APPENDICES:

These appendices are mostly about heuristics. Appendix A lists important properties of heuristics crucial to explaining characteristics of their use. Appendix B is a general argument for how reductionistic problem-solving leads to biases against recognizing relevant causes outside the system-boundary. Appendix C lists 20 problem-solving heuristics which are reductionistic or have that effect under specified circumstances.

Appendix D has a paragraph each explaining some key possibly unfamiliar concepts or assumptions used in the text. Appendix E reviews a number of "in principle" idealizations found in philosophical or scientific arguments which are unrealistic by assuming that our powers are infinitely (or sometimes just indefinitely) greater than we could possibly achieve. Were they true, these kinds of assumptions would make this book irrelevant. Their obvious falsity points to the need to re-engineer philosophy for limited beings.

Appendix A: **IMPORTANT PROPERTIES OF HEURISTICS:**

I use the term heuristics in a broader sense than is common in the Artificial Intelligence literature, but closer to Herbert Simon's original (and less formal) use when he got the term from mathematician Georg Polya's *Patterns of Plausible Inference* (1954). In AI, heuristics are formal procedures or inference rules. But it is only a small move, using the properties below to see, e.g., the extended waggling "bait"-like tongue in the open mouth of the anglerfish as a heuristic procedure for attracting smaller fish which then become prey. So also, to a migrating animal the presence in a strange environment of a conspecific is a useful indicator that the environment is suitable: organisms characteristically stay longer in places which are fit, and quickly move on in hostile environments. Here a heuristic is viewed most broadly as a regularity for action wherein a kind of action (behavior) is characteristically undertaken under specifiable kinds of circumstances to achieve an end, or as part of a larger plan which is designed to do so. These "action-patterns" have important characteristics which explain why heuristics are adopted, calibrated, combined in larger methodologies to correct for biases and increase robustness, and have the rich characteristics they do in use. These properties are shared with adaptations, and so heuristics are plausibly seen as problem-solving specializations of a broader class of adaptive tools. Items 1-3 appeared and were first applied in 1980b, with results discussed in chapter 5. Item 4 was first noted in Chapter 5, item 5 in Griesemer and Wimsatt1989, and item 6 here. This list may not yet be complete.

(1) By comparison with truth-preserving algorithms or other procedures for which they might be substituted, heuristics make **no guarantees** (or if substituted for a heuristic, different guarantees) that they will produce a solution or the correct solution to a problem. A truth-preserving algorithm correctly applied to true premises **must** produce a correct conclusion. A heuristic need not.

(2) By comparison with the procedures for which they may be substituted, heuristics are very **"cost-effective"** in terms of demands on memory, computation, or other limited resources. (This is why they are used instead of methods offering stronger guarantees.)

(3) Errors produced by using a heuristic are not random, but **systematically biased.** This means two things: (a) The heuristic will tend to break down in certain classes of cases and not others, but not at random. Indeed, with an understanding of how the heuristic works, it should be possible to **predict** the conditions under which it will fail. (b) Where it is meaningful to speak of a **direction of error**, heuristics will tend to cause errors in a certain direction, which is again a function of the heuristic and of the kinds of problems to which it is applied.

(4) Application of a heuristic to a problem yields a **transformation** of the problem into a non-equivalent but intuitively related problem. This means that answers to the transformed problem may not be answers to the original problem, even though various cognitive biases operative in learning and science may lead us to ignore this.

(5) Heuristics are useful **for** something: they are **purpose relative**. Tools which are very useful for one purpose may be very bad for another (Levins, 1968). This may give a way to identify or predict their biases: one would expect a tool to be relatively unbiased for the applications it was designed for, and perhaps more biased for others. One might also expect that increases in performance in one area will be accompanied by decreases elsewhere.

(6) Heuristics are commonly **descended from other heuristics**, often modified to work better in a slightly different environment. Thus they commonly come in evolutionarily related families, which may be drawn upon for other resources or tools appropriate for similar tasks. On different scales of resolution, a family of heuristics may look like a single heuristic, or conversely. (Lenat 1981 provides a beautiful (and unintended?) example).

