

Chapter 4

Darwinian Algorithms

4.1 Another view of human rationality

If adherence to the canons of formal logic is the measure of human rationality, then humans are not very rational. But there is another, teleological view of rationality: An organism is behaving rationally when it is behaving purposefully -- when it is employing means that are likely to accomplish its goal. By this criterion, humans may indeed be rational beings. On this view the question of human rationality becomes: Are our reasoning processes appropriate to the problems they were designed to solve?

* * *

Unless you are a creationist, you probably believe that the human mind -- like the rest of the body and its functions -- is the product of evolution.

This insight was of little value to psychologists at the time of William James, or when John B. Watson was debating William McDougall, because evolutionary theory was in its infancy -- it was too hazy, too imprecise.

This situation has changed dramatically in the last 20 years. Now, the dynamics of natural selection can be mathematically modeled with great precision. This allows evolutionary biologists to determine what kinds of traits will be quickly selected out, and what kinds of traits are likely to become universal and species-typical.

Consequently, evolutionary theory now can be used as a heuristic guide for psychological theory. This heuristic rests

on the recognition that natural selection has produced psychological mechanisms as responses to various selection pressures in a species' "environment of evolutionary adaptedness" (Bowlby, 1969) -- the environment in which the species evolved. The more important the adaptive problem, the more intensely selection will have specialized and improved the performance of the mechanism for solving it.

Our species spent over 99% of its evolutionary history as Pleistocene hunter-gatherers. During that time, the dynamics of natural selection should have operated in the production of information processing mechanisms just as they did in the production of morphological and physiological mechanisms. The Pleistocene savannahs are the human environment of evolutionary adaptedness; our cognitive processes should be adapted to it, not necessarily to the 20th century industrialized world.

Recently, a number of cognitive scientists -- Chomsky, Fodor, Marr -- have argued that the best way to understand any mechanism, either mental or physical, is to first ask what its purpose is, what problem was it designed to solve (e.g., Chomsky, 1975; Fodor, 1983; Marr & Nishihara, 1978).

That is exactly what evolutionary theory allows you to do -- it allows you to pinpoint what kind of problems the human mind should be very good at solving. And although it cannot tell you the exact structure of the algorithms that solve these problems, it can suggest what design features they are likely to have. It allows you to develop a "computational theory" for that problem domain: a theory specifying what functional characteristics a mechanism capable of solving that problem must have (Marr &

Nishihara, 1978).

From the point of view of evolutionary theory, it is very unparsimonious to assume that the human mind is a general purpose computer with domain general, content-independent processes. There are domains of human activity for which the evolutionarily-appropriate information processing strategy is complex, and deviations from this strategy result in large fitness costs. For such domains, humans should have evolved "Darwinian algorithms" -- specialized learning mechanisms that organize experience into adaptively meaningful schemas or frames. When activated by appropriate problem content, these innately specified "frame-builders" should focus attention, organize perception and memory, and call up specialized procedural knowledge that will lead to domain-appropriate inferences, judgments, and choices. Like Chomsky's language acquisition device, these inference procedures allow you to "go beyond the information given" -- to reason adaptively even in the face of incomplete or degraded information (Bruner, 1973).

There are many domains of human activity that should have Darwinian algorithms associated with them. Aggressive threat, mate choice, sexual behavior, pair-bonding, parenting, parent-offspring conflict, friendship, kinship, resource accrual and distribution, disease avoidance, and predator avoidance are but a few. Social exchange is another. The dynamics of natural selection rigidly constrain the kinds of social exchange that can evolve, providing insight into the structure of the mechanisms that regulate it. This structure, and its consequences for performance on the Wason selection task, will be explored in

Chapters 5 and 6.

4.2 A brief primer on natural selection

We came into existence through the process of evolution, and the single ordering process in evolution is natural selection. Therefore, whatever systematic properties we have were produced by natural selection. If we wish to understand these properties, we need to understand the process that produced them.

