-measure, will rapidly and robustly converge upon the actual optimal layout.[14] Also, a force-directed placement ("mesh of springs") algorithm, with each connection approximated as a microspring acting between components, attains the actual layout as a minimum-energy state, with relatively little trapping in local minima.[14] (See Fig.5.)

(There is also statistical evidence that this "brain as microchip" wire-minimization framework applies in the worm down to the level of clustering of individual neurons into ganglionic groups, and even to cell body positioning within ganglia to reduce connection costs.[8])

Finally, the wiring-minimization approach can be applied to placement of functional areas of mammalian cerebral cortex. Since wirelengths of intrinsic cortex connections are difficult to derive, an alternative approach is to examine instead a simpler topological measure of connection cost, conformance of a layout

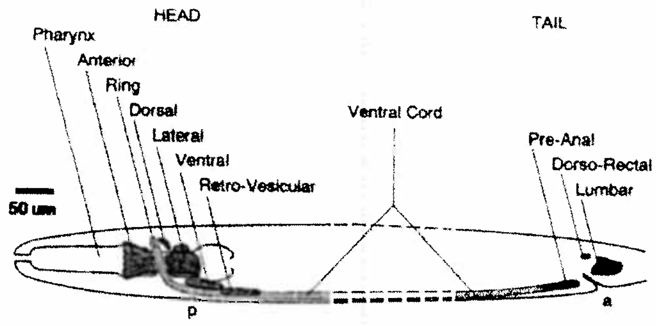

Figure 3: *C. elegans* ganglion components: their body locations and schematized shapes.

to a wire-saving heuristic Adjacency Rule: If components a and b are connected, then a and b are adjacent. Exhaustive search of all possible layouts is still required to identify the cheapest ones. A promising calibration of this approach is that the minimum wirecost actual layout of the nematode ganglia is among the layouts with fewest violations of this adjacency rule. For 17 core visual areas of macaque cortex, the actual layout of this subsystem ranks in the top 10^{-7} layouts best fitting this adjacency-costing; for 15 visual areas of cat cortex, the actual layout ranks in the top 10^{-6} of all layouts.[15]

In general, a Size Law seems to apply to cases like macaque, cat (and worm) with local-global tradeoffs: The larger proportion of a total system the evaluated subsystem is, the better its optimization. (We have observed this Size Law trend recently also for rat olfactory cortex and for rat amygdala.[16]) For the largest systems studied (visual, auditory, plus somatosensory areas of cat cortex), there is evidence of optimization approaching limits of current detectability by brute-force sampling techniques. (See Fig.6.) A similar Size Law pattern also appears to hold for Steiner subtree optimization of neuron arbor topologies. [13]

5 Optimization: Mechanisms and Functional Roles

The innateness hypothesis is typically expressed in the DNA era as a thesis that some cognitive structure is encoded in the genome. In contrast, an idea of "non-genomic nativism"[11] can be explored, that some biological structure is inborn, yet not genome-dependent; instead, it arises directly from simple physical processes. – Not only, then, is the organism's *tabula rasa* in fact not blank, it is "pre-formatted" by the natural order: a significant proportion of structural information is pre-inscribed via physical and mathematical law.

Noam Chomsky describes a strong Minimalist thesis, that "a principled account" of language is possible: "If that thesis were true, language would be something like a snowflake, taking the form it does by virtue of natural law, ..." [17–20] The snowflake reference recalls D'Arcy Wentworth Thompson's *On*