Appendix B: **THE ORIGINS OF BIAS IN REDUCTIONISTIC HEURISTICS:**

This section was to show how the assumption of limited powers plus a minimal interest in reductionism (seeking to explain system behavior in terms of interactions between parts of the system) biases the problem-solver against recognizing environmental causes of system behavior. But this is covered in the section "Reductionistic Strategies and their Biases" of chapter 5, Heuristics and the Study of Human Behavior, so read it there.

Appendix C**: COMMON REDUCTIONISTIC HEURISTICS:**

I suppose (with most scientists) that a reductionistic analysis offers a lower level mechanistic account of a higher-level phenomenon, entity, or regularity. To do so, one commonly decomposes a complex system into its parts, analyzes them in isolation, and then re-synthesises these parts and the explanations of their behavior into a composite explanation of some aspect of the behavior of the system. [Decomposition and recomposition (Bechtel and Richardson, 1992) is a "near-decomposeability" meta-heuristic for reductionistic problem-solving.] In using this approach, we use heuristic strategies with systematic biases which ignore or downplay the context-sensitivity of the results and the importance of the environment. They are numbered in order of their discovery. Heuristics 1-9 appear in 1980b. Some of these are:

A. Biases of conceptualization:

(1) ***descriptive localization*****:** describe a relational property as if it were monadic, or a lower order relational property. Thus, e.g., describe fitness as if it were a property of phenotypes or genes, ignoring the fact that it is a relation between organism and environment. (This strategy may be justified/facilitated (and its strong assumptions hidden) by fixing the environment, making it artificially disappear as a variable—see complementary heuristics #5, 7, 8 below.)

(2) ***meaning reductionism*****:** assume that lower-level redescriptions change the meanings of terms, but higher-level redescriptions do not. This reflects a kind of atomistic essentialism. Thus we suppose that the meaning of 'gene' is changed when we discover the structure of DNA, but that 'iron' is not when we discover that it occurs as a crucial chelating ion in hemoglobin. The result: philosophers (who view themselves as concerned with meaning relations) are inclined to a reductionistic bias.

(3) ***interface determinism*****:** Assume that all that counts in analyzing the nature and behavior of a system is what comes or goes across the system-environment interface. This has two complementary versions: (a) **black-box behaviorism** -all that matters about a system is how it responds to given inputs; and (b) **black-world perspectivalism** -

all that matters about the environment is what comes in across system boundaries and how the environment responds to system outputs (e.g., Fodor's "methodological solipsism" or Searle's Chinese room). Either can introduce reductionistic biases when conjoined with the assumption of "white box" analysis--that the order of study is from a system with its input-output relations to its subsystems with theirs, and so on. The analysis of functional properties, in particular, is rendered incoherent and impossible by these assumptions.

(11) ***entificational anchoring*****:** Assume that all descriptions and processes are to be referred to a entities at a given level, which are particularly robust, salient, or whatever. This is the ontological equivalent of assuming that there is a single cause for a phenomenon, or single level at which causation can act. Thus the tendency to regard individual organisms as primary, and more important than entities at either higher or lower levels (cf. methodological individualism for rational decision theorists and other social scientists. Similarly for genes for some reductionist neo-Darwinians.) cf. ***perceptual focus*** (#19 below) and ***multi-level reductionistic modelling*****.**

B. Biases of Model-Building and Theory Construction:

(4) ***modelling localization*****:** look for an intrasystematic mechanism to explain a systemic property rather than an intersystemic one. *Corollary 4a*: ***Structural*** properties are regarded as more important than ***functional*** ones (since functional ones require reference to embedding systems).

(5) ***contextual simplification*****:** in reductionistic model building, simplify environment before simplifying system. Thus the environment may be treated as homogeneous or constant (in space or in time), regular in some other way, or random. This strategy often legislates higher-level systems out of existence, (see the "migrant pool assumption" in models of group selection, Wimsatt, 1980b) or leaves no way of describing systemic phenomena appropriately.

(6) ***generalization*****:** When starting out to improve a simple model of the system in relation to its environment, focus on generalizing or elaborating the internal structure, at the cost of ignoring generalizations or elaborations of the existing structure. *Corollary 6a*: If a model doesn't work, it must be because of simplifications in the description of internal structure, not because of simplified descriptions of external structure.

