The Meaning of Life: A Merleau-Pontian Investigation of How Living Bodies Make Sense

Author: Noah Moss Brender

Persistent link: http://hdl.handle.net/2345/bc-ir:103615

This work is posted on eScholarship@BC, Boston College University Libraries.

Boston College Electronic Thesis or Dissertation, 2012

Copyright is held by the author, with all rights reserved, unless otherwise noted.
Boston College

The Graduate School of Arts and Sciences

Department of Philosophy

THE MEANING OF LIFE: A MERLEAU-PONTIAN INVESTIGATION OF HOW LIVING BODIES MAKE SENSE

a dissertation

by

NOAH MOSS BRENDER

submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

May 2013
THE MEANING OF LIFE: A Merleau-Pontian Investigation of How Living Bodies Make Sense
By Noah Moss Brender
Supervisor: Jeffrey Bloehl

This dissertation takes up Maurice Merleau-Ponty’s unfinished project of developing an ontology of nature whose concepts are drawn from the phenomenon of life, rather than from human techne. I argue that the question of life has been hopelessly obscured by the collapse, in the Modern era, of the distinction between nature and artifice. We cannot hope to understand the difference between life and non-life until we understand the difference between the living body and the machine.

Merleau-Ponty’s constant aim was to show that the living body is not a blind mechanism, and that the body has its own endogenous sense which is not projected onto it by a disembodied consciousness. Central to these efforts were the phenomena of learning and development, and the concept of form or Gestalt. Development is what distinguishes the living body, which is an open-ended process of becoming, from the machine, whose possibilities are determined in advance by its creator. In order to conceptualize the phenomenon of development, Merleau-Ponty appropriated from psychology the concept of form (Gestalt): a dynamic, self-organizing whole that cannot be decomposed into independent parts. Where the conception of nature as mechanism implies that everything is determined in advance, Merleau-Ponty’s conception of nature as Gestalt allows for the genesis of genuinely new phenomena through nature's own self-organizing movement. We would thus be able to understand the genesis of sense in nature as a process of morphogenesis—the genesis of form.

However, Merleau-Ponty struggled to clarify the ontological status of form. He lacked the conceptual resources to explain form in its own terms, rather than by contrast with the decomposable wholes of human artifice. This dissertation attempts to locate these conceptual resources in the science of complexity that has emerged since Merleau-Ponty’s death, and whose descriptions of complex systems are uncannily anticipated in Merleau-Ponty’s writings. I take from this new science the conception of form as asymmetry or difference, and of morphogenesis as symmetry-breaking or self-differentiation. In order to investigate how meaning emerges out of form, I turn to recent work in biology and psychology that applies the concept of symmetry-breaking to the phenomena of anatomical growth and motor development. By studying the development of the living body and its behavior, I show how nature articulates itself into perceiver and perceived. In the movement of the living body, form folds back upon itself, giving rise to a new kind of meaning: a pre-reflective, motor significance that is neither mechanism nor mental representation.

In Chapter One, I distinguish the living body from a machine or artifact by distinguishing between manufacturing and growth. This distinction, which seemed obvious to the Ancients, has been obscured by Modern science's pivotal decision to treat nature as if it were a product of human artifice. This decision has committed us to an atomistic ontology, which takes nature to be a synthetic whole composed of mutually indifferent parts. However, this ontology faces a basic problem, which I call the problem
of form: how to explain the synthesis of indifferent atoms into the complex, harmonious wholes we observe in nature, without appealing to an intelligent designer. Nowhere is this problem more acute than in the phenomenon of anatomical development or embryogenesis. I argue that biology has been unable to explain this phenomenon in mechanical or atomistic terms: the Neo-Darwinist view of the living body as a synthetic whole determined in advance by a genetic blueprint or program has succeeded not by explaining development, but rather by ignoring it.

