CHAPTER 1

New perspectives on development

The problem of change

Things change. When things change in a positive direction (i.e., more differentiation, more organization, and usually ensuring better outcomes), we call that change "development." This is Heinz Werner's orthogenic principle (Werner, 1948).

Ironically, in the past several decades of developmental research there has been relatively little interest in the actual mechanisms responsible for change. The evidence of surprising abilities in the newborn, coupled with results from learning theory which suggest that many important things which we do as adults are not learnable, have led many researchers to conclude that development is largely a matter of working out predetermined behaviors. Change, in this view, reduces to the mere triggering of innate knowledge.

Counterposed to this is the other extreme: Change as inductive learning. Learning, in this view, involves a copying or internalizing of behaviors which are present in the environment. "Knowledge acquisition" is understood in the literal sense. Yet this extreme view is favored by few. Not only does it fail to explain the precocious abilities of infants and their final mature states, but it also fails to provide any account of how knowledge is deposited in the environment in the first place.

The third possibility, which has been the position advocated by classic developmentalists such as Waddington and Piaget, is that change arises through the *interaction* of maturational factors, under genetic control, and the environment. The interaction at issue here is not the banal kind where black and white yield gray, but a much more challenging and interesting kind where the pathways from genotype to phenotype may be highly indirect and nonobvious. The

problem with this view in the past has been that, lacking a formal and precise theory of how such interactions might occur, talk of "emergent form" was at best vague. At worst, it reduces to hopeless mysticism.

Two recent developments, however, suggest that the view of development as an interactive process is indeed the correct one, and that a formal theory of emergent form may be within our grasp. The first development is the extraordinary progress that has been made in the neurosciences. The second has been the renascence of a computational framework which is particularly well suited to exploring these new biological discoveries via modeling.

Advances in neuroscience

The pace of research in molecular biology, genetics, embryology, brain development, and cognitive neuroscience has been breathtaking. Consider:

- Earlier theories of genes as static blueprints for body plans have given way to a radically different picture, in which genes move around, recombine with other genes at different points in development, give rise to products which bind directly to other genes (and so regulate their expression), and may even promote beneficial mutation (such that the rate of mutation may be increased under stressful conditions where change is desirable).

- Scientists have discovered how to create "designer genes." Human insulin can be produced in vats of bacteria, and caterpillar-resistant tomatoes can be grown. And plants have been created which produce biodegradable plastic!

- We now possess a complete and detailed picture of the embryology of at least one relatively complex organism (the nematode, *C. Elegans*). Scientists know, on a cell-by-cell basis, how the adult worm develops from the fertilized egg.

- Scientists have carried out ingenious plasticity experiments in which plugs of brain tissue from visual cortex (in late fetal rodents) are transplanted to sensorimotor cortex. This has led to the discovery that the old visual cortex neurons start to act like sensorimotor neurons. In other cases, researchers have shown that if information from the eyes is routed to auditory cortex early enough, regions of auditory cortex will set up retinotopic maps, and the organism will start to respond to visual stimuli based on messages going to the "borrowed" cortex. The conclusion many neuroscientists are coming to is that neocortex is basically an "organ of plasticity." Its subsequent specification and modularization appear to be an outcome of development—a result, rather than a cause.
- Although the degree of plasticity observed in the developing brain is surprising, the discovery of plasticity in adult mammals has come as an even greater surprise for those who believed in fixed and predetermined forms of neural organization. Studies have shown that somatosensory cortex will reorganize in the adult primate to reflect changes in the body surface (whether resulting from amputation or from temporary paralysis of a single digit on the hand). At first, this kind of reorganization seemed to be restricted to a very small spatial scale (a few microns at most) which suggested that a more transient local phenomenon could be responsible for the change. More recent evidence from adult animals that underwent amputation more than a decade prior to testing shows that this reorganization can extend across several centimeters of cortex. There are only two possible explanations for a finding of this kind: New wiring can be manufactured and established in the adult brain, or old patterns of connectivity can be converted (i.e., reprogrammed) to serve functions that they never served before.
- Sophisticated techniques have been developed for "eavesdropping" on brain activity with extraordinary spatial and temporal detail. Structural Magnetic Resonance Imaging (MRI), for example, provides enough spatial resolution to reveal a flea dancing on the corpus callosum (assuming there were such a flea). Evoked response potentials (ERP) gives us a temporal localiza-

tion of brain processes to within thousandths of a second. Positron emission tomography (PET), magneto-encephalography (MEG), and new functional MRI techniques provide a bridge between the precise spatial resolution of structural MRI and the fine temporal resolution of EEG, showing us which parts of the brain are most active during various cognitive tasks. Taken together, these techniques provide us with potentially powerful tools both for examining the structure and functioning of the living brain, and its development over time.

These techniques make available a range of data which were simply not accessible even a decade ago. But although some might like to believe that theory follows inevitably from data, in fact it is usually the case that data may be interpreted in more than one way. What are needed are additional constraints. These come from a second development, which is a computational framework for understanding neural systems (real or artificial).

Neural computation: the connectionist revolution

Coinciding (but not coincidentally) with the dramatic advances in neuroscience, a second dramatic event has unfolded in the realm of computational modeling. This is the re-emergence of a biologically oriented framework for understanding complex behavior: Connectionism. The connectionist paradigm has provided vivid illustrations of ways in which global behaviors may emerge out of systems which operate on the basis of purely local information. A number of simple but powerful learning algorithms have been developed which allow these networks to learn by example. What can be learned (without being prespecified) has been surprising, and has demonstrated that a great deal more information and structure is latent in the environment than has been realized. Consider:

within thousandths of a second.
hy (PET), magneto-encephalogra-
tional MRI techniques provide a
patial resolution of structural MRI
ution of EEG, showing us which
t active during various cognitive
techniques provide us with poten-
examining the structure and func-
d its development over time.

a range of data which were simply
. But although some might like to
ably from data, in fact it is usually
reted in more than one way. What
raints. These come from a second
ational framework for understand-
ial).

connectionist

ly) with the dramatic advances in
event has unfolded in the realm of
the re-emergence of a biologically
nding complex behavior: Connec-
digm has provided vivid illustra-
aviors may emerge out of systems
ely local information. A number of
lgorithms have been developed
learn by example. What can be
fied) has been surprising, and has
nore information and structure is
s been realized. Consider:

- Visual cortex in mammals is well known to include neurons which are selectively sensitive to highly specific visual inputs. These neurons include edge detectors, center-surround cells, and motion detectors. Biologically plausible network models have been constructed which demonstrate that such specialized response properties do not have to be prespecified. They emerge naturally and inevitably from cells which are initially uncommitted, simply as a function of a simple learning rule and exposure to stimulation (Linsker, 1986, 1990; Miller, Keller, & Stryker, 1989; Sereno & Sereno, 1991). These artificial networks even develop the characteristic zebra-like striped patterns seen in ocular dominance columns in real cortex (Miller, Keller, & Stryker, 1989).

- When artificial networks are trained to compute the 2-D location of an object, given as inputs the position of the stimulus on the retina and the position of the eyeballs, the networks not only learn the task but develop internal units whose response properties closely resemble those of units recorded from the parietal cortex of macaques while engaged in a similar task (Zipser & Andersen, 1988).

- Networks which are trained on tasks such as reading or verb morphology demonstrate, when "lesioned," symptoms and patterns of recovery which closely resemble the patterns of human aphasics (Farah & McClelland, 1991; Hinton & Shallice, 1991; Marchman, 1993; Martin et al., 1994; Plaut & Shallice, 1993; Seidenberg & McClelland, 1989).

- The rules of English pronunciation are complex and highly variable, and have been difficult to model with traditional Artificial Intelligence techniques. But neural networks can be taught to read out loud simply by being exposed to very large amounts of data (Sejnowski & Rosenberg, 1987).

- In learning a number of tasks, children frequently exhibit various "U-shaped" patterns of behavior; good early performance is succeeded by poorer performance, which eventually again

improves. Networks which are trained on similar tasks exhibit the same patterns of behavior (MacWhinney et al., 1989; Plunkett & Marchman, 1991, 1993; Rumelhart & McClelland, 1986).

- Children are known to go through phases in which behavior changes slowly and is resistant to new learning. At other points in time children show heightened sensitivity to examples and rapid changes in behavior. Networks exhibit similar "readiness" phenomena (McClelland, 1989).
- Networks which are trained to process encrypted text (i.e., the words are not known to the network) will spontaneously discover grammatical categories such as noun, verb, as well as semantic distinctions such as animacy, human vs. animal, edible, and breakable (Elman, 1990). A curious fact lurks here which points to the importance of a developing system: Such networks can be taught complex grammar, but only if they undergo "maturational" changes in working memory or changes over time in the input (Elman, 1993).

Our perspective

Taken together, these advances—in developmental and cognitive neuroscience on the one hand, and neural computation on the other—make it possible for us to reconsider a number of basic questions which have challenged developmentalists, from a new and different perspective:

What does it mean for something to be innate? What is the nature of the "knowledge" contained in the genome?

Why does development occur in the first place?

What are the mechanisms which drive change?

What are the shapes of change? What can we infer from the shape of change about the mechanisms of change?

re trained on similar tasks exhibit
r (MacWhinney et al., 1989; Plun-
Rumelhart & McClelland, 1986).

rough phases in which behavior
nt to new learning. At other points
ened sensitivity to examples and
Networks exhibit similar "readi-
d, 1989).

o process encrypted text (i.e., the
network) will spontaneously dis-
such as noun, verb, as well as
animacy, human vs. animal, edi-
1990). A curious fact lurks here
ce of a developing system: Such
plex grammar, but only if they
anges in working memory or
(Elman, 1993).

-in developmental and cognitive
and neural computation on the
econsider a number of basic ques-
elopmentalists, from a new and

be innate? What is the nature of the
ome?

rst place?

change?

at can we infer from the shape of
ange?