C. Observation and Experimental Design:

(7) **focussed *observation*:** Reductionists will tend not to monitor environmental variables, and thus will often tend not to record data necessary to detect interactional or larger scale patterns.

(8) **environmental *control*:** Reductionists will tend to keep environmental variables constant, and will thus tend to miss dependencies of system variables on them. (*Ceteris paribus* is regarded as a qualifier on environmental variables.) Mill's methods applied with this heuristic (vary the system variables one at a time while keeping all others (always including the environmental variables) constant) will yield as a systematic bias apparent independence of system variables from environmental variables, though the right experiments won't have been done to establish this.

(9) **locality of *testing*** Test a theory only for local purturbations, or only under laboratory conditions, rather than testing it in natural environments, or doing appropriate robustness or sensitivity analyses to suggest what are important environmental variables or parameter ranges.

(10) ***abstractive reification*****:** Observe or model only those things that are common to all cases; don't record individuating circumstances. Losses: (1) sense of natural (or populational) variation; (2) lose detail necessary to explain variability in behavior, or exploitable in experimental design. [Raff (1996) notes that evolutionary geneticists focus

on intraspecific variability, while developmental geneticists focus only on genes which are invariant within the species; which produces problems both of methodology and of focus when trying to relate micro-evolution and macro-evolution or evolution and development. Similarly, cognitive developmental psychologists tend to look only for invariant features in cognition, or major disfunctions, rather than populational variation.]

(12) ***Articulation-of-Parts (AP) coherence*** (Kauffman/Taylor/Schank): Assuming that studies done with parts studied under different (and often inconsistent) conditions are *context-independent*, and thus still valid when put together to give an explanation of the behavior of the whole. (Schank: Checking this gives a non-trivial use for computer simulation.)

(13) ***behavioral regularity*** (Schank/Wimsatt): The search for systems whose behavior is relatively regular and controllable will result in selection of systems which may be uncharacteristically stable because they are relatively insensitive to environmental variations (Schank: regular 4-day cyclers among Sprague-Dawley rats are insensitive to conspecific pheromones; Wimsatt: Mendel's selection of 7 out of 22 characters which are relatively constant and insensitive to the environment probably resulted in unconscious selection against epistatic traits, which (happily) made his model ignoring them less problematic.

D. Functional Localization Fallacies:

(14) Assuming that the function of a part is to produce whatever the system fails to do when that part is absent, (e.g., Spark plugs as "sputter suppressors"), or produced when that part is activated or stimulated. More generally, the error involves *reifying added or subtracted behaviors* of the system *as functional properties* of the manipulated unit. Gregory (1962) notes that the things not done with lesion or deletion experiments may simply be the most obviously affected (rather than the most important). The part could have more importance to functions which are strongly canalized or whose defecits are not revealed under the testing conditions. More generally, even if a part does realize a function, it does so usually only against a background of activities by other interacting components. Judgements of modularity are often insufficiently justified.

(15) Assuming ***simple 1-1 mappings between*** recognizeable ***parts*** and ***functions.*** This can lead to problems in two ways: (1) ignoring pleiotropy: stopping search for functions of a part when you find one [e.g., the newly discovered region of hemoglobin implicated in NO^+ transport, because it was assumed that *the* function of hemoglobin was oxygen transport]; (2) ignored division of labor (when a part's necessity is shown thru deletion studies, etc.) [missing other parts' roles in the hypothesized function because they are part of the constant context, so they are always there to provide it]. See #8 above.

(16) ***Ignoring interventive effects*** and damage due to experimental manipulation. This was first noticed as relevant in neurophysiological studies, but it occurs also in many other places (e.g. marking specimens in mark-recapture studies may affect their fitness).

(17) ***Mistaking*** lower-level ***functions for*** higher-level ***ends***, or misidentifying system which is benefited. This is common in units of selection controversies—either of the apocalyptic variety as with Dawkins (1976) who denies all units of selection at higher levels than the gene, or for eliminative reductionists, who want to deny the existence or significance of large domains of cognitive function. There are legitimate concerns of level in both disputes, but the extremists are commonly seriously wrong.