In Chapter Two, I argue that the problem of form—and of living form in particular—can only be resolved by abandoning our atomistic ontology, and with it our synthetic understanding of form as a shape imposed on an indifferent material. Recent developments in the science of complexity have yielded a new definition of form as asymmetry or difference. On this view, the genesis of form in nature is not the synthesis of wholes out of pre-existing parts, but the self-differentiation of wholes into parts through symmetry-breaking. In order to understand how natural wholes become less symmetrical over time, I introduce three further concepts from the science of complexity: nonlinearity, stability, and instability. With these concepts in hand, I return to the problem of embryogenesis, in order to show how complex living forms can develop reliably and robustly without being determined in advance by a design or program.

In Chapter Three, I turn from anatomical development to the development of behavior, in order to see how the genesis of form becomes a genesis of sense. I begin by criticizing three mechanistic theories of behavior—Behaviorism, Cognitivism, and Connectionism—which suffer from the same problem of form that plagues mechanistic theories of anatomical development. Behavior grows like an organ: by symmetry-breaking, not by synthesis. Learning is not a matter of association, but of differentiation: the perception of increasingly subtle asymmetries in the body's environment through increasingly asymmetrical movements. It is the world that teaches the organism how to move—but a world that is only revealed to the organism by its own movements. Thus the living body and its world grow together dialectically, each driving the other to become more determinate through its own increasing determinacy.
# Table of Contents

Analytic Table of Contents ........................................................................................................... ii

ABBREVIATIONS ........................................................................................................................... iv

PREFACE .......................................................................................................................................... v

INTRODUCTION: Returning to life ............................................................................................... 1

CHAPTER ONE: Is the living body a machine? .............................................................................. 12

CHAPTER TWO: How does form arise in nature? .......................................................................... 54

CHAPTER THREE: How do living bodies make sense? ............................................................... 110

CONCLUSION: The meaning of life ............................................................................................ 190

APPENDIX A. Dynamic systems theory: A different kind of scientific model ..................... 205

WORKS CITED .............................................................................................................................. 239
# Analytic Table of Contents

<table>
<thead>
<tr>
<th>ABBREVIATIONS</th>
<th>iv</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREFACE</td>
<td>v</td>
</tr>
<tr>
<td>INTRODUCTION: Returning to life</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER ONE: Is the living body a machine?</td>
<td>12</td>
</tr>
<tr>
<td>1. What are living bodies made of?</td>
<td>13</td>
</tr>
<tr>
<td>2. What is the difference between manufacturing and growth?</td>
<td>15</td>
</tr>
<tr>
<td>3. Atomism and the problem of form</td>
<td>23</td>
</tr>
<tr>
<td>4. The paradox of development</td>
<td>31</td>
</tr>
<tr>
<td>A. Epigenesis vs. preformation</td>
<td>32</td>
</tr>
<tr>
<td>B. Darwinism, genetics, and the Modern Synthesis in biology</td>
<td>34</td>
</tr>
<tr>
<td>C. DNA and the return of development</td>
<td>39</td>
</tr>
<tr>
<td>D. The concept of a genetic program</td>
<td>45</td>
</tr>
<tr>
<td>5. Conclusion</td>
<td>48</td>
</tr>
<tr>
<td>CHAPTER TWO. How does form arise in nature?</td>
<td>54</td>
</tr>
<tr>
<td>1. What is form?</td>
<td>55</td>
</tr>
<tr>
<td>2. Form as asymmetry</td>
<td>59</td>
</tr>
<tr>
<td>3. Symmetry-breaking as the genesis of form</td>
<td>64</td>
</tr>
<tr>
<td>A. Symmetry-breaking bifurcations in nonlinear dynamic systems</td>
<td>65</td>
</tr>
<tr>
<td>B. Stability and instability</td>
<td>75</td>
</tr>
<tr>
<td>C. Limit cycles and Hopf bifurcations</td>
<td>84</td>
</tr>
<tr>
<td>D. Nonlinearity, indecomposability, and unpredictability</td>
<td>88</td>
</tr>
<tr>
<td>4. The genesis of living form</td>
<td>92</td>
</tr>
<tr>
<td>A. Division</td>
<td>94</td>
</tr>
<tr>
<td>B. Differentiation and the role of DNA in development</td>
<td>98</td>
</tr>
<tr>
<td>C. Morphogenesis</td>
<td>103</td>
</tr>
<tr>
<td>D. Summary</td>
<td>105</td>
</tr>
<tr>
<td>5. Conclusion</td>
<td>106</td>
</tr>
</tbody>
</table>
CHAPTER THREE: How do living bodies make sense?.......................................................... 110