Can we talk meaningfully about "partial knowledge?"

How does the environment affect development, and how do genetic constraints interact with experience?

Our purpose in writing this book is to develop a theoretical framework for exploring the above questions and understanding how and why development occurs. We will cover a number of different specific topics in this book, but there are some central themes which recur throughout. We would like to identify these issues explicitly from the outset and foreshadow, briefly, what we will have to say about each one.

We begin with a discussion of genes. Although we are primarily concerned with behavior, and behavior is a very long way from gene expression, genes obviously play a central role in constraining outcomes. When we contemplate the issue of innateness, it is genes that we first think of. This discussion of genes will also help us to set the stage for what will be a recurring theme throughout this book: The developmental process is—from the most basic level up—essentially dependent at all times on interactions with *multiple* factors.

From genes to behavior

There is no getting around it: Human embryos are destined to end up as humans, and chimpanzee embryos as chimpanzees. Rearing one of the two in the environment of the other has only minimal effects on cross-species differences. Clearly, the constraints on developmental outcomes are enormously powerful, and they operate from the moment of conception. Furthermore, although there is a great deal of variability in brain organization between individuals, the assignment of various functions (vision, olfaction, audition, etc.) is not random. There are predictable and consistent localizations across the majority of individuals.

It is easy to state the obvious conclusion, which is that genes play the central role in determining both interspecies differences and intraspecies commonalities. This is true, but the real question is how, and what the genes are doing. Most developmentalists agree

that a preformationist version of an answer (that these outcomes are contained in an explicit way in the genome) is unlikely to be correct (although some version of preformation may come close to capturing the nature of development in certain organisms, e.g., nematodes). There is simply too much plasticity in the development of higher organisms (as we shall discuss in Chapter 5) to ignore the critical effect of experience. We know too that there aren't enough genes to encode the final form directly, and that genes don't need to code everything. So how do genes accomplish their task?

How genes do their work

Asked what genes do, most people will report the basic facts known since Mendel (although he did not use the term gene), namely, that genes are the basic units of inheritance and that genes are the critters that determine things like hair color, gender, height. Such a view of genes is not incorrect, but it is woefully incomplete, and lurking beneath this view are a number of commonly held myths about genes which are very much at odds with recent findings in molecular genetics.

For instance, according to conventional wisdom, genes are discrete in both their effects and their location. Thus, one might imagine a gene for eye color which in one form (allele) specifies blue eyes, and in another specifies brown eyes. Genes are also thought of as being discrete with regard to location. As with the memory of a computer, under this view one should be able to point to some region of a chromosome and identify the starting point and ending point of a gene (which is itself made up of a sequence of base pairs).

In fact, the reality of genes and how they function is far more complex and interesting. Consider the following.

Genes are often physically distributed in space. In eukaryotes (e.g., humans, fruitflies, and corn are eukaryotes), DNA has been found to be made up of stretches of base pairs called exons, which code for the production of proteins, but which are interrupted by sequences of noncoding base pairs called introns. In some cases, the quantity of noncoding DNA may be more than 100 times greater than the coding DNA. What happens during protein synthesis, which is how

most genes actually accomplish their work, is that the RNA copy of the gene-to-be—which includes introns—has to be cut up and respliced by specialized molecular machinery (see Figure 1.1). The

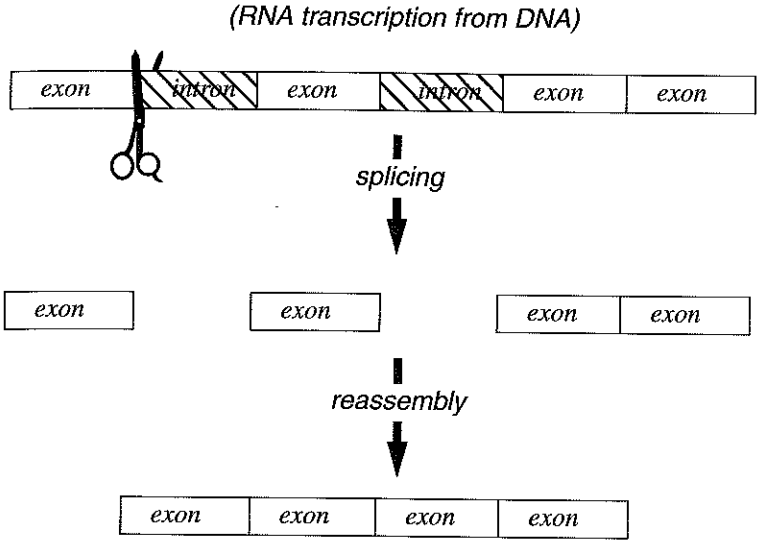


FIGURE 1.1 DNA often includes nonfunctional base pairs (*introns*) as well as sequences which code for proteins and other products (*exons*). During synthesis, the RNA transcript (but not the original DNA) is cut and spliced so that only the exons remain; the revised RNA is then used for actual synthesis.

result is a temporary "cleaned up" version of the gene transcript which can then be used for protein synthesis.

Moreover, the same portion of DNA can be spliced in different ways. For some purposes, a sequence of base pairs may be treated as an intron (noncoding), but for other purposes, the same region may be spliced to yield a different gene transcript and end up as an exon (coding). Finally, although the structure of DNA base pairs is basically stable, some sequences move around. This movement turns out to play a much more important role in genetic expression than was thought when "jumping genes" were first discovered.

Genes are not binary in their effects. What about the view of genes as discrete in their effects? This too turns out to be a misleading idea. To be sure, there are base pair sequences which code directly for specific and well-defined traits. But in many cases the encoding is continuously valued. A subsequence of base pairs may be repeated or there may be multiple copies of the gene; this causes more of the protein product to be produced and may result in a more strongly expressed trait.

Genes do their work with other genes. Sometimes, but rarely, it is possible to tie the effects of a single gene's products to some clearly defined trait. However, such "single action" genes either tend to be associated with evolutionarily primitive mechanisms or they work as switches to turn on and off some other function which is coded by a group of genes.

For example, the fruitfly, *Drosophila melanogaster*, has a gene called Antennapaedia (Antp). If the Antp gene undergoes a certain mutation, then instead of antennae the fruitfly will develop an extra pair of feet growing out of its head where the antennae would be. Notice that this bizarre effect relies on the fact that what the Antp gene does is to regulate the expression of other gene complexes which actually produce the feet (or antennae). Even simple traits such as eye color in the fruitfly may depend on joint action of 13 or more genes. Thus, while there are single-action genes, more typical are cases where multiple genes are involved in producing any given trait, with some genes playing the role of switches which control and regulate the expression of other genes.

Genes are often reused for different purposes. A very large number of genes in an animal's genome are what one might call "housekeeping genes." They code for the production of basic proteins which function as enzymes, form cellular organelles, carry out cellular metabolic activities, etc.

But Nature is stingy with her solutions. Things which work in one species frequently turn up in very distantly related species. All together, probably something like 5,000 genes are needed by cells in all eukaryotes for housekeeping purposes. Essentially the same genes, modified here and there, are shared by all species and cell types. The lesson here is that there is remarkable conservatism and

reusing of old solutions. By rearranging and slightly modifying only a few thousand interacting gene complexes, enormous diversity of structure is possible.

This conservatism does not rule out the abrupt appearance of what seem to be radically new structures, be they language or flippers or wings. There is a great deal of genetic redundancy in eukaryotes. The same gene may appear many times in the genome, and often slightly different genes produce similar or identical products. This redundancy accommodates many small changes in the genome before there is a dramatic shift in phenotype. Thus the appearance of abrupt changes in phenotypic outcomes may be misleading, and result from much tinier changes at the genetic level. This brings us to the next point.

The relationship between genome and phenotype is highly nonlinear. Although a linear increase in genome size (measured as the number of DNA base pairs) which correlates with phenotypic size can be observed for simple species (e.g., worms), this does not hold for so-called higher species (see Table 1.1). In the latter case, the relationship is highly nonlinear. In Chapter 4 we will discuss nonlinear phenomena in some detail. For the moment suffice it to note that one of the most dramatic nonlinear relationships in nature is that which exists between the genome and the phenotype.

Compare, for example, the genome of the chimpanzee, the Old World monkey, and the human. To the layman's (admittedly biased) eye, the Old World monkey and the chimp resemble each other much more closely than either species resembles us. Yet genetically the chimp and the human are almost indistinguishable: We have 98.4% of our genetic material in common, compared with only approximately 93% shared by the chimp and Old World monkey. Humans are also closer to chimps, genetically, than chimps are to gorillas. Whatever differences there are between us and the chimp therefore come down to the effects of the 1.6% difference.

In Chapter 7, we will discuss the implications of the above facts for what it might mean for a trait or a behavior to be innate. For the moment, the above—which reveals only the most modest glimpse of the complexity which underlies genetic functioning—is enough to help us make a simple point. Even the simplest questions, *what* a

TABLE 1.1

Organism	Amount of DNA (base pairs) 10^6 10^7 10^8 10^9 10^{10} 10^{11}
Flowering plants	
Birds	
Mammals	
Reptiles	
Amphibians	
Bony fish	
Cartilaginous fish	
Echinoderms	
Crustaceans	
Insects	
Mollusks	
Worms	
Molds	
Algae	
Fungi	
Gram-pos. bacteria	
Gram-neg. bacteria	
Mycoplasma	

Adapted from Edelman (1988). For simpler species (e.g., mycoplasma through worms), there is an approximately linear increase in DNA with increasing organism size. For more complex species (the upper portion of the table), there is no obvious relationship between amount of DNA and the size or complexity of the organism.

gene is, and *where* it resides, have answers which are very much at variance with the commonplace notion of the "digital gene." Genes are fluid and graded, which gives them a distinctly analog character, and they rarely work in isolation. Genes work in concert with large numbers of other genes, and tracing a particular gene's contribution to the emerging phenotype is very indirect and rarely possi-

ble without considering the whole network of interactions in which that gene participates.