(18) Imposition of ***incorrect*** set of ***functional categories***. (Common in philosophy of psychology when it ignores ethology, ecology, and evolutionary biology.)

E. Other important biases: (#'s 10, 11 and these can generate either reductionistic or holistic biases in different contexts.)

(19) ***Extra-perspectival blindness*** *or* ***perceptual focus:*** Assuming that a system can be exhaustively described and explained from a given perspective because it has been very successfully and powerfully so described. (Not all problems of biology are problems of genetics, or of molecular biology, physiology, or anatomy (to cite other past excesses) and (as we can now see from a safe distance), not all problems of psychology are problems of behavior. Perceptual focus can artificially inflate the number of properties attributed to a level of organization. Thus, the individual psyche—though perfectly real—has attracted social properties through improper (reductionistic) functional localization fallacies, and other phenomena better explained at lower neurological levels. This bias nteracts with #11 to give **extra-level blindness**, which can be counteracted by doing **multi-level reductionistic modeling,** in which a process is modeled at several levels with results which are then cross-checked.

(20) ***Tool-binding:*** Becoming sufficiently bound to a specific (usually very powerful) tool that one chooses problems for it, rather than conversely ("The right job for the organism", rather than "The right organism for the job"!) This applies to theoretical models and skills as well as to material tools and model organisms. An efficient division of labor if mastery of the tool is very demanding--it is problematic only when it facilitates errors #11 or 16.

Appendix D: **GLOSSARY OF KEY CONCEPTS AND ASSUMPTIONS**
(With references to articles where more can be found; primary ref in **bold**)

AGGREGATIVITY The condition of a system property in which it can legitimately be said that it is "nothing more" than the properties of its parts, justifying ***nothing-but-ism.*** For this to be true, roughly, the system property must not depend upon the mode of organization of the system's parts. (One productive way of defining a property as emergent is to say that it does depend upon the mode of organization of the parts, so aggregativity can be regarded as the opposite of emergence).Being aggregative requires meeting 4 conditions which are almost never satisfied, but these conditions provide useful tools to determine how a system property depends upon the organization of the parts. Reduction is sometimes mistaken for aggregativity, which is a much stronger condition. That is, Reductionism is **not** "nothing-but-ism". Our tendencies to make this confusion (are explicable in terms of the use of reductionistic problem-solving heuristics, and are discussed in **chapter 11.**

BOUNDARIES Robustness for objects involves coincidence of boundaries. This brings boundaries into focus. Strong gradients in values of many properties can cause other properties to develop gradients in the same place. This reinforcement of boundaries can lead to the "spontaneous" emergence of robust systems as objects, natural and biological (Platt 1969) and social (Abbott 1995). But systems may have robust boundaries while their parts do not, leading to a situation which naturally promotes functional localization fallacies and failures of near-decomposability (**Chapters 8**, 9, 11). Systems with multiple partially overlapping boundaries have richer possibilities for interpenetrating interactions, but are for that reason harder to individuate. These constitute a new class of COMPLEX OBJECTS found much more commonly in the biotic and social realms. The multiple boundary crossings yield causal thickets and disciplinary conflicts (Chapter 9).

ENGINEERING PERSPECTIVE. A cluster of theses derived from the assumption that theory has much to learn from practice and application. Teleological. Design is design for an end. View scientific activities as functional, and evaluate their designs for that supposed end. (1979, chapter 10) Relation to practice: focus not only on theiry and arguments in principle, but on what the practical implications of a view of science is, and how to apply it, and how it must be adjusted or qualified to make it applicable. The central role of heuristics as fallible inferential tools, rather than sources of certainty. Applied not only to our theories and methods as instruments, but also to our mental capabilities and inferences. Most engineering as Re-engineering, recognizing that we rarely start from scratch, but will use what comes readily to hand, as quicker, cheaper, more convenient. This has two consequences: (1) History matters; to understand our methods we must understand where they came from and how. The genetic fallacy is not a fallacy. (2) There is no "perfect adaptation" ex nihilo: adaptation is an adaptation of something else to a new role, so exaptation is common. This view is profoundly instrumental, but denies any necessary tension between instrumental usefulness and truth or realism.