1. Introduction...................................................................................................................... 111

2. Mechanistic theories of behavior ................................................................................ 115
   A. The reflex theory of behavior ...................................................................................... 115
   B. Cognitivism .................................................................................................................. 120
   C. Connectionism ............................................................................................................ 127

3. Learning to move ........................................................................................................... 129
   A. The view from above and the view from below ......................................................... 129
   B. The moving body as a dynamic system .................................................................... 132
   C. Motor development as morphogenesis .................................................................. 139

4. Learning to perceive ...................................................................................................... 147
   A. Perception and the self-organization of behavior ..................................................... 147
   B. Motor programs vs. sensorimotor habits .................................................................. 162
   C. Perception and motor development ....................................................................... 168
   D. Learning as differentiation ...................................................................................... 178

5. Conclusion ...................................................................................................................... 187

CONCLUSION: The meaning of life .................................................................................. 190

APPENDIX A. Dynamic systems theory: A different kind of scientific model .......... 205

WORKS CITED .................................................................................................................... 239
ABBREVIATIONS

Works by Maurice Merleau-Ponty:

La structure du comportement (3rd ed.). Quadrige/puf, 2006 [1942].

Phénoménologie de la Perception. Gallimard, 2008 [1945].


Works by other authors:


PREFACE

Thus, on the one hand it is necessary to follow the spontaneous development of the positive sciences by asking whether man is really reduced to the status of an object here, and on the other hand we must reconsider the reflexive and philosophical attitude by investigating whether it really gives us the right to define ourselves as unconditioned and timeless subjects. It is possible that these converging investigations will finally lead us to see a milieu which is common to philosophy and the positive sciences, and that something like a third dimension opens up, this side of the pure subject and the pure object, where our activity and our passivity, our autonomy and our dependence no longer contradict one another.  

[M]y general aim is always to… confront the concepts which we are accustomed to using, and which I am accustomed to using, with the realities they are supposed to designate.  

The phenomenological method is founded on Husserl's ideal of a "presuppositionless science": an inquiry that would not prejudge the phenomena it is investigating by deciding on the terms of this investigation in advance. Instead, a phenomenological inquiry would allow the phenomena themselves to generate the concepts we use to think about them. This would be a thoroughly empirical philosophy: rather than trying to deduce a priori truths from first principles, phenomenology would return "to the things themselves" and allow them to speak for themselves. To understand the meaning of the phenomena, however, we must learn to speak their language. Phenomenology thus sets itself the task of learning the logos of the phainomena—letting the phenomena teach us their native tongue, rather than forcing them to speak in terms that we already understand.  

However, phenomenology does not fall into the typical Modern error of trying to wipe the slate clean, to rid ourselves all at once of the concepts and presuppositions we have inherited and start fresh from a new foundation. Nor does phenomenology follow

---

1 TT 13 (translation from Waldenfels 1980, p. 21).
2 Merleau-Ponty, "Les voyages du philosophe," eleventh interview with Georges Charbonnier, R.T.F., first broadcast July 31, 1959. (Citation and translation from Noble 2011, p. 93.)
the Romantic path of trying to abandon concepts and theories altogether in search of an immediate, intuitive coincidence with nature. Instead, phenomenology starts just where we are, with the phenomena as they appear to us from our present theoretical standpoint, and looks for those *aporia*—those puzzles, problems, and antinomies—that indicate the places where our concepts are inadequate to the phenomena they describe. Thus phenomenology is a *dialectical* philosophy: it uses the phenomena revealed by our present concepts and questions to challenge those very concepts and formulate better questions, which in turn disclose new phenomena.³ This is an endless task: each new generation of phenomenologists seeks to unearth foundational assumptions that their predecessors have left unexamined, and to hold them up to the light of the phenomena. In this way, each generation strives to penetrate more "radically" into the roots of its own thought.