Let us turn now to a slightly higher stage of organization from genes, namely, cells and tissues. We shall see that the same sorts of interactions which are observed in genes occur in cell differentiation and tissue formation.

How cells come to be

The human body contains roughly 100 trillion cells; these are made up of about 200 different cell types, which are more or less the same types as are found in snakes, birds, and other mammals. The major difference between species lies in changes in cell number and topology. Some species have more cells than others, and those cells are arranged in different ways.

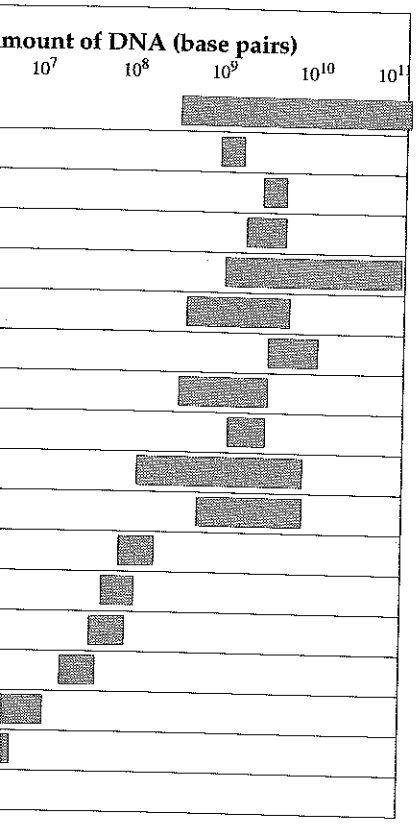
However, there is a mystery here. All cells (with one or two minor exceptions) contain the same genetic information. So the question we must ask is how cells come to be different. This process of differentiation occurs early in development. While the process is far from fully understood, enough is known for us to give examples which illustrate the importance of interactions.

Mosaics and regulators. Not all species rely on interactions to the same degree. Embryologists distinguish between two very different styles of cellular development: Mosaic development and regulatory development.

In mosaic development, cells develop more or less independently. They tend to be largely unaffected by each other and by the environment. The fate of each cell or group of cells is determined early on by their location. When and what a cell becomes is under relatively tight genetic control.

A mosaic organism that has been particularly well-studied is the nematode, *C. Elegans*. This worm has been the subject of a long-term research project, resulting in a largely complete picture of its embryological development from zygote to adult. *C. Elegans* makes a good subject for study for several reasons. It develops very quickly (about two days to reproductive maturity). It is also reasonably complex. Further, its body is transparent; this makes it possible

ABLE 1.1



simpler species (e.g., mycoplasma
mately linear increase in DNA with
complex species (the upper portion of
nship between amount of DNA and the

ve answers which are very much at
notion of the "digital gene." Genes
es them a distinctly analog charac-
ation. Genes work in concert with
d tracing a particular gene's contri-
be is very indirect and rarely possi-

to view internal development with a light microscope. Finally, each member of the species is virtually identical. Every *C. Elegans* contains exactly 959 somatic cells, and cell connectivity is very regular across different individuals of the species.

The path by which each *C. Elegans* arrives at adulthood is both highly circuitous but also highly invariant. Cells differentiate in ways that seem neither tidy nor logical. Organs and tissues are formed from what seem like random groupings, and cell lineage is quite eccentric. But as byzantine as the process may seem, it also seems to be largely predetermined. Genetically identical individuals have essentially the same cell morphology, cell position, and cell connectivity (by comparison, this is not at all the case for human monozygotic twins). By and large, cell lineage is not influenced by the fate of other cells; molecular level interactions are sufficient to determine its fate.

It is easy to see the advantages of the mosaic style of development. By largely prespecifying cell fate, nature ensures that evolutionarily tested solutions are well preserved. Mosaic development is independent and fast. Each cell has its own timetable and need not wait for other cells to develop. Growth can occur just as quickly as cells are able to consolidate the materials they need to divide and differentiate. Each cell does this on its own, confident that when its job is done, its fellows will be ready and waiting in place to join up. If humans developed as mosaics we could be up and running in just a short time after conception because the approximately 100 trillion cells could be produced in just under 47 binary divisions.

In fact, the human species—and most other higher organisms—have not opted for the mosaic style of development, but instead rely on what is called regulatory development. Why should this be? What are the limitations of mosaic development?

The problems with mosaic development are the flip side to their advantages. First, lack of interaction at the cellular level permits speedy development under good conditions, but it also means lack of flexibility in the face of abnormal conditions. If some cells are damaged, others cannot compensate for their absence and the organism may no longer be viable. In environments which are

unstable or likely to change, there may be no single developmental solution which can be hardwired into the genome.

There is another, and probably more serious, price which is paid by mosaic organisms. The burden on the mosaic genome is considerable. The genome comes close to being a blueprint for the body; it must specify everything. For relatively small and simple organisms—up to, say, the worm—this may not be much of a problem, and may be compensated by the other advantages of mosaic organization. However, the blueprint approach to development puts an upper bound on the complexity which can be achieved. Such direct specification of the human brain alone, for example, could plausibly require something on the order of 10 trillion base pairs of DNA, which is far in excess of what is structurally feasible. Indeed, in many organisms the relationship between the amount of a species' genetic material and its morphological and behavioral complexity is highly nonlinear. Indeed, many plants have more genetic material than do humans (recall Table 1.1).

The alternative to mosaic development is regulatory development. Regulatory systems rely heavily on cellular-level interactions. The orchestration of cell differentiation and the final outcome are under broad genetic control, but the precise pathway to adulthood reflects numerous interactions at the cellular level that occur *during development*.

While most species show some elements of both types of development (see Edelman, 1988), higher vertebrates generally show more regulatory development. The effects of regulatory development may be quite dramatic. Earlier we pointed out that the average human and chimpanzee DNA differ only by 1.6%, which is less than the difference between two species of gibbons. Yet the morphologies and behaviors of humans and chimps differ considerably. It seems reasonable to believe that these differences depend far more on the evolution of regulatory mechanisms than on the evolution of new structural genes.

One advantage of regulatory development is that it allows for greater flexibility. Damage to a group of cells can often be compensated for by their neighbors. More significantly, regulatory systems probably permit far greater complexity of phenotypes than can be

achieved in mosaic developmental systems. The cellular level interactions provide an extra source of constraint which makes it possible to develop more complex phenotypes.

However, such interactions impose their own cost. Regulatory systems require a period of interdependent development, and this may slow down the process of development since some events will require the completion of other events before they can start. The organism is likely to have a prolonged period of immaturity during which time it is vulnerable. Mosaics are simple but fast to develop; regulators are complex and slow to develop. When the phenotype is relatively simple, genetic and molecular level information are sufficient to allow parallel development of a large number of cells. This is the mosaic approach. But there are constraints on the amount of genetic material which can be safely housed in a cell and reliably replicated across generations. In order to attain a more complex outcome (phenotype) with a roughly similar number of genes, it is necessary to create hierarchical intermediate stages, which, in the case of regulatory systems, occur at the cellular level of interaction.

It is for this reason that developmental timing becomes more crucial as the hierarchical complexity of an ontogenetic system increases. And genes are algorithms which operate sequentially in time. In Chapter 6, we discuss just how timing can be controlled, and how it can be exploited in the service of building complexity. Thus the action of genes—in particular, those associated with the development of more complex behaviors—may be very indirect. The genome is, first, algorithmic rather than descriptive. The algorithms often rely on predictable regularities in the input (so that the algorithms do not need to encode information which can be counted on to be made available through experience). Two of the lessons we have learned from connectionist research are, first, that considerably more information may be latent in the environment—and extractable, using simple learning algorithms—than was previously thought; and second, that useful self-organization may occur in the absence of explicit guidance from the environment. Certain problems have a natural solution; all that may be required are a few gentle nudges in the form of prewired biases and constraints.

tal systems. The cellular level inter-
of constraint which makes it possi-
enotypes.

impose their own cost. Regulatory
rdependent development, and this
evelopment since some events will
events before they can start. The
onged period of immaturity during
aics are simple but fast to develop;
to develop. When the phenotype is
olecular level information are suffi-
ent of a large number of cells. This
e are constraints on the amount of
afely housed in a cell and reliably
n order to attain a more complex
ghly similar number of genes, it is
intermediate stages, which, in the
at the cellular level of interaction.

elopmental timing becomes more
plexity of an ontogenetic system
ams which operate sequentially in
ust how timing can be controlled,
ne service of building complexity.
rticular, those associated with the
ehaviors—may be very indirect.
rather than descriptive. The algo-
egularities in the input (so that the
ode information which can be
through experience). Two of the
nnectionist research are, first, that
ay be latent in the environment—
ning algorithms—than was previ-
seful self-organization may occur
ce from the environment. Certain
all that may be required are a few
ired biases and constraints.

The problem of interaction

We have taken some time talking about genes and cells for two reasons. First, as we already pointed out, genes are what most people think of when they think of innateness. And second, even a cursory discussion reveals how much genes and cells depend on interactions.

This is not a new insight. Developmentalists have long acknowledged the role of interaction. The problem has been that these interactions are either so trivial as to be of little interest, or so complex as to resist analysis. So the interactionist position—while in principle agreed upon by virtually every developmentalist—remains a difficult one to pursue in practice.

Several decades ago, Waddington tried to illustrate this conception of development with his picture of the "epigenetic landscape" (see Figure 1.2, left panel). Embryologists knew that phenotypically very similar individuals might have wildly different genotypes; and an organism with a single (apparent) phenotype might emerge from a genome that contains a much larger array of possibilities than are ever realized. How could this be? Waddington offered the following account:

> *I envisage [development] as a set of branching valleys in a multidimensional space that includes a time dimension, along which the values extend. The development of the phenotype of an individual proceeds along a valley bottom; variations of genotype, or of epigenetic environment, may push the course of development away from the valley floor, up the neighboring hillside, but there will be a tendency for the process to find its way back.* (Waddington, 1975; p. 258)

Waddington's image of the epigenetic landscape and his account of how development proceeds bear, at one level of description, an eery resemblance to another familiar image—the way neural networks change over time. In the case of the network, the "environment" is usually a training set; the "phenotype" is the evolving set of connection strengths between artificial synapses (Figure 1.2, right panel). Over time, the network seeks the "low ground" of error in its weight space, but there are often many paths to equivalent solutions.