ERROR: The Axiomatic (or 'Euclidean') world view, and its descendant, the Computational world view are specialized adaptations for dealing with structures which don't need to be changed. Ideally (or 'in principle') they have no errors. In the real world, they (i.e., their implementations) have small error rates and relatively easily localized errors, but this is an uncommon and profoundly non-biological model of organization. Optimal performance in such networks requires quite different strategies than are appropriate for structures with higher error rates. Evolution and evolutionary ecology are laboratories for the design of 'error tolerant' (and even 'error metabolizing') systems. We have invested lots of effort in investigating the potential of Euclidean systems, but almost nothing in theoretically characterizing Organic/Robust systems (**chapters 1, 2, 3, 4**) or THE METABOLISM OF ERROR, the deliberate use of error (in model-building) to increase knowledge (chapter 6).

EXAPTATION: When something is used to serve a function which it wasn't originally designed (or selected) for, it is called an exaptation rather than an adaptation--a term first introduced by Gould and Vrba (1982). Gould also championed the idea that most elaboration of functional organization is through exaptation rather than adaptation (design for that function). Thus the reasons for a given functional design are inextricably historical and complex, and seldom transparent to someone looking only at current function. So in evolutionary history, organisms are not so much engineered as re-engineered through a succession of kluges interpolated with retunings to make them work better together.

FUNCTIONAL LOCALIZATION FALLACIES A functional localization fallacy occurs when a function is attributed or localized incorrectly in a system. With reductionist methodologies, the most common kind of mis-localization is to attribute a function properly attributed to a whole system to just a part of that system. This error is particularly easy when one part is particularly central to that function, as e.g., the brain is to consciousness. This bias may misattribute many social properties as individual psychological ones. Related errors occur in treating a relational property as monadic, or in underestimating the degree of a relational property. For different kinds of functional localization fallacy, see **appendix C** and (in passing) chapter 11. More fully discussed in Bechtel and Richardson 1992.

GENERATIVE ENTRENCHMENT (or **GE**): A measure of how many things depend upon an element and therefore are subject to change if it changes. (In an abstract network, if robustness (means of access) are paths to a node, then GE is the reach of paths from a node. Thus Robustness and GE are complementary measures of local order in a complex system. What is derived from what depends upon the manipulations or inferences involved, so whether a given relation involves one or the other may depend on the specified operations.) Things with higher GE are more evolutionarily conservative because the chance that random changes in them will be adaptive declines exponentially with increasing GE. They also generate more massive changes when they do change. Things that stay around long enough get entrenched and more resistant to change because they have more things depending on them and depending on them to greater degrees. Evolution is an ongoing dialectic of local adaptation to an (internally and externally) changing environment which is also partially a biotic product. This correlates with the layered cyclic process process through which adaptations to some things become exaptations, and then through modification, adaptations to others. This process is the major way through which the evolutionary history of an adaptive system becomes relevant to its current form. (1981b, **1986**, 1988a, 1988b, 1997b, **2001, chapter 13**).

Our adaptations meet the defining characteristics of HEURISTICS—or to put it in the phylogenetic order; our heuristic problem solving principles are specialized cognitive adaptations, and are still marked by five important characteristics of this origin. (1980b, **chapter 5**, appendices A, B, C).

KLUGE: Originally a programmers term; an unpretty but conditionally effective fix for a program or design failure or "bug". It may be inelegant or unpretty by violating common principles of design (and thus prone to other failures), by using something not intended for that purpose (see "exaptation"), or by doing the fix in a way which is not efficient or robust. In conscious design processes, a kluge is more likely if the cause of the failure it "fixes" cannot be determined, leading one to look for other ways to block the failure than the preferred way of redesigning to remove the cause. Since a kluge (because inelegant) is itself harder to understand (especially if undocumented), use of kluges increases the probability of having to use more. Inelegant solutions which are products of past historical commitments are kinds of kluges which have analogues in evolution. Mutations have diverse effects "randomly" dispersed throughout the phenotype. They are "selected" when the net benefit of these effects exceeds net loss, but may fail anywhere due to sampling processes (genetic and ecological "drift"). The standard of selection is more likely met for changes with greater positive excess,but applies only locally and incrementally: there is no design process directing the accumulation of successive kluged increments, so they generally are not "systematic". Thus they are incorporated not for ultimate optimality or efficiency, but on a "first come" basis, leading naturally to generative entrenchment of the earlier modifications.