Of course, the concepts and theories that phenomenology takes as its starting point do not come from nowhere; we have inherited them from a long philosophical tradition. But this tradition is not transparent to itself: our inheritance largely takes the form not of explicit claims but of unstated assumptions, which seem so obvious to us that they no longer require justification. Indeed, these assumptions are for the most part so familiar that they are no longer visible to us at all—they have become part of the conceptual bedrock of our lives, buried under centuries of sedimented thought. Thus the phenomenological project of tracing our *aporia* to their sources is a *historical* or

---

³ I have chosen the term "dialectical" over the more common "hermeneutical" deliberately, in order to suggest that there is a certain methodological continuity between Plato's Socrates, Hegelian phenomenology, and twentieth-century phenomenology—though I will not argue for this continuity here. On Socratic dialectic, see *Republic* 510b-511d. On the relation between Hegelian and twentieth-century phenomenology, see Merleau-Ponty, "Hegel's Existentialism" (*SN*, pp. 63-70); Russon 2010.
genealogical task: it brings to light conceptual prejudices which have been hidden by their long familiarity, and makes them appear strange and new again. Studying the history of our tradition as it has been preserved and passed down to us in historical documents, is an indispensable part of this project. To read the history of philosophy phenomenologically is to see the past neither as completely alien—an object of mere curiosity; nor as utterly familiar—a story of progress culminating in the present. Rather, phenomenology reads the tradition as a series of choices that could have been made otherwise and questions that could have been answered differently. For phenomenology, the past is not over and done with, but alive with possibilities that we can take up anew in the present.

Certain readers may be surprised to see a description of the phenomenological method that makes no mention of "qualia," "lived experience" or the "first-person perspective." To define phenomenology in these terms would be to presuppose the distinctions between introspection and scientific observation, first-person and third-person, or "lived experience" and scientific theory. But there is nothing self-evident about these distinctions: they are highly theoretical, embodying a whole complex of inherited ontological assumptions. Thus any inquiry that defines itself in terms of these oppositions is not a phenomenological inquiry in the sense that I am defending here. Far from presupposing these concepts or treating them as self-evident, a phenomenological philosophy would have to take them as objects of investigation. We would have to ask whether these distinctions can be justified empirically, or whether, on the contrary, the phenomena themselves demand that these traditional concepts be rejected.
This three-fold determination of the phenomenological project as empirical, dialectical, and historical is likely to give rise to two apparently fatal objections. First, someone will say that phenomenology has arrived on the scene several hundred years too late. The method it proposes might have been novel or radical before the Scientific Revolution; but nothing could be more commonplace today than to suggest that philosophy ought to be empirical, or that our theories should be progressively improved by testing them against the phenomena they reveal. This is precisely what the natural sciences have been doing, with consummate success, for centuries. To call, today, for a return "to the things themselves" seems absurd—one might as well call for the end of mercantilism, or try to write Don Quixote.