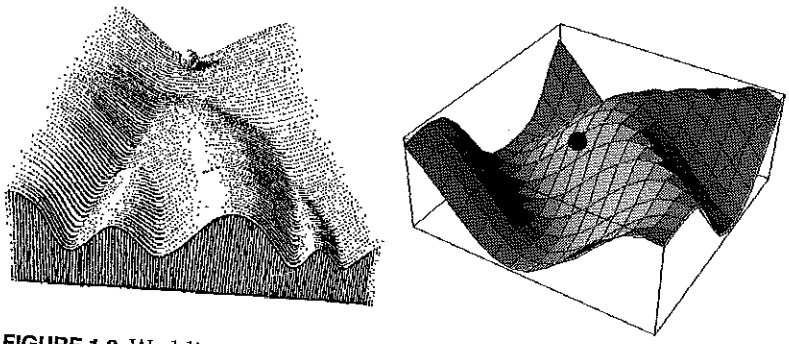


FIGURE 1.2 Waddington's epigenetic landscape on the left (Waddington, 1975); hypothetical error surface from a neural network, right.

We suspect that developmentalists would have little difficulty recognizing the image on the right as just another epigenetic landscape; while connectionists would look at Waddington's drawing and immediately think of neural network error surfaces. Is this mere coincidence?

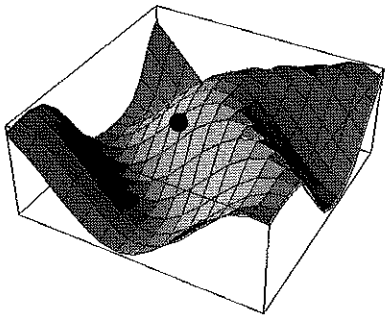
We believe not. In fact, we shall argue in this book that the two images are more than coincidentally similar. We believe that these two images capture what is fundamentally the same process.

Taking a biological perspective

Our perspective is connectionist, but it is equally a biological perspective. Since our views probably differ from those of some connectionists, as well as of some biologists, we should explain just what our brand of "biologically oriented connectionism" entails.

(1) *We think it is crucial to pay attention to what is known about the genetic basis for behavior and about developmental neuroscience.*

Remarkable strides have been made recently in understanding the relationship between gene expression and development. If we want to know in what senses behavior might be innate, we should strive for an account which is broadly consistent with what is known about what genes do and how they work in other domains.



ıetic landscape on the left (Waddington,
from a neural network, right.

ntalists would have little difficulty
ght as just another epigenetic land-
uld look at Waddington's drawing
al network error surfaces. Is this

nall argue in this book that the two
tally similar. We believe that these
amentally the same process.

tive

, but it is equally a biological per-
ɔly differ from those of some con-
biologists, we should explain just
ɔriented connectionism" entails.

attention to what is known about the
developmental neuroscience.

n made recently in understanding
xpression and development. If we
navior might be innate, we should
broadly consistent with what is
how they work in other domains.

As we shall argue in the final chapter, the way genes work probably precludes anything like "genes for language."

Through imaging, animal models, and focal lesion data, we now know a great deal more about how the brain develops. There is good evidence that neocortex is initially equipotent (or at least, multipotent, since some outcomes are more likely than others but many outcomes are possible). In Chapter 5 we review the dramatic results of experiments involving transplanting and rewiring bits of cortex to other regions. The results of these experiments argue strongly against solutions which depend upon innately specified populations of neurons prewired for complex cognitive functions such as language.

(2) *At the level of computation and modeling, we believe it is important to understand the sorts of computations that can plausibly be carried out in neural systems.* While there is an inevitable (and desirable) tension between the abstract models and the specific systems they attempt to model, there are basic principles which should be observed.

Computation is distributed; but the information available to the computing elements is mostly local. (There are, in fact, global regulators and chemical gradients which extend over large spatial regions; but these are typically low-dimensional sources of information, and the same regulators, e.g., morphogens which trigger cell differentiation, are frequently used across systems and even across species.)

Information also is distributed; representations are graded, continuous, and spread over wide areas. Moreover, representations are often superposed, such that the same tissue participates in representing multiple pieces of information.

Finally, neural computation is often highly nonlinear. This means that under certain circumstances, small differences in input may have little effect; whereas under other circumstances, small differences in input produce qualitatively different behaviors. This is discussed in Chapter 4.

(3) *We take a broad view of biology which includes concern for the evolutionary basis of behavior.* This approach can be traced back to classical traditions in development, as articulated by pioneers like

Baldwin, Piaget, and Vygotsky. Just as it is hard to imagine studying behavior from snapshots frozen in time, it makes no sense to try to understand development without taking into account the environment within which development unfolds or the evolutionary history which gives rise to the behavior.

The study of evolutionary change is important in itself because it can give insight into the mechanisms which underlie complex behaviors. If one is interested in innate constraints, then it is important to understand possible origins for those constraints. Furthermore, if one is interested in development, evolutionary change is important because it interacts with individual development. In fact, it can be plausibly argued that individual development cannot be fully understood without understanding its evolutionary basis. An evolutionary perspective examines (through reconstruction or simulation) the process by which forms and behaviors become constrained, either by virtue of initial starting point or channeled development.

(4) *Finally, a broader biological perspective emphasizes the adaptive aspects of behaviors, and recognizes that to understand adaptation requires attention to the environment.* An important goal for simulation studies, therefore, is taking an ecological view in which behaviors are situated in context and carried out in concert. We see this as an important counterweight to the reductionist approach, which has tended to fractionate the study of behavior. Studying cells (to paraphrase David Marr) may be useful for understanding life; but understanding how a cell works will not tell us what it means to be alive.

What does it mean to be innate?

This is a question we pose now, and consider again in the final chapter. The term "innate" has a very checkered history in science. In some fields, such as ethology, the term has been virtually banned from use for the past 20 years. One of the main reasons why this has happened is that ethological studies revealed that most of the examples of behavior originally claimed to be innate (by Lorenz and his collaborators) in fact turn out to be dependent on prior interactions

Just as it is hard to imagine studying
n in time, it makes no sense to try to
out taking into account the environ-
ent unfolds or the evolutionary his-
avior.

change is important in itself because
echanisms which underlie complex
innate constraints, then it is impor-
gins for those constraints. Further-
velopment, evolutionary change is
ith individual development. In fact,
individual development cannot be
standing its evolutionary basis. An
nes (through reconstruction or sim-
forms and behaviors become con-
nitial starting point or channeled

l perspective emphasizes the adaptive
izes that to understand adaptation
ent. An important goal for simula-
an ecological view in which behav-
arried out in concert. We see this as
the reductionist approach, which
dy of behavior. Studying cells (to
useful for understanding life; but
will not tell us what it means to be

ate?

, and consider again in the final
very checkered history in science.
ne term has been virtually banned
e of the main reasons why this has
es revealed that most of the exam-
d to be innate (by Lorenz and his
e dependent on prior interactions

with the pre- or postnatal environment. For similar reasons the term has dropped from use in genetics. Since it has become evident that genes interact with their environment at all levels, including the molecular, there is virtually no interesting aspect of development that is strictly "genetic," at least in the sense that it is exclusively a product of information contained within the genes.

Nonetheless, many cognitive scientists and developmentalists have continued to use the term innate and to speak of such things as the "language instinct." We feel this reflects the entirely justifiable desire to understand how behaviors come to be; and in cases where outcomes seem inevitable, it is tempting to call these outcomes "innate." So what is wrong with this?

The problems, from our perspective, are two-fold and have to do with *mechanism* on the one hand, and *content* on the other.

First, calling a behavior innate does very little to explain the mechanisms by which that behavior comes to be inevitable. So there is little explanatory power to the term. If all that is meant by saying a behavior is innate is that it is (under normal circumstances) inevitable, then we have gained little.

What is often meant by innate is somewhat stronger: "that which is specified within the genome," under the assumption that genes code for innate behaviors. From this perspective the challenge is to elucidate what aspects of cognition or behavior, if any, are the direct result of genetic information. But as we have already argued and will repeatedly stress throughout this book, there are a multitude of molecular, cellular and system interactions that occur between the gene and its developmental products. While aware of a great many interactions at the molecular and cellular level, some developmental psychologists choose to think of these as biological details of minimal relevance to those interested in behavior and cognition. It is a reasonable approximation, they argue, to state that the component of cognition in question is coded for in the genes. As we hope to demonstrate in this book, however, this is not a position which leads to insightful accounts of development.

The second way in which in which claims regarding innateness can be problematic has to do with the *content* of what it is that is presumed to be innate. Does the fact that the vast majority of

humans end up speaking (or signing) some language mean that language is innate? Possibly, but universal outcomes are not sufficient diagnostics of innate mechanisms (since the vast majority of humans living in the British Isles end up speaking English—yet their genomes obviously do not encode for one particular language). Does the fact that some individuals have difficulty learning a language at all mean that the deficit lies with an innate faculty for language learning? Or does the deficit arise more indirectly from some other problem which has a deleterious consequence for language learning?

These two problems are not easily dealt with, but lacking a more precise specification of the mechanisms which constrain development, and of the content domains over which they operate, any use of the term innate is bound to be muddled and counterproductive. To this end, we propose in the next two sections first, a way of thinking about possible mechanisms by which developmental outcomes might be constrained; and second, a way of thinking about the domain specificity of those mechanisms.