LAWS The compositional multi-level materialism I favor takes mechanisms as more fundamental than laws at all higher levels of organization than the physically most fundamental. Whether expressed as "principles" (Hardy-Weinberg principle) or law (Mendel's law of independent assortment) these are mechanism based expressions of the operation of causal factors under specifiable standard or idealized conditions. (discussed here in **chapter 10**, but argued more fully in Glennan 1996, 1997).

LEVELS: Natural systems are often found in compositional levels of organization, with the entities at neighboring levels related by part-whole relations. *Size scale* is a pivotal indicator of the effectiveness of whole families of causal processes (because of the robustness of most objects at levels), and also largely determines (given the forces at work) a correlative characteristic time scale or "relaxation time"—a relatively narrow range of rates over which most processes at that level happen or go to equilibrium. As a result, our theories come in levels (language tracks Nature rather than conversely), because it is in the relations among robust objects that you can get the "biggest bang for a buck" in theory construction, and most of what "happens" between them takes characteristic times within that range. The aim of science is the articulation and coordination of different entities and phenomena at different levels—the discovery of mechanisms [static mode] or processes [dynamic mode] relating phenomena and regularities at different levels rather than eliminative "nothing-but"-ism (1976a, **Chapter 9.**)

MULTI-PERSPECTIVAL REALISM: Realism is connected with robustness (detectability via multiple independent means). A particularly important variant of this occurs when the different means derive from different perspectives (see below).When a conceptuasl scheme claims exclusiveness or exhaustiveness, or treats other conceptual schemes as competitors when it cannot establish its primacy, this leads to relativism. But if these schemes are recognized as perspectives (severally incomplete, mutually complementary, and possibly co-calibrating), the consequence is a realism recognizing their respective co-referring objects of study as robust multi-dimensional trans-perspectival objects. Chapters 8, 9 for analysis and examples.

NEAR DECOMPOSEABILITY (Simon, 1996): The ability to break structures into parts, and then reassemble them to solve or engineer problems is an impressively powerful heuristic. When this is possible, the system is said to be "nearly-decomposable" (**Chapter 8**, **1986**). Reductionism in science, structured programming, engineering, mass production from stable sub-assemblies, and tinkertoys all use (and teach) this strategy. It is so powerful and so endemic to Western industrial society that we find it hard to see when it breaks down, but it often does. Reductionistic problem-solving heuristics, their strengths, limits, and systematic biases are covered in1980b, **chapters 5** (heuristics), **11** (aggregativity)] See also aggregativity and functional localization fallacies.

PERSPECTIVES: In complex systems, "perspectives" give organized approaches to a cluster of problems and techniques, often span levels, cross-cutting levels and each other, and give knowingly incomplete descriptions of the systems to which they are applied. (Levels can be viewed as special cases of perspectives ordered by hierarchial part-whole composition relations.) Their relation to each other, how they partition the systems to which they are applied, and the degree to which problems for a given system are bounded within individual perspectives on it are crucial to characterizing the complexity of a system. This kind of analysis is also required to understand the behavior of robust systems which must be simultaneously described using multiple boundaries and decompositions (1974, 1976a, **Chapter 9**). Reality is multiperspectival and robust. Some systems get so complex (the causal interactions among their variables are sufficiently disordered) that levels and perspectives break down, failing to have the partial dynamical and explanatory closure characteristic of both. Problems can (and must) be approached from a variety of directions, any one of which makes competing methodological claims with the others. In this situation, all we have is "CAUSAL THICKETS": methodological disputes are rife between researchers from different levels and perspectives over the common territory that they claim, and functional localization fallacies are hard to avoid.