An allele is a gene that occupies a particular location (locus) on a chromosome. Only one allele can occupy one locus on a particular chromosome. Let's say that in a given population of individuals, 50% of the individuals in the population have allele A at a particular locus, and 50% have allele B at that locus. A and B are alternative alleles -- a given individual has either A or B at a particular locus, but not both. Evolution is a change in the proportion of alternative alleles in a population. Hence, if the ratio of A:B alleles changes from 50:50 to 60:40, then evolution has occurred. Evolution is, therefore, a zero sum game: any increase in the proportion of A alleles in the population comes at the expense of alternative alleles, like B.

Evolution can occur in only two ways. The proportion of alternative alleles can change either via random processes (such as genetic drift) or via natural selection. Natural selection is the process whereby an allele increases its representation in the gene pool by virtue of the effect it has on the individual carrying it. If allele A has an effect on its carrier that, for any reason, causes an increase the proportion of A over B alleles in the population, then natural selection has occurred -- A has

been selected for, and B has been selected against. Technically, "fitness" refers to genes, not to individuals. Allele A is considered more "fit" than allele B if A codes for a trait that increases the proportion of A alleles over the proportion of B alleles in the population.

But genes are located in individuals. This means that there is one, and only one, way that an allele can increase its relative frequency -- by coding for traits that enhance the reproduction of the individual carrying that allele, or of its relatives, who may also bear that particular allele. An allele's fitness is a direct function of the relative number of offspring produced who have a copy of that allele.

Selection does not occur for the "good of the species", the "good of the group", or even for the "good of the individual" -- in fact, it is not even clear what these expressions mean. Genes that code for traits that enhance their own replication will spread through the population, even if this eventually causes the species to become extinct (for excellent discussions see Williams, 1966, and Dawkins, 1982).

In short: Genes -- through the traits they influence -- can influence the rate of their own replication, and hence their frequency in subsequent generations. If they code for traits that influence their own replication positively, then there is positive feedback and their representation in the population increases; if they code for traits that influence their own replication negatively, then there is negative feedback, and their representation in the population decreases. Natural selection is the process by which positive and negative feedback

of genes on themselves regulate the existence and frequencies of those genes over generations. Genes that code for traits that tend to maximize their rate of replication -- their "inclusive fitness" -- will tend to spread through the population, displacing alternative alleles each generation, until the traits those genes code for become fixed in the population. When this happens, the trait coded for is called a "species-typical" trait.

By analyzing the dynamics of gene flow through populations in terms of natural selection theory, one can determine what kinds of traits are most likely to become species-typical. In cognitive psychology, such species-typical traits have been called "cognitive competences" (Chomsky, 1975).

"Adaptation" means something precise: An adaptation is an aspect of the organism that has come into being because it had the effect of promoting the frequency of the genes that code for it. Natural selection theory concerns itself with adaptive function: Why does one trait come to dominate a population rather than another? How susceptible is the population to invasion by mutant genes coding for a different trait? How does one trait compare with another in terms of replicating the genes that code for it? In speaking of mental algorithms as adaptations, one is asking: How does one information processing strategy compare to another, as measured in terms of the replication of the genes that underlie each?

The traits of an organism are there either as a result of random processes (the stochastic dimension of evolution) or because they were selected for because they promote the spread of the genes that code for them. This second process is natural

selection, and it is major shaping force of any ordered relation in organisms.

What does this mean for social exchange? If social exchange was important in human evolution, then the mechanisms that regulate it will have been shaped by natural selection. This means they must promote the spread of the genes that code for them. This has profound consequences for what kinds of strategies for engaging in social exchange can evolve, and these will be discussed in depth in Chapter 5.

Natural selection theory is not behavior genetics

Natural selection theory is not behavior genetics. Behavior genetics attempts to ascribe differences in the behavior of individuals to differences in their genes. In contrast, natural selection theory's primary utility is in discovering what sort of traits are likely to become universal and species-typical.* Differences in behavior between individuals are presumed to be the facultative responses of species-typical traits to differences in those individuals' environments.