An alternative definition is to reserve the term innate to refer to developmental outcomes that are more or less inevitable for a given species. That is, given the normal environment for a species, outcomes which are more or less invariant between individuals are innate.

In considering these issues, Johnson and Morton (1991) suggest that it is useful to distinguish between the various levels of interaction between genes and their environment. Some of these are shown in Table 1.2. Here, the term innate refers to changes that arise as a result of interactions that occur within the organism itself during ontogeny. That is, interactions between the genes and their molecular and cellular environments without recourse to information from outside the organism. We adopt this working definition of the term in this book. Interactions between the organism and aspects of the external environment that are common to all members of the species, the species-typical environment, (such as patterned light, gravity, etc.) were referred to as "primal" by Johnson and Morton. Clearly, the boundary between innate and primal is often difficult to draw, and there are many instances from ethological studies where

behaviors thought to be innate were found to be primal (i.e., requiring interaction at the level of the species-typical environment) on closer study.

TABLE 1.2

Level of interaction	Environment	Outcome
molecular	internal environment	INNATE
cellular	internal environment	
organism-external	species-typical environment	PRIMAL
organism-external	individual environment	LEARNING

In this book we use the term innate in the same sense as Johnson and Morton (1991) to refer to putative aspects of brain structure, cognition or behavior that are the product of interactions internal to the organism. We note that this usage of the term does not correspond, even in an approximate sense, to genetic or coded in the genes.

Ways to be innate: A framework

This brings us to a crucial question. If development truly is an interactive process, and if emergent form is the rule rather than the exception, how do we keep from lapsing into the hopeless mysticism which is too often invoked in place of real explanation? Can we be more specific about the nature of the interactions and about the things which interact? Put more generally, what are the ways in which behaviors can be innate?

We suggest that it is useful to think about development as a process which occurs on multiple levels (we use the word levels here in a heterarchical rather than hierarchical sense). For an outcome to be innate, in our terms, means that development is constrained at one or more of these levels. Interactions may occur within and also across levels. And outcomes which are observed at one level may be produced by constraints which occur at another.

Given our perspective, we have also found it useful to consider how these constraints might be implemented both in natural sys-

tems (brains) as well as in the artificial systems (networks) we use to model those natural systems. Sometimes the correspondence is ready and obvious, but there are also cases where we do not yet have a clear understanding of how a constraint might be implemented. This tension is useful, we feel.

Two words of warning. First, we recognize that Nature has no particular obligation to honor any taxonomy, so we expect to find cases where our distinctions are blurred. The primary value of this taxonomy is as a conceptual framework for formulating hypotheses about where the major determinants of cognitive behavior may lie.

Second, much of what follows necessarily makes reference to terms and concepts having to do with brains and networks. Some of these terms may not be clear to the reader, but will be defined later in the book. (Chapter 2 describes network terms, and Chapter 5 deals with brain organization; the terms themselves are also listed in the Index.) We hope this will not cause the reader undue confusion. Because the taxonomy we propose is basic to our perspective, we felt it important to present it up front. We also hope that in so doing, we will have provided a framework which will assist the reader in understanding all that follows.

So, what are the ways in which things can be innate?

We propose that constraints may operate at three different levels: **representations**, **architectures**, and **timing**. Again, we emphasize that we take these to be heterarchical rather than hierarchical levels. Typically, behaviors reflect constraints which operate at multiple interacting levels. The levels themselves may be distinguished by the degree of specificity and directness of consequence for behavior and knowledge. Representational constraints have the most specific and direct relationship to knowledge; architectural constraints operate at a more general level with less directly obvious relationship to resulting knowledge; and timing constraints are typically the most opaque with regard to outcome. Let us consider each of these possibilities. (The reader may wish to refer to Table 1.3 on page 35 for a summary of these three types of constraints.)

(1) *Representational constraints*

One of the strongest hypotheses one might advance is that knowledge and behaviors are innate by virtue of the representations which subserve them being "hard-wired" in advance. For example, it has been argued that children are born with innate knowledge about basic principles of (for example) grammar (Crain, 1992; Lightfoot, 1989; Pinker, 1994a,b; Pinker & Bloom, 1990), physics (Spelke, 1994) or mathematics (Wynn, 1992). (We discuss these claims in greater detail in Chapter 3, and again in Chapter 7.) To be sure, these authors do not argue for a simple one-to-one relationship between genes and behavior; the knowledge is taken to be shaped by experience to some extent (perhaps in the form of "triggering" or "selecting" among predetermined options, e.g., Piatelli-Palmarini, 1989); and some maturation may have to take place before the innate knowledge can be implemented in the service of some behavior (e.g., Borer & Wexler, 1987; Spelke et al., 1992). However, most of these investigators have been clear in their belief that children are born with domain-specific representations laid out somewhere in the brain.

What might this mean, in network terms and in brain terms?

In a connectionist network, representations are patterns of activations across a pool of neuron-like processing units. The form of these activation patterns is determined by the nature of the connections between the units. Thus, innate representational knowledge—by which we mean the potential to produce representations of specific sorts—would take the form of prespecified weights on the inter-unit connections.

In the brain, the most likely neural implementation for such innate knowledge would have to be in the form of fine-grained patterns of synaptic connectivity at the cortical level, i.e., cortical micro-circuitry. To the best of our knowledge at the present time, this is how the brain stores its representations, whether they are innate or acquired. In this regard, Pinker (1994b) suggests that the "language instinct" is indeed based upon specific microcircuitry, and that the same is probably true for many other cognitive processes:

> *It is a certain wiring of the microcircuitry that is essential.... If language, the quintessential higher cognitive process, is an instinct, maybe the rest of cognition is a bunch of instincts too—complex circuits designed by natural selection, each dedicated to solving a particular family of computational problems posed by the ways of life we adopted millions of years ago.* (Pinker, 1994b; pp. 93, 97)

Assuming that representations are defined in terms of cortical microcircuitry, what might it mean to say that knowledge and/or representations are innate? Although it is theoretically possible to set the weights of a network (natural or artificial) in advance through natural selection, we will argue that representational innateness (so defined) is relatively rare in higher organisms, at least at the cortical level (for some possibilities at the subcortical level, see Chapter 6). Indeed, there are many reasons to think that the cortex in higher vertebrates (especially humans) has evolved as an "organ of plasticity" which is capable of encoding a vast array of representational types.

In fact, as we shall see in some detail in Chapter 5, evidence has been mounting against the notion of innate domain-specific microcircuitry as a viable account of cortical development, i.e., against what we call "representational nativism."

In a number of recent studies with vertebrate animals, investigators have changed the nature of the input received by a specific area of cortex, either by transplanting plugs of fetal cortex from one area to another (e.g., somatosensory to visual, or vice-versa, O'Leary, 1993; O'Leary & Stanfield, 1989), by radically altering the nature of the input by deforming the sensory surface (Friedlander, Martin & Wassenhove-McCarthy, 1991; Killackey et al., 1994), or by redirecting inputs from their intended target to an unexpected area (e.g., redirecting visual inputs to auditory cortex (Frost, 1982, 1990; Pallas & Sur, 1993; Roe et al., 1990; Sur, Garraghty & Roe, 1988; Sur, Pallas & Roe, 1990; see also Molnar & Blakemore, 1991).

Surprisingly, under these aberrant conditions, the fetal cortex takes on neuroanatomical and physiological properties that are appropriate for the information it receives ("When in Rome, do as the Romans do..."), and quite different from the properties that would have emerged if the default inputs for that region had occurred. This suggests that cortex has far more representational

plasticity than previously believed. Indeed, recent studies have shown that cortex retains representational plasticity into adulthood (e.g., radical remapping of somatosensory cortex after amputation, in humans and in infrahuman primates (Merzenich et al., 1988; Pons et al., 1991; Ramachandran, 1993; see also Greenough, Black, & Wallace, 1993).

One cannot entirely rule out the possibility that neurons are born "knowing" what kinds of representations they are destined to take on, but right now the case for innate representations does not look very good. However, this does not mean that there is no case for innate constraints on higher cognitive processes. Instead, it means that we have to search for other ways that genes might operate to ensure species-specific forms of brain organization, and the thoughts and behaviors mediated by that form of brain organization—which brings us to the next two sources of constraint on development.

(2) *Architectural constraints*

Although there is no good evidence that we know of that knowledge and behavior are constrained at the level of representations, it is much more likely that such constraints operate at the level of architectures. The architecture of a system encompasses a number of different characteristics. Thus, the term architecture is potentially ambiguous, meaning different things to different people. What is true of all the features we consider architectural, however, is that they operate at a higher level of granularity than representational constraints, which take the form of prespecified connections between individual neurons or nodes.

Architectural constraints can vary along a number of dimensions and degrees of granularity (from cell types to overall brain structure), but in general fall into three broad categories. We call these **unit-based architectural constraints**, **local architectural constraints**, and **global architectural constraints**. We recognize that these subdivisions do not perfectly classify all types of architectural features. This is simply a provisional attempt to impose some order

on what is in reality a very complex and somewhat disordered system.

(a) *Unit level architectural constraints.* The lowest level of architecture deals with the specific properties of neurons (in brains) or nodes (in connectionist networks). This is the level that many neuroscientists would consider cytoarchitectural (but note that cytoarchitecture is often used in the neuroscience literature to refer to higher-level, areal organization as well).

In the brain, unit level constraints include the specification of neuron types which are found in different regions of the brain; response characteristics of neurons, including firing threshold, refractory period, etc.; type of transmitter produced (and whether it is excitatory or inhibitory); nature of pre- and postsynaptic changes (i.e., learning), etc. In network terms, unit level constraints might be realized through node activation functions, learning rules, temperature, momentum, etc.

It is clear that unit level constraints operate in brain development. There are a relatively small number of neuron types, for instance, and they are neither randomly nor homogeneously distributed throughout the brain. The unit level constraints are fundamental to brain organization, since they concern the lowest level of computation in the brain.