This first objection will quickly be followed by a second: that the phenomenological project of reaching "the things themselves" is hopelessly naïve. It seems that the phenomenologist is oblivious not only to the existence of Modern science, but also to the history of Modern philosophy, which has firmly established that we can only know our own representations. Sense perception cannot give us the things themselves, but only their effects on us; and there is no reason to suppose that these effects resemble their causes. Moreover, the testimony of our senses is confused and contradictory; it can only be made intelligible by the operations of the mind. Thus even our perceptions, to the extent that anything meaningful or objective is to be found in them, are products of consciousness' own synthetic activity. We do not perceive reality with our bodies, but rather construct it with our minds. But this means that we are never conscious of things as they are in themselves, but only of our own ideas. If there is a
world beyond thought, it is unknowable to us, for thought knows only itself and is subject only to its own internal necessities.⁴

Either one of these objections on its own would suffice to dissuade us from pursuing phenomenology any further. And yet there is something strange, even paradoxical, about the two objections taken together. One claims that phenomenology is unnecessary; the other, that it is impossible. One dismisses phenomenology because its project has already been accomplished; the other, because it cannot be accomplished. How is it that both of these objections seem persuasive to us? How can phenomenology's principles appear at once trivially true and demonstrably false?

These contradictory objections are an indication of the paradoxical situation in which we find ourselves today. When we are engaged in the practice of science, we meet with success after success. We know more about the natural world than ever before, and this knowledge manifests itself in an unprecedented mastery over nature—including the human body. Each day brings new discoveries and new inventions, and the pace of this progress seems only to increase with time. And yet, when science turns its attention to the phenomena of perception and cognition, it arrives at a shocking and paradoxical conclusion: that the human mind cannot know the external world. The brain, after all, is encased inside the skull; its only access to the world is through the nervous signals it receives from the body's sense organs. Thus the brain must piece together a picture of the external world from its effects on the body. All of our knowledge consists of inferences

⁴ On this point, Post-Modernism simply reaffirms the conclusions of Modern philosophy. We can replace "thought" with "language" or "power" without altering the force of this objection to phenomenology: if there is nothing outside of discourse, or if we can only know our own social constructs, then phenomenology's naïve attempt to return to "the things themselves" will succeed only in uncritically reinscribing existing regimes of power.
from given effects to hidden causes: we are constantly constructing an internal model of
the world that would explain the sensations we have received. But this model is never
more than a hypothesis, for we can never get outside of our own heads to compare our
representations with the things they are supposed to represent. The only measure of the
model—the only standard to which it can answer—is its predictive power: whether it
allows us to correctly infer from our past sensations what sensations will arrive next.
Thus we find ourselves in the same position as the prisoners in Plato's cave, staring at the
shadows of material things and trying to guess which shadow will appear next. But our
situation is even worse than that of Plato's prisoners: they could be freed from their
chains and dragged out into the light to confront the material world directly; but only
death awaits the brain freed from the confines of its skull. Our brains cannot confront the
external world directly. They require the mediation of the senses. And yet this very
mediation seems to leave us stranded inside of our own heads, permanently cut off from
the objects of perception and knowledge. Thus the nature that science had naïvely taken
to be out there in the world, independent of our knowing activity, turns out—on science's
own account of cognition—to be a mental representation. In an abrupt and dizzying
change of perspective, nature and consciousness switch places: where consciousness had
been just another object within nature, suddenly nature becomes just another idea within
consciousness. It is as if nature were a bag that science had tried to sew shut, and
consciousness were its lining; the bag was all but closed except for a little thread of
consciousness still dangling outside of it—but when we yanked on this thread the whole
bag turned inside out, and nature was trapped within the very consciousness it was
supposed to enclose.
What is the cause of this uncanny reversal? The problem arose as soon as science attempted to explain human perception and cognition. These appear to be natural phenomena like any other; as such, they ought to be susceptible to the same scientific methods that have succeeded so well in every other area of nature. However, there is something special about these phenomena that sets them apart from the rest of the natural world: perception and cognition are the activities that the scientist herself is engaged in when she investigates nature. In studying these phenomena, then, the scientific method was unwittingly turned back upon itself. So long as science's gaze is turned resolutely outward, its progress is unchecked; but when knowledge becomes its own object—when science attempts to reflect on itself, to explain itself as it has explained everything else—it becomes mired in the most baffling aporia. This is not to say that the scientific study of human perception and cognition has ground to a halt; on the contrary, these investigations proceed confidently, seeming to make steady progress alongside every other specialized science. Contradictions only appear when we recall that the cognition we are studying is our own—that the cognitive scientist is herself engaged in the very act that she is investigating. The problem, in other words, is that we cannot seem to reconcile the human being as knower with the human being as object of knowledge. When we try to complete our picture of nature—to sew the bag shut, as it were—we always find ourselves qua knowers standing on the outside of nature looking in. As in a dream or an out-of-body experience, we can see ourselves—or at least our bodies—there within the system of nature, moving according to the same scientific laws as everything else; and we can even be quite content with this view, so long as no one asks us where we take
ourselves to be standing. Then suddenly the bottom falls out, and the abyss yawns beneath us; we realize that we must be dreaming, and yet we cannot wake.