All traits have a "genetic basis", and can therefore be acted on by natural selection; This claim has nothing to do with issues of nature "versus" nurture

Phenotypic traits must have a "genetic basis" if natural selection is to act on them. Although few people have difficulty accepting this premise when the trait in question is eye color, leg number, or height, they balk when the trait in question is a

* Except in cases of frequency dependent selection, in which a population is expected to manifest a balanced polymorphism.

psychological mechanism or a behavioral tendency. Many social scientists believe this premise entails vast and unwarranted assumptions about a hotly debated empirical issue: To what extent can environmental manipulations alter the phenotypic expression of a trait? They believe that evolutionary thinking requires a "nativist" stand on the "nature/nurture" issue. It does not.

If properly understood, the premise that traits have a genetic basis entails quite minimal claims that should be acceptable to any cognitive psychologist. In fact, these claims should be acceptable to anyone who realizes that it is impossible, in principle, for a tabula rasa to learn anything (e.g., Hume, 1777/1748; Kant, 1766/1781; Quine, 1969; Popper, 1972).

Genes are the blueprint for the development of an organism. More precisely, they are molecules of DNA which, by virtue of their physical structure, organize surrounding molecules into an enveloping organism -- an entity that can replicate the DNA that built it. The phenotype is the manifest organism, the collection of morphological, physiological, mental, and behavioral properties -- however construed -- that make up the organism. Genotypes specify phenotypes. That is, by virtue of its molecular structure, the genotype carries "instructions" for building an organism in a given environment. Thus there is a causal link between genotype and phenotype.

Natural selection can be thought of as the process whereby environments shape the characteristics of organisms by affecting the frequency of alternative alleles. There are both constancies and regular variations in any species' environment of evolutionary adaptedness that natural selection can be expected

to take advantage of. Thus, there are two causal links between environment and phenotype: 1) environments select genotypes, which specify phenotypes; and 2) the genotype's instructions are environment specific -- the phenotype built "assumes" certain environmental characteristics. Change those characteristics, and you will very likely change the phenotype.

For example, the arrowleaf plant's environment of evolutionary adaptedness included both watery and dry habitats.* When the arrowleaf plant sprouts in water, it develops wide leaves; when a genetically identical clone sprouts on dry land, it develops narrow leaves. An arrowleaf plant whose genes could not respond facultatively to these regular variations in its environment would be at a selective disadvantage -- its seeds could prosper in one habitat, but not in the other.

However, the fact that the leaf's width varies with wetness does not mean that it varies with every environmental dimension, nor does it mean that its shape is infinitely plastic. Leaf width does not vary, for example, with the amount of poetry read to the plant. Similarly, there is probably no environmental factor that would cause the arrowleaf's leaves to grow into the shape of the Starship Enterprise.

Certainly it would be correct to say that "leaf width" has a "genetic basis" in the arrowleaf. The plant's genetic blueprint specifies how wide or narrow the leaf will be in various environments. Furthermore, the arrowleaf's genetic blueprint prevents its leaves from assuming the shape of the Starship Enterprise, and prevents them from being affected by poetry

* From a 1976 lecture by E.O. Wilson on norms of reaction.

readings. Thus, it would also be correct to say that the traits "failure to assume the shape of the Starship Enterprise" and "failure to be affected by poetry readings" have a "genetic basis" in the arrowleaf plant. However, the claim "Leaf width in the arrowleaf has a genetic basis, therefore it cannot be affected by variations in the environment" is clearly false. Leaf width has a genetic basis, yet a simple environmental manipulation -- planting it in a wet or dry habitat -- changes the width of the arrowleaf's leaves.

All traits have a "genetic basis." No matter how far from "direct" genic influence a trait seems to be, it is still built from structures and processes that were built from structures and processes that were, ultimately, specified in the organism's genetic blueprint.

One of the things a genetic blueprint does, however, is specify how a trait will develop under different environmental conditions. Traits differ in terms of which environmental factors can affect their development, and in what ways their development can be affected. This, then, is why evolutionary theory takes no position on nature/nurture questions: Although evolutionary theory requires the assumption that the physical, mental, or behavioral traits it discusses have a genetic basis, it frames no hypotheses as to how these traits may be affected by environmental manipulations. That is a question for physiologists, developmental biologists, and psychologists: for scientists who study the structure of the particular physical and mental mechanisms an organism has. Evolutionary biologists do not study mechanisms, they study questions of adaptive function

-- questions like why genes that code for leaf width that varies with wetness would outcompete genes that do not.