(b) *Local architectural constraints.* At the next higher level of granularity we have local architectural constraints. In brains, such constraints might take of the form of differences in the number of layers (e.g., the six-layered organization of cortex), packing density of cells, types of neurons, degree of interconnectivity ("fan in" and "fan out"), and nature of interconnectivity (inhibitory vs. excitatory). In network terms, local architectural differences would include feedforward vs. recurrent networks, or the layering of networks. (However, one should resist the obvious temptation to associate layers of cortex with layers in a network. We feel this would be a mistake, since the layers have traditionally served different purposes in the two domains.) In reality, current connectionist models have tended not to exploit differences in local architecture in a manner which comes anywhere near to the diversity observed in real brains.

Interestingly, the cortex itself initially appears to display relatively little in the way of local architectural differences. That is, the local architecture of the cortical mantle does differ significantly from other regions of the brain, but there is good reason to believe that initially this architectural plan is mostly the same through the cortex.

At the same time, it is also true that the adult cortex displays much greater differences in local architecture. Thus, and not surprisingly, experience and learning can significantly alter the local cortical architecture over time. What is at present unclear and controversial is the extent to which such differences are predetermined or emerge as a consequence of postnatal development. One likely candidate for possible intrinsic differentiation within the cortex is the thickening of cortex in the primary visual area (arising primarily from an enlarged layer 4, which is known to contain an unusually large number of neurons). More recently, work with mouse embryos suggests that while certain aspects of areal identity of cortical tissue can be changed by early transplant (e.g., connectivity and some cytoarchitectural features; see Chapter 5), there are other subtle characteristics of cells which may not be alterable and thus be intrinsic to that region of cortex (Cohen-Tannoudji, Babinet, & Wassef, 1994). The computational consequences of such regional differences are as yet unknown, but it has been suggested that (among other things) this form of innate constraint gives rise to the left/right-hemisphere differences in language processing that emerge in human beings under default conditions (i.e., in the absence of focal brain lesions; see Chapter 5 for details).

(c) *Global architectural constraints.* Finally, a major source of constraint arises in the way in which the various pieces of a system—be it brain or network—are connected together. Local architecture deals with the ways in which the low-level circuitry is laid out; global architecture deals with the connections at the macro level between areas and regions, and especially with the inputs and outputs to subsystems. If one thinks of the brain as a network of networks, global architectural constraints concern the manner in which these networks are interconnected.

In this form of nativism, knowledge is not innate, but the overall structure of the network (or subparts of that network) constrains or determines the kinds of information that can be received, and hence the kinds of problems that can be solved and the kinds of representations that can subsequently be stored. In other words, the macrocircuitry—meaning principally the areal patterns of input/output mappings—may be prespecified even if the microcircuitry is not.

In brain terms, such constraints could be expressed in terms of (e.g., thalamo-cortical) pathways which control where sensory afferents project to, and where efferents originate. Very few network models employ architectures for which this sort of constraint is relevant (since it presupposes a level of architectural complexity which goes beyond most current modeling). One might imagine, however, networks which are loosely connected, such that they function somewhat modularly but communicate via input/output channels. If the pattern of inter-network connections were prespecified, this would constitute an example of a global architectural constraint.

As we noted in the previous section, the representations developed by specific cortical regions are strongly determined by the input they receive. On this assumption, one good way to ensure that a region of cortex will be specialized for (say) vision, audition or language would be to guarantee that it receives a particular kind of information (e.g., that visual cortex receives its information from the retina, and auditory cortex receives its information from the ear). In addition, one might guarantee further specialization by placing more constraints on the input that a particular region receives. For instance, the differential projection of dopaminergic fibers to the frontal cortex from the substantia nigra and ventral temgental nuclei may provide constraints on what types of representations emerge in this part of cortex, since dopamine levels are thought to influence the firing thresholds of neurons.

Before continuing with the third and final source of constraint, we pause to note that although it is rarely acknowledged, architectural constraints of one kind or another are necessarily found in all connectionist networks. Specific choices are made regarding computational properties of individual units (unit level architectural

constraints), and although local architectural assumptions are usually very simple and uniform throughout a network (for example, a three-layer, feedforward network in which all input units report to all units in the hidden layer, and each unit in the hidden layer is connected with all the units in the output) assumptions of one kind or another are made. These initial architectural characteristics strongly constrain the behavior of the network and are almost always critical to the success (or failure) of the model in question. (When we consider all the variations of input/output relations and cytoarchitectures that are employed in real brains and compare those with the limited set of options used in current modeling efforts, we may wonder why such simulations ever succeed at all.) The point we stress, therefore, is that almost all connectionist models assume innate architectural constraints, and very few assume innate representations.

(3) *Chronotopic constraints*

One of the crucial ways in which behavioral outcomes can be constrained is through the timing of events in the developmental process. Indeed, as Gould (and many other evolutionary biologists) has argued eloquently, changes in the developmental schedule play a critical role in evolutionary change (Gould 1977; see also McKinney & McNamara, 1991).

In networks, timing can be manipulated through exogenous means, such as control of when certain inputs are presented. Or timing can arise endogenously, as in the Marchman simulation (discussed below). In Marchman's simulation, the gradual loss of plasticity in a network comes about as a result of learning itself. In brains, very occasionally, timing is under direct genetic control. More commonly, the control of timing is highly indirect and the result of multiple interactions. Hence the onset and sequencing of events in development represents a schedule that is the joint product of genetic and environmental effects. Another example in which developmental timing plays a role can be found in the "growing networks" of Cangelosi, Parisi, and Nolfi (1994), in which nodes divide according to a genetically determined schedule.

At the level of the brain, variations in timing can play an important role in the division of labor described above, determining the specialization of cortical regions for particular cognitive functions. For example, a region of cortex may be recruited into a particular task (and develop subsequent specializations for that task) simply because it was ready at the right time. Conversely, other areas of the brain may lose their ability to perform that task because they developed too late (i.e., after the job was filled).

To offer one example, differential rates of maturation have been invoked to explain the left-hemisphere bias for language under default conditions (Annett, 1985; Corballis & Morgan, 1978; Courchesne, Townsend & Chase, 1995; Kinsbourne & Hiscock, 1983; Parmelee & Sigman, 1983; Simonds & Scheibel, 1989). Because the maturational facts are still very shaky, arguments have been offered in two opposing directions! (See Best, 1988, for a detailed discussion and an alternative view.) Some have suggested that the left hemisphere matures more quickly than the right, which leads in turn to a situation in which the left hemisphere takes on harder jobs (i.e., cognitive functions that require more computation, at greater speeds). Other have made the opposite claim, that the right hemisphere matures more quickly than the left during the first year of life, and as a result, the right hemisphere takes over visual-spatial functions that begin to develop at birth, leaving the left hemisphere to specialize in linguistic functions that do not get underway until many weeks or months after birth. Variants have been proposed that incorporate both these views, e.g., that the right hemisphere starts out at a higher level of maturity (defined in terms of synaptic density and speed of processing), but the left hemisphere grows more quickly after the first year of life, leading to a division of labor in which critical aspects of visual-spatial functioning are handled on the right while linguistic functions are mediated to a greater degree on the left.

Genetic timing has also been invoked to explain critical-period effects in language learning (Johnson & Newport, 1989; Krashen, 1973; Lenneberg, 1967; Locke, 1993). However, there are at least two versions of the critical-period hypothesis that need to be considered here, one that requires an extrinsic genetic signal and another that

does not (Marchman, 1993; see also Oyama, 1992). On the "hard" maturational account, plasticity comes to an end because of some explicit and genetically determined change in learning capacity (such as a reduction in neurotrophic factors). In this case, the genetically timed stop signal is independent of the state of the system when the critical period comes to an end (see also Locke, 1993, 1994). On the "soft" maturational account, no extrinsic stop signal is required. Instead, reductions in plasticity are an end product of learning itself, due to the process of progressive cortical specialization described above. In essence, the system uses up its learning capacity by dedicating circuits to particular kinds of tasks, until it reaches a point at which there are serious limitations on the degree to which the system can respond to insult.

An example of soft maturation, mentioned above, comes from Marchman (1993), who simulated aspects of grammatical development in neural networks that were subjected to lesions (the random elimination of 2% to 44% of all connections) at different points across the course of learning. Although there were always decrements in performance immediately following the lesion, networks with small and/or early lesions were able to recover to normal levels. However, late lesions (if they were large enough) resulted in a permanent impairment of language learning. Furthermore, this impairment was more severe for some aspects of the task than it was for others (e.g., regular verb inflections were more impaired than irregular verbs). Notice that these findings mimic classical critical-period effects described for human language learning (e.g., Johnson & Newport, 1989), but without any extrinsic ("hard") changes in the state of the system. Instead, the network responds to the demands of learning through specializaion, changing its structure until it reaches a point of no return, a point at which the system can no longer start all over again to relearn the task without prejudice.

As Marchman points out, the respective hard and soft accounts of critical-period effects are not mutually exclusive. Both could contribute to the reductions in plasticity that are responsible for differences between children and adults in recovery from unilateral brain injury (see also Oyama, 1992). However, if the soft account is at

least partially correct, it would help to explain why the end of the critical period for language in humans has proven so difficult to find, with estimates ranging from one year of age to adolescence (Krashen, 1973; Johnson & Newport, 1989; for a discussion, see Bialystok & Hakuta, 1994).

One major goal of Chapter 6 will be to illustrate how the brain can solve difficult problems by "arranging" the timing of input. The idea we develop is that many complex problems have good solutions which can be best found by decomposing the problem temporally. In this way a solution may be innate not by virtue of being encoded from the start, but by guaranteeing the brain will develop in such a way that the solution is inevitable. We will refer to this kind of constraint as "chronotopic nativism." This is a powerful form of innateness which plays a central role in the evolution of complex behaviors, and is similar at an abstract level to well-known examples of timing effects in the evolution and ontogeny of physical organs (e.g., Gould, 1977; McKinney & McNamara, 1991).