For the most part, of course, we are content to go on sleeping, beguiled by the dream of mastery over nature. The *aporia* I have described is seen as a philosophical problem, not a scientific one. The scientist is not asked to reflect on her own standpoint, and indeed is discouraged from doing so; such reflections do not produce results, but only conundrums. Thus science and philosophy go their separate ways: science goes about its business with continued success, and it is left to philosophy to consider the consequences. Since philosophy for the most part feels itself incapable of questioning science's results, it can only counsel resignation: we must accept that we have no ground on which to stand, and try to overcome our fear of falling. Thus we arrive at our present, paradoxical situation: science assures us that we know more than ever, while philosophy tells us with equal assurance that there is no such thing as knowledge. There is only calculation, only power: power over nature, over others, and over ourselves. For the most part, we hold both of these views simultaneously—unable to reconcile them, but equally unable to choose between them, since they share a common root. If the problem could be located entirely within philosophy, then we could simply abandon this outmoded practice in favor of its superior offspring, the natural sciences. But this is impossible; for what seems to be a conflict between science and philosophy is really a conflict within science itself, which philosophy merely calls to our attention. It is science that tells us both that we are masters of nature and that free will is an illusion. If nature is a machine, its movements determined in advance by immutable laws, then we can predict and control it; but if nature is a machine then so are we, and our actions too must be determined in advance—
including those of the scientist. Thus science affirms the self-contradictory claim that we are masters of nature and also, at the same time, its slaves.

It is this paradoxical situation that calls for phenomenology, and to which phenomenology attempts to respond. Rather than resigning ourselves to the antinomies of Modern thought, phenomenology takes them as a sign that our thinking has somewhere gone astray. If our arguments generate contradictions, then at least one of our premises must be false. We must therefore go back and examine our assumptions to see which of them are truly grounded in the phenomena, and which are merely prejudices that we have inherited uncritically from our predecessors. As I have briefly indicated, the perplexities in which we find ourselves are tied inextricably to the principles and practices of Modern science. Thus phenomenologists have been driven repeatedly to confront the origins and development of the natural sciences. In *The Crisis of European Sciences and Transcendental Phenomenology* (1970 [1954]), Husserl traces our present perplexities to Galileo's claim, at the very origin of Modern science, that nature itself is mathematical, and that the qualities we perceive therefore reside not in things but in us, as the effects of quantitative reality on our bodies. Heidegger, for his part, observes that by construing nature as an object of knowledge, Modern science implicitly locates the subject of this knowledge outside of nature (1982b). Thus Modern science is implicitly theological, or in Heideggerian terms, "metaphysical": it presupposes a God's-eye-view, a transcendent standpoint which would apprehend the true world behind the world of mere appearances. Heidegger (1982a) also traces the development of this scientific worldview into our present, "technological" worldview: the cleavage between subject and object ultimately
undermines the very possibility of knowing nature as it is in itself; thus the only possible relation to nature becomes one of calculation and manipulation, prediction and control. Nature no longer even appears to us as an object of knowledge, but only as a resource for our own activities, and the question of truth is thus supplanted by the question of what works.