If one's goal is to change a trait in some particular way, one must understand the mechanisms by which environmental factors influence the phenotype. Environmental factors can influence the phenotype either by acting directly on the genes (mutations, viruses) or by acting on mediating structures or processes built according to the genes' specifications -- the phenotype. The mediating structures can be physical organs and physiological processes, innately specified mental structures and processes, or "higher level" algorithms constructed by learning processes constructed by innately specified structures and processes. Environmental factors cannot act on a vacuum. If you want to understand how certain kinds of information will affect learning, you have to study the mind's information processing mechanisms, and natural selection theory can help in this endeavor: the structures and processes that environmental factors act on are, ultimately, built by the genes, and were shaped by natural selection. But natural selection theory cannot, by itself, tell you which manipulations will produce which effects.

4.3 Why should Darwinian algorithms be specialized and domain specific?

Nature has kept us at a great distance from all her secrets, and has afforded us only the knowledge of a few superficial qualities of objects; while she conceals from us those powers and principles, on which the influence of these objects entirely depends. Our senses inform us of the colour, weight, and consistence of bread; but neither sense nor reason can ever inform us of those qualities, which fit it for the nourishment and support of a human body.

-- David Hume (1977/1748, p. 21)

Genes coding for psychological mechanisms that maximize the inclusive fitness of their bearers will outcompete those that do not, and tend to become fixed in the population. The maximization of inclusive fitness is an evolutionary "end"; a psychological mechanism is a means by which that end is achieved. Can a psychological mechanism be domain general and content-independent, yet realize this evolutionary end?

Consider how Jesus explains the derivation of the Mosaic code to his disciples:

Jesus said unto him, "Thou shalt love the Lord, thy God, with all thy heart, and with all thy soul, and with all thy mind. This is the first and great commandment. And the second is like it, Thou shalt love thy neighbor as thyself. On these two commandments hang all the law and the prophets."

Matthew 22: 37-40 (emphasis added)

Jesus has given his disciples a domain general, content-independent decision rule to be used in guiding their behavior. But what does it mean in practice? Real life consists of concrete, specific, situations. How, from this rule, do I infer what counts as "loving my neighbor as myself" when, for example, my neighbor's ox falls into my pit? Should I recompense him, or him me? By how much? How should I behave when I find my neighbor sleeping with my spouse? Should I fast on holy days? Should I work on the Sabbath? What counts as fulfilling these commandments? How do I know when I have fulfilled them?

In what sense does all the law "hang" from these two commandments?

It doesn't. That is why the Talmud was written. The Talmud is a "domain specific" document: an interpretation of the commandments that tells you what actions count as "loving God"

and "loving your neighbor" in the concrete, specific situations you are likely to encounter in real life. The Talmud solves the "frame problem" (e.g., Boden, 1977; Fodor, 1983) posed by a "domain general" rule like Jesus's.

A domain general decision rule like "Do that which maximizes your inclusive fitness" cannot guide behavior in ways that actually do maximize fitness, because what counts as fit behavior differs from domain to domain. Therefore, like the Talmud, psychological mechanisms governing evolutionarily important domains of human activity must be domain specific.

The easiest way to see that Darwinian algorithms must be domain specific is to ask whether the opposite is possible: In theory, could one construct a domain general, content-independent decision rule, that, for any two courses of action, would evaluate which better serves the end of maximizing inclusive fitness?

Such a rule must include a criterion for assessing inclusive fitness: there must be some observable environmental variable against which courses of action from any domain of human activity can be measured. The simplest variable that correlates with inclusive fitness is number of grand-offspring produced by the end of one's life. Using this criterion, the decision rule can be rephrased more precisely as, "Choose the course of action that will result in more grand-offspring produced by the end of your life."

But how could one possibly evaluate alternative actions using this criterion? Consider a simple, but graphic example: Should I eat feces or fruit?