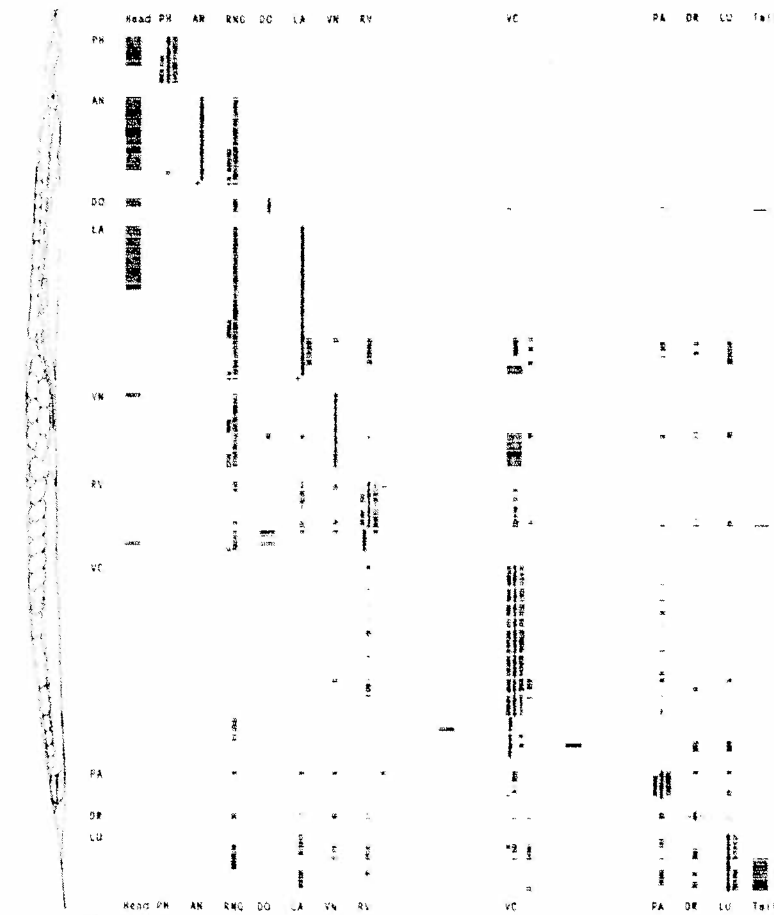

Figure 4: Complete ganglion-level connectivity map for *C. elegans* nervous system. (This may be the first depiction of approximately complete connectivity of a nervous system down to synapse level.) Each horizontal microline represents one of the 302 neurons. This actual ganglion layout requires the least total connection length of all 39,916,800 alternative orderings.[8]

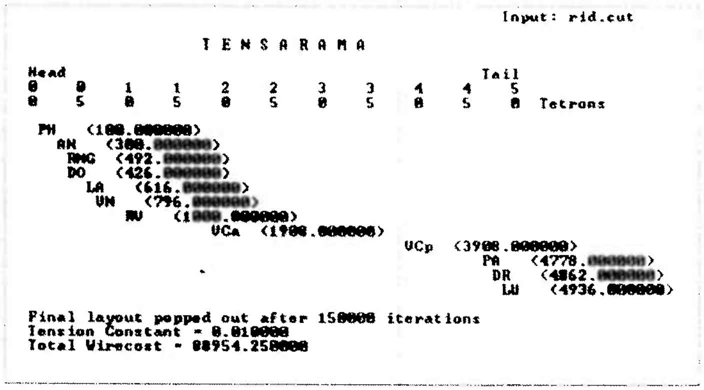

Figure 5: Tensarama, a force-directed placement algorithm for optimizing layout of *C. elegans* ganglia. This "mesh of springs" vector-mechanical energy-minimization simulation represents each of the worm's ~ 1,000 connections (not shown) acting upon the movable ganglia PH, AN, etc. The key feature of Tensarama performance for the actual worm connectivity matrix is its low susceptibility to local minima traps. [14] - Unlike Tensarama performance for small modifications of the actual connectivity matrix, and unlike such force-directed placement algorithms in general for circuit design. Here Tensarama is trapped in a slightly sub-optimal layout, by a "killer" connectivity matrix that differs from the actual matrix by only *one* less connection.

Growth and Form (1917/1961)[27], where the paradigmatic example of mathematical form in nature was the hexagonal packing array, of which snow crystals are an instance. However, even the thousand pages of the unabridged 1917 edition of Thompson's opus contained few neural examples. Similarly, Alan Turing's study[28] of biological morphogenesis via chemical diffusion processes began an inquiry that needs to continue. In effect, we have seen here how far this type of idea presently can be seen to extend for biological structure at the concrete hardware level of neuroanatomy. The key concept linking the physics and the anatomy is optimization of brain wiring.

The picture here has been of limited connections deployed very well – a predictive success story. The significance of ultra-fine neural optimization remains an open issue. Another question is where it does in fact occur, and how good it is. Tradeoffs of local optimality for better cost minimization of a total system are one way in which global optimization can be obscured.

The high levels of connection optimization in the nervous system seem unlike levels of optimization common elsewhere in organisms. Optimization to nearly absolute physical limits also can be observed in human visual and auditory sensory amplitude sensitivities, and in silk moth olfactory sensitivity to pheromones[14] – i.e., at the very boundary of the neural with its environment. Why should the neural realm sometimes demand optimization, rather than the more familiar biological satisficing?