Finally, we note that there is one more very important source of constraint on development, namely the constraints which arise from the problem space itself. Because this source is entirely external to the developing organism, we have not included it in our taxonomy of "ways to be innate." But this fourth source may interact crucially with the other three endogenous constraints.

We refer here to the fact that certain problems may have intrinsically good (or sometimes, unique) solutions. For example, the logical function Exclusive OR (see Chapters 2, 3, and 6) readily decomposes into two simpler functions, OR and AND. Given a range of possible architectures, networks will typically "choose" this solution without being explicitly instructed to find it. The hexagonal shape of the beehive (see Chapter 3) is another example. The hexagonal cell shape is a natural consequence of rules of geometry having to do with maximizing packing density of spheres (which then deform under pressure into hexagons). Thus, in neither the network nor the bee do the solutions (AND and OR units; hexagonally shaped cells) need to be internally specified. These outcomes are immanent in the problems themselves.

TABLE 1.3

Source of constraint		Examples in brains	Examples in networks
Representations		synapses; specific microcircuitry	weights on connections
Architectures	*unit*	cytoarchitecture (neuron types); firing thresholds; transmitter types; heterosynaptic depression; learning rules (e.g., LTP)	activation function; learning algorithm; temperature; momentum; learning rate
	local	number of layers; packing density; recurrence; basic (recurring) cortical circuitry	network type (e.g., recurrent, feedforward); number of layers; number of units in layers
	global	connections between brain regions; location of sensory and motor afferents/efferents	expert networks; separate input/output channels
Timing		number of cell divisions during neurogenesis; spatiotemporal waves of synaptic growth and pruning/decay; temporal development of sensory systems	incremental presentation of data; cell division in growing networks; intrinsic changes resulting from node saturation; adaptive learning rates

Most specific/direct → *Least specific/indirect*

On domain specificity

Armed with these distinctions between forms of innateness, we can turn to the final and perhaps most hotly disputed aspect of the nature-nurture issue, which is the *content* of what is presumed to be

innate. To what extent have we evolved mental/neural systems that serve only one master, i.e., are uniquely suited to and configured for a particular species-specific task, and no other task? This is the issue that is usually addressed with the twin terms "modularity" and "domain specificity." These are important but slippery issues. Indeed, the authors of this book have spent a good portion of their careers dealing with these vexing problems (e.g., Bates, Bretherton, & Snyder, 1988; Karmiloff-Smith, 1986, 1992a). And these issues are directly relevant to the question of innateness. We will deal with these issues again in Chapters 3, 5, and 7. For present purposes, let us struggle with a definition of terms.

It is vitally important to note that the word "module" is used in markedly different ways by neuroscientists and behavioral scientists. This has led to considerable confusion and unfortunate misunderstandings in the course of interdisciplinary discussions of brain and cognition. When a neuroscientist uses the word "module," s/he is usually referring to the fact that brains are structured, with cells, columns, layers, and regions which divide up the labor of information processing in various ways. There are few neuroscientists or behavioral scientists who would quibble with this use of the word module. Indeed, Karl Lashley himself probably had something similar in mind, despite his well-known claims about equipotentiality and mass action (Lashley, 1950).

In cognitive science and linguistics, on the other hand, the term module refers to a very different notion. Here, the term embodies a far stronger and more controversial claim about brain organization, and one that deserves some clarification before we proceed.

The strongest and clearest definition of modularity in cognitive science comes from Jerry Fodor's influential book, *The Modularity of Mind* (Fodor, 1983; see also Fodor, 1985). Fodor begins his book with an acknowledgment to the psycholinguist Merrill Garrett, thanking him for the inspiring line, "parsing is a reflex." This is, in fact, the central theme in Fodor's book, and it represents the version of modularity that most behavioral scientists have in mind when they use this term. *A module is a specialized, encapsulated mental organ that has evolved to handle specific information types of particular relevance to the species.* Elaborating on this definition, Fodor defines

modules as cognitive systems (especially perceptual systems) that meet nine specific criteria.

Most of these criteria describe the way modules process information. These include *encapsulation* (it is impossible to interfere with the inner workings of a module), *unconsciousness* (it is difficult or impossible to think about or reflect upon the operations of a module), *speed* (modules are very fast), *shallow outputs* (modules provide limited output, without information about the intervening steps that led to that output), and *obligatory firing* (modules operate reflexively, providing predetermined outputs for predetermined inputs regardless of the context). As Fodor himself acknowledges (Fodor, 1985), these five characteristics can also be found in acquired skills that have been learned and practiced to the point of automaticity (Norman & Shallice, 1983; Schneider & Shiffrin, 1977).

Therefore, another three criteria pertain to the biological status of modules, to distinguish these behavioral systems from learned habits. These include *ontogenetic universals* (i.e., modules develop in a characteristic sequence), *localization* (i.e., modules are mediated by dedicated neural systems), and *pathological universals* (i.e., modules break down in a characteristic fashion following insult to the system). It is assumed that learned systems do not display these additional three regularities.

The ninth and most important criterion is *domain specificity*, i.e., the requirement that modules deal exclusively with a single information type, albeit one of particular relevance to the species. Aside from language, other examples might include face recognition in humans and certain other primates, echo location in bats, or fly detection in the frog. Of course, learned systems can also be domain specific (e.g., typing, driving, or baseball), but according to Fodor they lack the instinctual base that characterizes a "true" module. In the same vein, innate systems may exist that operate across domains (see below for examples). However, in Fodor's judgment such domain-general or "horizontal" modules are of much less interest and may prove intractable to study, compared with the domain-specific or "vertical" modules such as language and face recognition.

We discuss these claims at length in Chapters 3 and 7. For present purposes, we point out that a serious problem with such claims regarding domain specificity is their failure to recognize that specificity may occur on (at least) five levels: *tasks*, *behaviors*, *representations*, *processing mechanisms*, and *genes*.

(a) *Specificity of tasks and problems to be solved.* We will define a "task" as a problem that the organism must solve in order to achieve some goal. Each task or problem can be defined in terms of a "problem space," a set of parameters that includes a well-defined goal (G), specification of the environment in which the organism must work (E), the resources and capabilities that the organism brings to bear on the problem (R), and a description of the sequence of operations that will lead to G given E and R. Of course such parameters are implicit in the situation; they are not necessarily represented explicitly anywhere in the organism or in the environment.

Our point for present purposes is this: Most problems are unique, and thus form the starting point for an analysis of domain specificity. For example, human languages emerged within a rich problem space that has little in common with the many other things we do. Put in the simplest possible form, languages represent solutions to the problem of mapping inherently nonlinear patterns of thought onto a linear sequence of signals, under a severe set of processing constraints from human perception, motor coordination and production, and memory. Of course all complex behaviors must be organized in time, but the temporal constraints on language may be unique due to the complexity of the information that must be conveyed and the multiplex of constraints on use of that information in real time—to say nothing of the problem of learning, addressed in more detail in Chapters 2 and 6.

(b) *Specificity of behavioral solutions.* Let us be clear on this point: language is "special," a unique problem unlike any other that we face, and unlike the problems faced by any other species. Similar claims can be made for other complex systems as well. This does not mean that every complex problem has forced special solutions all the way down to the genetic level. However, it is quite likely that specific problems will require specific solutions at the behavioral

level. There is, for example, nothing outside of language that resembles past tense morphology in English, case marking in Turkish, nominal classifiers in Navajo, or relative clauses in German and Dutch. Grammars are complex behavioral solutions to the problem of mapping structured meanings onto a linear string of sounds. Any resemblance to other cognitive domains (real or fictional) is purely coincidental. But the same is true for domains that cannot conceivably be innate (e.g., tennis; chess; chicken sexing). The domain specificity of behavior does not in itself constitute evidence for the innateness of that behavior. The real arguments lie at the next few levels down.

(c) *Specificity of representations.* An individual who can produce domain-specific behaviors on a reliable basis must (within the framework we have outlined here) possess a set of mental/neural representations that make those outputs possible. We are persuaded that such representations must also have a high degree of domain specificity. But as we have suggested here, and shall elaborate in detail in Chapter 5, we are skeptical that detailed representations of any kind are innate, at least at the cortical level. So, in our view, the argument for "vertical modules" (Fodor's term for innate and domain-specific systems) must lie at another level.

(d) *Specificity of processing mechanisms.* Assuming that higher (and lower) cognitive processes require domain-specific representations, must we assume that a domain-specific representation is handled by a domain-specific processor, or that it is acquired by a domain-specific learning device? This is the crux of the debate about domain specificity, the point at which arguments about domain specificity and innateness cross in their most plausible and compelling form. This critical issue breaks down into two distinct questions: a "where" question and a "how" question.

The "where" question also breaks into two related issues (addressed in more detail in Chapter 5). Are the representations required for a specific domain coded in a particular (compact) region of the brain, or are they widely distributed across different cortical and/or subcortical areas? Do the representations required for a specific task occupy their own, unique, dedicated neural tissue, or must they share neural territory with other tasks?

It is obvious why these are related issues. On the one hand, if each cognitive function is localized to a specific, compact region, then it is not unreasonable to assume (or it could at least be true) that each region functions as a specialized processor, used always and only for a specific task (e.g., a face processor, a language processor, a music processor, a Grandmother cell, a yellow Volkswagen detector). As we have already noted, this claim is independent of the issue of innateness, since most of these specializations can be (and probably are) acquired through experience. In the limiting case, every concept (innate or otherwise) would be assigned to a specific neuron. With 10^{11} neurons to hand out to all comers, it would take a very long time to run out of neural capacity (see Churchland, 1995, for an enlightening discussion of this point).