Clearly, I do not have two parallel lives to lead for purposes of comparison, identical except that in one I eat feces and in the other, fruit. Will trial and error work? Although I do not know it, if I eat the feces, there is a good chance I will contract a disease and die -- a large fitness cost. And if I eat the fruit and do not die, I still do not know if I can eat feces: for all I know, feces could be a rich food source that would greatly increase my fecundity. Could I learn from others? If I watch some people eat fruit, and others eat feces, and notice that more feces-eaters than fruit-eaters die by some later point in time, how do I know whether their death was caused by eating feces or by one of the many other things they did before their illness? And why would they choose to learn this way when I do not? -- my population of "guinea pigs" would be selecting themselves out. Furthermore, if, like most Pleistocene hunter-gatherers, I am living among my close kin, their death through experimentation is also a fitness cost to me (see Chapter 5).

Perhaps I could smell both: I'll eat what smells good and avoid what smells bad. But this method violates the assumption that the information processing system is domain general, and side-steps the "grand-offspring produced" criterion entirely. Nothing smells "intrinsically" bad or good; feces smell just fine to dung flies. Moreover, why would I infer that foul-smelling entities should not be eaten? Admitting smell or taste preferences is admitting domain specific innate knowledge. Admitting the inference that foul-smelling or foul-tasting entities should not be ingested is admitting a domain specific innate inference.

Without domain specific knowledge like this, how would I possibly learn to avoid feces and ingest fruit? Even if this were possible, an individual with appropriate domain specific knowledge would enjoy a selective advantage over one who relies on "trial and possibly fatal error" (Shepard, 1985). The tendency to rely on trial and error in this domain would be selected out; domain specific Darwinian algorithms governing food choice would be selected for, and become a species-typical trait.

There is also the problem of deciding which courses of action to evaluate. The possibilities for action are infinite, and the best a truly domain general mechanism could do is generate random possibilities to be run through the inclusive fitness decision rule. When a saber-toothed tiger bounds toward you, what should your response be? Should you file your toenails? Do a cartwheel? Sing a song? Is this the moment to run an uncountable number of randomly generated response possibilities through the decision rule? And again, how could you compute which possibility would result in more grandchildren? The alternative: Darwinian algorithms specialized for predator avoidance, that err on the side of false positives in predator detection, and, upon detecting a potential predator, constrain your responses to flight, fight, or hiding.

The domain general "grandchildren produced" criterion fails even in these simple situations. How, then, could it work in more complicated learning situations, for example, when an action that increases your inclusive fitness in one domain decreases it in another? Suppose your domain general learning mechanism somehow allowed you to figure out that sexual intercourse is a

necessary condition for producing offspring. Should you, then, have sex at every opportunity?

According to evolutionary theory, no. There are large fitness costs associated with, for example, incest.* Given a potential partner with a physique, personality, or resources that would normally excite you sexually: The information that he or she is close kin must inhibit your sexual impulses.

Again, if you engaged in incest, and lost the baby after a few months, how would you know what caused the miscarriage? Your life is a series of many events (perhaps including sex near the time of conception with non-kin as well as kin), any one of which is a potential cause. Why conclude that sex with one individual, who physically and psychologically resembles other members of his sex in many respects, caused you to lose your baby?

The need to avoid incest implies the ability to spontaneously and automatically acquire the category "kin versus non-kin" by merely observing the world -- even if it were possible to learn it by engaging in incest, the fitness costs would be too high. But the "number of grand-offspring produced" decision rule cannot be used to acquire evolutionarily crucial categories through mere observation: unless a categorization

* Each person has, on average, four lethal equivalent genes: having only one lethal equivalent does not adversely affect your health, but individuals die when they are homozygous for one of these genes. The probability that a random individual has one of the same lethal equivalents as you is very small; however, the probability that a full sib shares a given lethal equivalent with you is 50%. If you only had one lethal equivalent, and mated with your full sib, on average, half your children would die in utero or at a very young age. As each person has about four, the selective cost is even higher. The reduced sexual recombination that attends inbreeding also imposes selective costs in long-lived species, having to do with parasite load (Tooby, 1982).

scheme is used to guide behavior, it has no consequences on fitness.