Mechanisms of neural optimization are best understood against the background sketched earlier, that the key problems of network optimization theory are

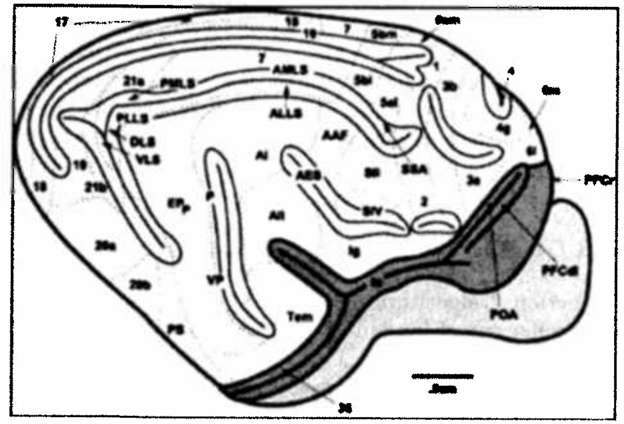

Figure 6: Cerebral cortex of cat. (Lateral aspect; rostral is to right.) Placement of 39 interconnected functional areas of visual, auditory, and somatosensory systems (in white). Exhaustive search of samples of alternative layouts suggests this actual layout ranks at least in the top 100 billionth of all possible layouts with respect to adjacency-cost of its interconnections.[15] – "Best of all possible brains"?

NP-complete, hence exact solutions in general are computationally intractable. For example, blind trial and error exhaustive search for the minimum-wiring layout of a 50 component system (such as all areas of a mammalian cerebral cortex hemisphere), even at a physically unrealistic rate of one layout / picosecond, would still consume more than the age of the Universe.[9] Thus, to avoid these Cosmos-crushing costs, even terrestrial evolution instead must exploit "quick and dirty" approximation / probabilistic heuristics.

One such possible strategy discernible above is optimization "for free, directly from physics". That is, as some structures develop, physical principles cause them automatically to be optimized. There is some evidence for neuron arbor optimization via fluid dynamics, as well as for nematode ganglion layout optimization via "mesh of springs" force-directed placement simulation. For the neural optimization examples above, some of this structure from physics depends in turn on exploiting anomalies of the computational order.[12] For component placement optimization, there is the chicken-egg question of whether components begin in particular loci and make connections, or instead start with their interconnections and then adjust their positions, or some mix of both causal directions. It is worth noting that both a force-directed placement algorithm for ganglion layout, and also genetic algorithms for layout of ganglia and of cortex areas, suggest that simple "connections → placement" optimization processes can suffice.

If the brain had unbounded connection resources, there would be no pressure to refine employment of wiring. So, to begin with, the very fact of neural finitude appears to drive "Save wire" fine-grained minimization of connections. Another part of the functional role of such optimization may be the picture here of "physics → optimization → neural structure". Optimization may be the means to anatomy. The human brain is often characterized as the most complex structure known in the Universe. Perhaps the harmony of neuroanatomy and physics provides an economical means of self-organizing complex structure generation, to ease brain structure transmissibility through the "genomic bottleneck"[4, 7] – the limited information carrying-capacity of the genome. This constitutes a thesis of non-genomic nativism, that significant innate complex biological structure is not encoded in DNA, but instead derives from basic physical principles.[7, 11]

6 Levels of Abstraction for Optimization Models

Marr's [24] "function / algorithm / hardware" levels in cognitive science were preceded by a similar set of levels in top-down methodology of computer engineering (e.g., Dijkstra, late 1960's). This in turn was reflected in the emergence of functionalism / computational psychology in 1960's philosophy of mind. And Chomsky [17] had distinguished between abstract competence and realistic performance models. These levels seem a good way to relate hardware optimization findings to main results of cognitive science at the computational (algorithm, software) level.

An uncomfortable instance of the levels approach has emerged for my research program in recent years. It concerns (in)consistency of optimal neural hardware observations with less idealized rationality models (at the function level). My own critique was that standard rationality idealizations – drawn from microeconomic, game, and decision theory – can benefit from a more realistic approach, particularly regarding limited available cognitive resources.

So, on the one hand, there is a tension or un-consilience – if that is the right euphemism – with the "best of all possible brains" neural hardware results. For what irony it is worth, as noted, the research program in computational neuroanatomy optimization in fact stemmed directly from my own prior philosophical critiques of classical rationality idealizations. – "We don't have God's brain." Our resources at all three levels are bounded. So, neural hardware optimization does not seem to entail rationality optimization. In the other direction, rationality optimization would require much more than brain optimization; indeed, hardware optimization could be dispensed with, if the ideal agent had unbounded resources. Efficiency is then not needed. (As noted, even absolute unsolvability results like Church's Theorem, etc. in fact no longer apply for such an ideal agent.)