REDUCTIONISM: There are three types of "reduction" or "reductionism" which are often confused: *successional* reduction (which applies either to same level theories or to universal theories) is a similarity relation (and is thus *intransitive*) in which one seeks to localize similarities and differences between a new theory and an older one in order to further the development of the newer theory and to delimit the useful range of the older one. This is plausibly thought of as a kind of theory-reduction. [1976a, **chapter 10,** Kauffman 1971]. *Inter-level* reduction involves an attempt to explain upper-level phenomena in terms of an articulated account of the operation of lower level MECHANISMS. It will commonly involve identification of descriptions from two levels of organization as referring to the same thing or more commonly, weaker species of identity relation, (qualified context-dependent identification, an instanciation, realization, or a localization). All of these forms of identification are assumed to be *transitive* as one moves up and down levels of organization. In spite of this important difference, neither the successional relation nor the identification is usefully captured as a truth-preserving deductive derivation because of the use of limiting case approximations (in both cases), and truncation of relevant contextual details (in the second). In the compositional sciences, (inter-level) reduction proceeds by developing mechanistic explanations through an iterative feedback process relating the levels (1976a, **chapter 9).** A common mistake of critics and advocates alike is to suppose that reduction supports "*nothing-but-ism*", (as in "the whole is nothing more than the sum of the parts") but this would require the much stronger conditions of *aggregativity*, which are rarely met **(chapter 11)**.

ROBUSTNESS is the existence of multiple means of determination, or access, to an entity, property, process, result, or theorem, which are at least partially independent. This allows lower reliability of the individual components, while still getting higher reliability and adaptability of the overall structure. [This is the fundamental organic design principle, not genetic determination of phenotype]. Robustness is the ultimate methodological criterion for reality *and is so in all sciences*. It also characterizes the appropriate kind of stability for ecosystems. (1974, 1976a, 1980a, **chapter 4,** 1991, chapter 9),

We are creatures (and cultures) engineered by SELECTION processes at a variety of intersecting—sometimes coupled, sometimes dynamically independent—levels. (1976a, **chapter 9**) Biological and Cultural evolution use different transmission channels and somewhat different principles, but show significant theoretical similarities (and also differences forcing different theoretical approaches to many of their problems). Both are still covered by "Darwin's Principles". However for virtually all significant evolutionary processes, "Darwin's Principles" must be supplemented by 2 additional principles which yield a developmental structure of the phenotype and a life cycle in terms of generative entrenchment (1999a, **2001, Chapter 13**). These 2 additional principles and the developmental organization they bring to bear are even more important for cultural evolution than for biological evolution. I see self-organization as continuous with selection, though it is often useful to treat them as separate processes (**1971, 1986-**last section).

Appendix E. A PANOPLY OF LAPLACEAN AND LEIBNITZIAN DAEMONS

19th century physicists invented "demons" to go along with the various perpetual motion machines forbidden by newly discovered fundamental laws. Philosophers, in their idealizations seem to have forgotten this, and biologists have contributed some as well. Here are some of my favorites. Add your own.

Heuristic for finding these: Look for ***in principle*** claims which are treated as if they could be delivered if anyone were sufficiently interested—but aren't and can't. (Given how important they are to the argument, if they can, why aren't they?) There is bound to be a daemon lurking in the background. (**Chapter 10**). Computability variants of this are OK and useful in mathematics, but when applied to the real world require impossible amounts of computation, degrees of knowledge, and (often ignored), a systematic way of ordering the states or alternatives so that they can be exhaustively specified. Dennett's (1995) versions of Borges' labyrinth contain several ingenious examples.

1. Laplace's original demon and the definition of determinism: a deterministic system is one in which, given the equations of motion of the system, and a complete description of it at any point in time, one can predict or retrodict its state arbitrarily far into the future or arbitrarily back into the past. Notice that genetic determinism cannot be formulated as true, subject to this constraint: LaPlace's demon would require non-genetic information since *even if one were to accept that all organic products are genetic products*, what happens at any given time requires the various levels of biological organization, (e.g., the cell) and thus requires the action of the genes over a temporally extended period of time (in a context of non-genetic elements.)