Kin recognition requires Darwinian algorithms tuned to environmental cues that are correlated with kin but not with non-kin. These cues must be used in a particular way: either they must be used to match self to other, as in facial or olfactory phenotype matching, or they must categorize others directly, as when one imprints during a critical period on those with whom one was raised -- this can be a reliable environmental cue in species where the individuals with whom one is raised are normally one's closest kin (Shepher, 1983). There are an infinite number of dimensions that could be used to carve the environment into categories; there is no assurance that a general purpose information processing system would ever hit on those useful for creating the kin/non-kin categorization scheme, and the "grandchildren produced" criterion cannot guide such a system toward the appropriate dimensions.

Then there is the problem of generalization. Suppose you were somehow able to figure out that avoiding sex with kin had positive fitness consequences. How would you generalize this knowledge about the kin/non-kin categorization scheme to other domains of human activity? Would you, for example, avoid any interaction with kin? This would be a mistake; selectively avoiding sex with kin has positive fitness consequences, but selectively avoiding helping kin has negative fitness consequences (given a certain envelope of circumstances, Hamilton, 1964).

Thus, not only must the acquisition of the kin/non-kin

categorization scheme be guided by domain specific Darwinian algorithms, but its adaptive use for guiding behavior is also domain specific. In the sexual domain, kin must be avoided; in the helping domain, they must be helped; when one needs help, kin should be among the first to be asked (Hamilton, 1964); when one is contagiously ill, kin should be selectively avoided (Tooby, 1982). The procedural knowledge governing how one behaves toward kin differs markedly from domain to domain. Only Darwinian algorithms with procedural knowledge specific to each of these domains can assure that one responds to kin in evolutionarily appropriate ways. Simply put, there is no domain general criterion of fitness that could guide an equipotential learning process toward the correct set of fit responses.

Trial and error learning is inadequate, not only because it is slow and unreliable, but because there is no domain-independent variable for signaling error. In the sexual domain, error = sex with kin. In the helping domain, error = not helping kin given the appropriate envelope of conditions. In the disease domain, error = infecting kin.

Consequently, there are only two ways the human mind can be built. Either:

1. All innate psychological mechanisms are domain general, and therefore do not track fitness at all,
- or
2. Some innate psychological mechanisms are domain specific Darwinian algorithms with procedural knowledge specialized for tracking fitness in the concrete situations hominids would have encountered as Pleistocene hunter-gatherers.

Clearly, the first alternative is no alternative at all. Advocates of this position would have to explain how genes coding for traits that impede their replication could possibly outcompete genes that code for traits that enhance their replication. In other words, they must explain how a complex of maladaptive traits was able to displace a complex of adaptive ones.

Darwinian algorithms solve the "frame problem"

Darwinian algorithms can be seen as frame-builders, as learning mechanisms that structure experience along adaptive dimensions in a given domain. Positing them solves the "frame problem" -- which is another name for the objections to domain general mechanisms that were raised in the above discussion.

Researchers in artificial intelligence have found that trial and error is a good procedure for learning only when an organism already has a well-specified model of what is likely to be true of a domain, a model that includes a definition of what counts as error. Programmers call this finding the "frame problem" (e.g., Boden, 1977; Fodor, 1983). To move an object, make the simplest induction, or solve a straightforward problem, the computer must already have a sophisticated model of the domain in question: what counts as an object or stimulus, what counts as a cause, how classes of entities and properties are related, how various actions change the situation. Unless the learning domain is severely circumscribed and the procedures highly specialized and content-dependent -- unless the programmer has given the computer what amounts to vast quantities of "innate knowledge" -- the computer can move nothing, learn nothing, solve nothing. The

frame problem is a concrete, empirical demonstration of the philosophical objections to the tabula rasa. It is also a cautionary tale for advocates of domain general, content-independent learning mechanisms.*