However, on the other hand, such suboptimal rationality of our cognitive system does not ipso facto exclude optimality at other levels higher than hardware. – For instance, computational subsystems. Chomsky has focused on syntax, pointedly eschewing even semantics. (E.g., [17].) Fodor extended the Chomskian program beyond syntax; but even he explicitly excludes the central cognitive system from such an approach.[21, p.101] So, the way still seems open for a Minimalist program; the hardware optimization findings might yet point to an approach. The conclusion remains inconclusive.

Bibliography

[1] Cherniak, C. *Pragmatism and Realism.* B. Litt. thesis, University of Oxford, 1973.

[2] Cherniak, C. Minimal rationality. *Mind*, 90:161–183, 1981.

[3] Cherniak, C. *Minimal Rationality.* Cambridge: MIT Press, 1986.

[4] Cherniak, C. Undebuggability and cognitive science. *Communications of Assoc for Computing Machinery*, 31:402–412, 1988.

[5] Cherniak, C. The bounded brain: Toward quantitative neuroanatomy. *Journal of Cognitive Neuroscience*, 2:58–68, 1990.

[6] Cherniak, C. Meta-neuroanatomy: The myth of the unbounded mind/brain. In Agazzi, E., Cordero, A. (Eds.), *Philosophy and the Origin and Evolution of the Universe*, 219–252. Boston: Kluwer, 1991.

[7] Cherniak, C. Local optimization of neuron arbors. *Biological Cybernetics*, 66:503–510, 1992.

[8] Cherniak, C. Component placement optimization in the brain. *Journal of Neuroscience*, 14:2418–2427, 1994.

[9] Cherniak, C. Philosophy and computational neuroanatomy. *Philosophical Studies*, 73:89–107, 1994.

[10] Cherniak, C. Neural component placement. *Trends in Neurosciences*, 18:522–527, 1995.

[11] Cherniak, C. Innateness and brain-wiring optimization: Non-genomic nativism. In Zilhao, A. (Ed.), *Cognition, Evolution, and Rationality*, 103–112. Londo: Routledge, 2005.

[12] Cherniak, C. Neuroanatomy and cosmology. In Bickle, J. (Ed.), *Oxford Handbook of Neuroscience and Philosophy.* New York: Oxford, 2009.

[13] Cherniak, C., Changizi, M., Kang, D. Large-scale optimization of neuron arbors. *Physical Review E*, 59:6001–6009, 1999.

[14] Cherniak, C., Mokhtarzada, Z., Nodelman, U. Optimal-wiring models of neuroanatomy. In Ascoli, G. (Ed.), *Computational Neuroanatomy: Principles and Methods*, 71–82. Totowa, NJ: Humana, 2002.

[15] Cherniak, C., Mokhtarzada, Z., Rodriguez-Esteban, R., Changizi, B. Global optimization of cerebral cortex layout. *Proceedings National Academy of Sciences*, 101:1081–1086, 2004.

[16] Cherniak, C., Rodriguez-Esteban, R. Information processing limits on generating neuroanatomy: Layout optimization of olfactory cortex and amygdala of rat. *Journal of Biological Physics*, 2009. In press.

[17] Chomsky, N. *Aspects of the Theory of Syntax.* Cambridge: MIT Press, 1965.

[18] Chomsky, N. Beyond explanatory adequacy. In Belletti, A. (Ed.), *Structures and Beyond: The Cartography of Syntactic Structures*, 2. New York: Oxford, 2004.

[19] Chomsky, N. Three factors in language design. *Linguistic Inquiry*, 36:1–22, 2005.

[20] Chomsky, N. Opening remarks. In Piattelli-Palmarini, M., Uriagereka, J., Salaburu, P. (Eds.), *Of Minds and Language.* New York: Oxford, 2009.

[21] Fodor, J. *Modularity of Mind.* Cambridge: MIT Press, 1983.

[22] Garey, M., Johnson, D. *Computers and Intractability.* San Francisco: WH Freeman, 1979.

[23] Kahneman, D., Slovic, P., Tversky, A. (Eds.). *Judgment under Uncertainty: Heuristics and Biases.* New York: Cambridge, 1982.

[24] Marr, D. *Vision.* San Francisco: WH Freeman, 1982.

[25] Ryle, G. *Concept of Mind.* Chicago: University of Chicago, 1949.

[26] Simon, H. *Models of Bounded Rationality.* Cambridge: MIT Press, 1982.

[27] Thompson, D. *On Growth and Form.* New York: Cambridge, 1917.

[28] Turing, A. Chemical basis of morphogenesis. *Phil Trans Roy Soc Lond B*, 237:37–72, 1952.