The phenomenologist who engages with Modern science in the most thorough-going and fruitful way, however, is Maurice Merleau-Ponty. Both Husserl and Heidegger offer insightful analyses of Modern science's foundational assumptions. But these analyses take place from a philosophical standpoint. They are external criticisms rather than immanent critiques. Merleau-Ponty, on the other hand, does more than simply expose the foundational assumptions of the natural sciences: he engages directly with their empirical results, in order to show how the sciences' own findings contradict their methodological principles, and so demand a reform at the level of basic concepts. Merleau-Ponty thus seeks to avoid perpetuating the conflict between philosophy and science by taking up the natural sciences on their own terms—arguing not that they have been insufficiently philosophical, but that they have been insufficiently scientific. At the same time, Merleau-Ponty challenges philosophy to be more resolutely empirical. His goal is not simply to criticize the natural sciences, but to learn from them:

How could one want to know what Nature is and not be interested in science? If Nature is something all-encompassing [un Englobant] then we cannot think it starting from concepts, let alone deductions, but must rather think it starting from experience, and in particular from experience in its most regulated form—namely, science. (N 87/122, translation modified)
Merleau-Ponty's first book, *The Structure of Behavior*, opens with the question of the relation between consciousness and nature—where nature is to be understood as "a multiplicity of events external to each other and bound together by relations of causality" (SB 3/1). This question, and the particular formulation that it takes, are no accident: this is the form that the most urgent philosophical questions have assumed in our time. Thus it is here that we must begin. However, Merleau-Ponty does not simply take this starting-point for granted. On the contrary, his aim—in the dialectical movement which is characteristic of phenomenology—is to subject this question and the terms in which it is posed to the most searching scrutiny.\

---\

5 Merleau-Ponty's first published work was unduly neglected in the secondary literature for many years. However, there has been a recent renewal of interest in *The Structure of Behavior*. For many years, the best resource on the *Structure* available in English was a special issue of Research in Phenomenology published in 1980 (X, 1). This situation has recently improved significantly with the publication of two books that include excellent exegeses of *The Structure of Behavior*: Evan Thompson's *Mind in Life* (2007) and Ted Toadvine's *Merleau-Ponty's Philosophy of Nature* (2009). I take the reading of the *Structure* that I will offer here to be broadly compatible with those of Thompson and Toadvine. (For other recent work on *The Structure of Behavior*, see Buchanan 2008; Low 2004.)

6 I am here disputing the influential reading of Renaud Barbaras (2001), which claims that Merleau-Ponty does not question this view of nature until much later in his career. For example, Barbaras writes that "it is only very late in Merleau-Ponty's work that the concept of nature becomes the object of a separate reflection. Until 1956–57 Merleau-Ponty utilized this notion in a non-critical way and conferred upon it the current philosophical meaning. Thus, *The Structure of Behavior* opens with these words: 'Our goal is to understand the relations between consciousness and nature: organic, psychological or even social. By nature we understand here a multiplicity of events external to each other and bound together by relations of causality.' This certainly is the classic conception of nature, common to Descartes and Kant, which Merleau-Ponty retains here, even if, to be sure, he inquires at the same time into the possibility of the upsurge of consciousness in the midst of this nature" (2001, p. 23).

This reading amounts to a denial of the *dialectical* character of the *Structure*: it takes Merleau-Ponty's initial formulation of the question to be a final and definitive statement, rather than the beginning of an investigation that will be transformed by its own results. I suspect that this way of reading the *Structure* stems from the view—widely held among Merleau-Ponty commentators—that Merleau-Ponty did not become concerned with *ontological issues* until his last, unfinished work, *The Visible and the Invisible*. My reading of the *Structure* will show, on the contrary, that Merleau-Ponty was dealing with ontological questions from the very beginning of his career—even though he does not begin to use the term "ontology" until later.