Unfortunately, the lesson has been lost on many. Although most cognitive psychologists realize that their theories must posit some innate cognitive architecture, a quick perusal of textbooks in the field will show that these still tend to be restricted to content-independent operating system characteristics: short term stores, domain general retrieval and storage processes, imagery buffers. Researchers who do insist on the necessity of positing content-dependent schemas or frames (e.g., Minsky, 1977; Schank & Abelson, 1977), seldom ask how these frames are built. They seem to presume that frames are the product of experience structured only by domain general learning mechanisms -- yet the building of frames must also be subject to the frame problem. Even Fodor (1983), a prominent exponent of the view that the mind's innate architecture includes specialized, content-dependent modules, restricts these to what he calls "input systems": perceptual or quasi-perceptual domains like vision, hearing, and language. He doubts the existence of modules governing "central" processes like reasoning and problem solving. Yet one wonders: Without domain specific inference processes, how can all this perceptual data be expected to guide

* Darwinian algorithms specify inference procedures, and can therefore be seen as constraining the theoretical set of all possible inferences to a few that are useful to the organism. However, they are not "constraints on learning" -- indeed, it is not clear that an organism could learn anything at all without such "constraints." See Appendix A: "The Frame Problem and So-called Constraints on Learning."

our behavior in adaptive directions?

Restricting the mind's innate architecture to perceptual systems, a content-independent operating system, a domain general concept learning mechanism, a content-independent hypothesis testing procedure, and a small ragbag of dimensions for construing similarity, might be OK if it did not matter what a person learned -- if, for example, learning that E is the most frequently used letter in the English language were as critical to one's inclusive fitness as learning that a saber-toothed tiger can eat you for lunch. But what a person learns does matter; not only what, but when, how reliably, and how quickly. And even more important is what a person does with that knowledge. The purpose of learning is, presumably, to guide behavior. Should I eat gravel? Should I engage in incest? Should I give others the only food I have for feeding my children? When my brother and my cousin are equally in need, should I satisfy those needs equally? Natural selection theory provides definite answers to questions like these, because the wrong decision can result in large fitness costs. How can an equipotential learning system that simply looks for relations in the world, provide information about the relative value, in inclusive fitness terms, of alternative courses of action? It cannot; it has no standard for assessing it.

Cognitive psychologists can persist in advocating such systems only because they are not asking what problems the mind was designed, by natural selection, to solve. The Darwinian view is that humans have innately specified mental algorithms that allow them to pursue goals that are or once were correlated with

their inclusive fitness. These innately specified mental algorithms cannot all be domain general. Behavior is a transaction between organism and environment; to be adaptive, specific behaviors must be elicited by evolutionarily appropriate environmental cues. Only specialized, domain specific Darwinian algorithms can insure that this will happen.

* * *

The proposal that Darwinian algorithms guide inference on reasoning tasks qualifies as a "family 1-b" explanation: Although we do not have a content-independent logic module, for evolutionarily important domains, we have extensive networks of "hypotheses" about what is true and what is "relevant", as well as rules of inference to guide reasoning within the domain. These hypotheses and rules are innate, or else the product of "experience" structured by domain specific innate algorithms. For these domains we do not face -- ontogenetically -- the problem of weeding out an infinite number of incorrect inductions, because this has already been accomplished phylogenetically, over 4 billion years of evolution. Hence, we spontaneously generate only a small subset of the class of all possible hypotheses, and those that we do consider are likely to be true -- or if not true, then adaptively useful. Their purpose is not merely to allow us to describe the world, but to pick up and process the information that is most salient for guiding our behavior in adaptive directions.

According to equipotential meta-theory, content is noise. According to Darwinian meta-theory, content is signal. If there are content-dependent Darwinian algorithms that guide reasoning

in evolutionarily important domains, then different content domains will "call up" different rules of reasoning. Choices on reasoning tasks should vary systematically with problem content, so long as problem content involves evolutionarily important domains. Non-important domains should show no systematic variation, because was no selection producing mechanisms sure to get such problems "right"; at best, reasoning about such areas should be weakly patterned by whatever domain general mechanisms do exist. To test this view, I chose a domain of human activity that should have been adaptively patterned by natural selection, and that appears to elicit consistent and robust content effects on a reasoning task: social exchange. Chapter 5 analyzes how the dynamics of natural selection apply to social exchange, and what this allows one to infer about the characteristics of the Darwinian algorithms that regulate social exchange in humans.