On the other hand, if the representations required for a specific task are widely distributed across brain regions, then the case for specialized processors is necessarily weakened. Simply put, we would not have enough brain to go around if we handed out huge tracts of territory for the exclusive use of language, faces, music, etc. If the mechanisms that store (for example) linguistic representations are broadly distributed in the brain, then it is quite likely that those mechanisms are also used for other cognitive functions. That is why claims about localization and claims about domain specificity so often go hand in hand, and why the connectionist notion of distributed representations is not well received by those who believe that domain-specific (perhaps innate) representations must be handled by domain specific (perhaps innate) processing mechanisms.

The "how" question has to do with the nature of the operations that are carried out by a processing device. Assuming that we have a set of domain-specific representations (innate or learned), handled by a dedicated processor (probably local, since distributed but domain-specific processors are limited by their size), does it necessarily follow that the processor carries out unique and domain-specific operations? Can a general-purpose device (or, at least, a multipurpose device) be used to learn, store, activate or otherwise make use of a domain-specific representation? Or must the processing characteristics of the tissue be tailored to the requirements of a specific task, to the point where it really cannot be used for anything

else? Even if the processor had rather plastic, general properties at the beginning of the learning process, does it retain that plasticity after learning is complete, or has it lost the ability to do anything else after years of dealing with (for example) language, faces, music, Grandmothers or yellow Volkswagens? The "big issues" of innateness, localization and domain specificity have clear implications for a fourth big issue addressed throughout this volume (but especially in Chapter 5): the issue of developmental plasticity.

We will insist throughout that connectionist models are not inherently "anti-nativist" and that they are certainly not the natural enemy of those who believe in domain specificity (in fact, a major criticism of current connectionist models is that most of them can do only one thing at a time; see Chapter 7, and Karmiloff-Smith, 1992a). However, these models do assume distributed representations, and those representations are usually distributed across processing units of considerable flexibility. To the extent that realistic developments can be modeled in a system of this kind, there is a prime facie case against the notion that domains like language, music, faces or mathematics *must be* carried out by dedicated, innate and domain-specific neural systems. Hence the issue of domain specificity at the level of processing mechanisms connects up with major controversies about innateness. This brings us to the sine qua non of "old-fashioned" nativism, the issue of genetic specificity.

(e) *Genetic specificity.* By definition, if we make a claim for innateness and domain specificity at any of the above levels, we mean that the outcome is ensured by and contained within the genome. But just how, and where, and when does this occur? There seems to be a widespread willingness to believe in single genes for complex outcomes (see Chapter 7 for a detailed accounting of the Grammar Gene Rumor). Things would be much simpler that way! But the evidence to date provides little support for this view. Alas, a complex cascade of interactions among genes is required to determine outcomes as simple as eye color in fruitflies or body types in earth worms. More often than not, the genes operate by controlling variables in timing. When the default schedule is followed (within some limits of tolerance), certain interactions between structures

inevitably occur. There is no need for genes to encode and control those interactions directly. Instead, they follow from the laws of physics, geometry, topology—laws of great generality, but laws that have very specific consequences for the actors on that stage (D'Arcy Thompson, 1917/1968).

If this is so demonstrably true for the embryogenesis of simple organisms, why should things be different for the embryogenesis of the human mind? It is unsettling to think that our beloved brain and all its products result from such a capricious game. That is why the idea that single genes and detailed blueprints are responsible for what we are is attractive to many. Albert Einstein once insisted that "God does not play dice with the Universe." We sympathize with this view, but we remember Nils Bohr's reply: "Stop telling God what to do."

The shape of change

If developmental paths were always straight and always uphill, they would not be nearly as interesting as they are. One usually implicitly assumes a linear model of growth and change; we presuppose that, all things being equal, development will be gradual (no sudden accelerations or decelerations), monotonic (always moves in the same upward direction), and continuous (reflecting quantitative changes in some constant measure or dimension). This is the "garden variety" model of change.

So when we observe developmental phenomena which deviate from this model, we find it interesting. But more: We assume that these deviations reflect changes in the underlying mechanisms. So, for example, when we see a U-shaped curve in children's acquisition of the past tense (Brown, 1973; Ervin, 1964; Kuczaj, 1977), it seems reasonable to infer that the child's similar performance at the early and late stages arises from different sources (e.g., rote learning early on, use of rules at the later stage). If we see a child ignoring certain inputs which are attended to later (for example, the relative importance of mass vs. size when attempting to balance weights on a scale), we infer that something internal has changed or appeared which makes it possible for the child to process what was previ-

ously ignored. And when we see children able to learn languages which adults learners can only imperfectly master, we assume that some critical component of the learning mechanism which is present in childhood is lost to the adult.

As reasonable as these inferences are, we know now—and this has been one of the powerful lessons of connectionist models—that nonlinear change in behavior can be generated by mechanisms which themselves undergo no dramatic or nonlinear changes in operation whatsoever. Connectionist models illustrate how the same learning mechanism can give rise to behaviors which differ dramatically at various points in time. The changes need not be linear nor monotonic, nor do they need to be continuous. We describe networks with these properties in Chapters 2 and 3, and analyze the basis for their nonlinear behavior in Chapter 4.

Partial knowledge

The idea of partial knowledge is central to the mystery of development. If an adult knows something, and an infant does not, then—unless we assume instantaneous change—the state in between can only be characterized as a state of partial knowledge. And notions of innateness frequently invoke the concept of partial knowledge which is prespecified, to be supplemented by experience. But what exactly is meant by partial knowledge? What could it mean to have "half an idea?" As intriguing as this concept is, it serves no useful or explanatory function if left vague and unspecified.

In a few domains, it is easy to characterize partial knowledge. For example, a child who knows how to multiply and subtract may be said to have partial knowledge of how to do long division. One might even try to quantify just how complete that knowledge is and claim the child has 75% of the knowledge required. But such cases are few and far between. Knowledge of complex domains is more often *not* a matter of assembling pieces in a jigsaw puzzle together. Having 90% of an idea does not usually mean lacking only 10% of the pieces.

In the connectionist models we shall discuss, there are certain key concepts which give us a new way to think about partial knowl-

edge. The idea of *distributed representations* is discussed in Chapters 2 and 3; we see what it means for knowledge to be distributed across a complex system, and the consequences of superposition of representation across the same processing elements. This makes it possible to talk about partial knowledge in situations where that knowledge is integrated across multiple domains. The properties of *gradient descent learning* and *nonlinear* processing are discussed in Chapters 2, 3, and 4. We see how it is possible for incremental learning to give rise to discontinuities in behavior; we can talk about "subthreshold knowledge."

The value of simulations

Although mathematical models are common in some areas of psychology, and computer simulations have a long tradition in fields such as artificial intelligence, many developmentalists may find the methodology of computer simulation of models to be strange. Yet such simulations play a central role in connectionism, and we deem them important enough that we have written a companion volume to this book in order to explain the methodology in detail. Why?

First, these simulations enforce a rigor on our hypotheses which would be difficult to achieve with mere verbal description. Implementing a theory as a computer model requires a level of precision and detail which often reveals logical flaws or inconsistencies in the theory.

Second, although connectionist models often appear simple—they are, after all, merely collections of simple neuron-like processing elements, wired together—their simplicity is deceptive. The models possess nonlinear characteristics, which makes their behavior difficult (if not impossible) to predict. The use of distributed representations also means that the models exhibit emergent behaviors which can usually not be anticipated. The simulation therefore plays the role of an empirical experiment in allowing us to study the behavior of our theory in detail. In the case of evolutionary simulations, there is the additional benefit that effects which normally unfold over hundreds of thousands, or millions or billions of years,

can be explored in a speeded-up model which may run in hours or days.

Third, the model's innards are accessible to analysis in a way which is not always possible with human innards. In this sense, the model functions much as animal models do in medicine or the biological sciences. After a computer model has been trained to generate a behavior which is of interest to us, we can inspect its internal representations, vary subsequently the input to it, alter the way it processes the input, and so forth. With humans, we can usually only guess at the nature of the mechanism responsible for a behavior by inference, but with the simulation we can directly inspect the network in order to understand the solution. Of course, it remains to be demonstrated that the model and human that it simulates do things the same way; but the model can be a rich source of hypotheses and constraints which we might not have stumbled across in human experimentation. Indeed, the conceptual role played by these models, in giving us new ways to think about old problems, is for us one of the most exciting and profitable reasons to do connectionist simulations.

* * *

In the remaining chapters of this book we shall attempt to flesh out the perspective we have outlined above. We begin in Chapter 2 with a minitutorial on connectionism, selectively emphasizing those concepts which are especially relevant to developmental issues. We stress the importance of connectionism as a conceptual framework, rather than simply a modeling methodology. In Chapter 3 we present what we see as some of the major challenges and mysteries to be solved in development, and discuss a number of connectionist simulations which start to address these issues. Chapter 4 focuses on the different "shapes of change" and introduces notions of nonlinearity and dynamics. We then show how connectionist networks provide accounts for the various patterns of change and growth which are observed in development. Chapter 5 is about brains: what they look like, how they develop, and how they respond to injury. These data bear directly on the question about the plausible

loci of constraints. In Chapter 6 we return to the theme of interactions and present simulations which illustrate how very specific outcomes can be produced through very indirect means. Finally, we sum things up in Chapter 7 and propose a new way of thinking about innateness.

We emphasize from the outset that the approach we take is not anti-nativist—far from it. Where we differ from some is that we view the major goal of evolution as ensuring adaptive *outcomes* rather than ensuring prior knowledge. The routes by which evolution guarantees these outcomes are varied and often very indirect. The mechanisms need not be optimal, nor tidy, nor easy to understand. It suffices that they barely work, most of the time. There is an evolutionary trade-off between efficiency and flexibility. We are prepared to call many universally recurring patterns of behavior—in languages, for example—innate, even though we find them nowhere specified directly in the genome. In this sense, our definition of innateness is undoubtedly broader than the traditional view. We also believe that it is richer and more likely to lead to a clearer understanding of how nature shapes its species. We hope, by the time you finish reading what we have to say, that you agree.