2. The rational demon and the definition of expected utility, requiring (1) mutually exclusive and (2) exhaustive specification of alternative actions, (3) mutually exclusive and (4) exhaustive specification of outcomes, and (5) their utilities; application of scientific theories and initial conditions to compute (6) complete specification of probabilistic arrays mapping actions to outcomes, (7) leading (then almost trivially) to computation of expected utilities of all actions and choice of maximal one. (Note that this in effect requires a state-space representation of the system.)

3. Descartes' evil demon and the sceptical regress (and modern brain-in-the-tank descendants. cf. Fodor's "methodological scepticism", and Searle's (reductio) "Chinese room".)

4. Epistemological daemons: knowledge-and-belief structures which are closed under entailment (and globally consistent).

5. Physical (or biological, or neurophysiological, or econo-psychological) reductionist demon, able to provide the *in principle* reductionist translations obviating the need for the relevant upper-level theories and descriptions of phenomena.

6. Lewontin's (1966) probabilistic demons (their estimates of means converge, and the means are sufficient for prediction of optimal behavior) vs. the capriciousness of nature

because of the decay of information on different time scales and pathological distributions (e.g. pareto distribution) for some kinds of events.

7. Darwinian instrumentalist demon: Like Darwin's view of natural selection, infinitely sensitive to detail (but only to functionally relevant properties), and selects always for the best. Although apparently weaker than LaPlacean demon, it is not clear that this demon can get by with knowing any less (because of the causal structure of the world and intertwining of functional, neutral, and disfunctional consequences).

8. Chaos has led to the denial of a "chaos demon", a LaPlacean demon who knows all parameters of the universe with the infinite precision necessary to accomplish Laplace's task in a chaotic world. (A temporally bounded version of same would only need to know what is going on within the bounded neighborhood necessary for prediction within the chosen time interval.)

9. The "meaning daemon"—able to keep track of all the semantic linkages (and to keep them updated in the light of new connections) in order to realize the maximally connected network required by holistic functional theories of meaning. Meaning holists don't distinguish importantly different cases here. Must everything be directly connected to everything (n^2 linkages for n elements, assuming pairwise connections only), or will it do just for everything to be reachable from every other element in the set (then as few as n-1 connections for n elements.) Also, the topology of connections should matter, as (in any more realistic models) their directionality.

10. The Gricean daemon—able to keep track of the recursively enumerated nested intentions necessary to unpack speakers' meaning on Grice's account, and to update the changing purposes of the parties of the conversation (42 times in 7 minutes, according to Nancy Stein's analysis of a recent conversation) sufficiently rapidly to do this. (The Gricean daemon—though respectable—is actually a dinky finite-state automaton by comparison. with the others.)

11. The Turing/Church daemon: able to complete any computable task. Lead to "Turing machine functionalism", and—with Kleene's proofs (in Shannon and McCarthy, *Automata Studies,* 1956) that mapped the powers of networks of McCulloch-Pitts neurons onto various classes of Turing machines, various "curve fitting" but radically biologically unrealistic "solutions" to the mind-body problem. Useful against someone who claims that there is a specifiable behavior which is *in principle* not realizeable by a machine, but otherwise irrelevant. Continued as plausible thru behavioristic arguments that only behavior could matter to an analysis of mind, since only that was accessible to us as we learn language. (Ignores the problem of how —i.e., with what hardware, and consequent limitations and failure modes—*we* accomplished it.) The dual of the brain in the tank or "chinese room" puzzles.

12. The Craigean elimination daemon. Use Craig's theorem and Ramsified sentences to eliminate the theoretical entities of any theory, therefore rendering it more "observationally justifiable" though essentially unusable.

13. Dawkins/Kitcher/Sterelny "bookkeeping daemon": Keep track of all contexts of all genes (including their genetic contexts to adjust for epistatic interactions) in all organisms so as to be able to calculate and update as necessary each generation (asynchronously, for different generation times of different organisms) the net selection coefficients of all genes so as to plug into the "bottom up" genic theory of natural selection required for dawkins reductionistic vision.