We find a more nuanced and sympathetic reading of the *Structure* in Barbaras' more recent work, such as "A Phenomenology of Life" (2005). Here Barbaras sees in the *Structure's* focus on *life* important anticipations of Merleau-Ponty's later ontology, and especially of the questions Merleau-Ponty grapples with in his lectures on *Nature*.
Might mechanistic science have missed the definition of objectivity? Might the cleavage between the subjective and the objective have been badly made; might the opposition between a universe of science—entirely outside of itself—and a universe of consciousness—defined by the total presence of self to self—be untenable? (SB 10/8)

Merleau-Ponty pursues these questions by returning to the phenomena themselves—specifically, the phenomenon of animal and human behavior [comportement]. Again, this choice is far from accidental. The living body is the site of all the aporia of Modern thought: it is both a natural phenomenon and the seat of consciousness, the place where subject and object meet, mix, and become confused. It is in human behavior that we witness the appearance of consciousness in nature. Thus it is here that we must look to see if our received concepts of nature and consciousness, body and mind, are empirically justified. As Merleau-Ponty puts it, behavior "taken in itself… is neutral with respect to the classical distinctions between the 'mental' and the 'physiological' and thus can give us the opportunity of defining them anew" (SB 4/2).

In order to avoid prejudicing the issue, Merleau-Ponty begins from within the "universe of science," studying the phenomenon of behavior as it appears to the psychologist and the physiologist. This is a methodological necessity. If the conception of nature that science begins with is inadequate to the natural phenomena it studies, then this must be evident from within the scientific perspective itself. To challenge the assumptions of natural science from any point of view other than its own would merely be to oppose one perspective against another, with no way of adjudicating between them. But this juxtaposition of incompatible perspectives, each valid in its own right, is precisely the aporia that Merleau-Ponty is trying to resolve. To criticize natural science from the point of view of philosophy, or that of "lived experience," would only be to re-
inscribe existing dualisms, or generate new ones. Thus the only way to criticize the
assumptions of the sciences is to do so on from within, on their own terms. Merleau-
Ponty's argument in the *Structure* is that biology and psychology have been *insufficiently empirical*: that they have maintained, contrary to their own principles, certain
metaphysical presuppositions that cannot be justified empirically. Specifically, Merleau-
Ponty argues that they have uncritically inherited the Cartesian view that nature in
general, and the living body in particular, are *machines*.

The essence of the machine is its *decomposability*: to say that nature is a machine
is just to say that it can be analyzed into independent parts. From a "naïve," pre-
scientific point of view, behavior—both human and animal—appears to be a coordinated,
goal-directed activity that responds to the meaning of its situation in a more-or-less
intelligent way. If the living body is a machine, however, then its behavior must admit of
a mechanical explanation: both behavior and its causes must be decomposable into
simple parts in such a way that the same elementary cause always produces the same
elementary effect.

To explain nerve functioning can only be to reduce the complex to the simple, to
discover the constant elements of which behavior is constituted. Thus one would
decompose the stimulus as well as the reaction until one encountered the
'elementary processes' composed of a stimulus and a response which were always
associated in experience. (SB 11/9).

---

7 Merleau-Ponty's description of these parts as "outside" of one another is a reference to Descartes' famous
definition of extension as "parts outside of parts" (*partes extra partes*). The exteriority in question is not
just spatial, but *ontological*: the parts of the machine are not only spread out in space, but also exist
independently of one another. See Descartes' letter to Henry More (5 February 1649): "'I call extended
only what is imaginable as having parts outside of parts" ["ita illud solum quod est imaginabile, ut habens
partes extra partes, . . . dico esse extensum"] (CSMK Vol. 3, p. 362; AT V, p. 270). (I have modified the
English translation, which for some reason renders "partes extra partes" as "parts within parts.") Merleau-
Ponty cites this passage in his lectures on The Incarnate Subject (IS 50/33).
This was precisely the project of the reflex theory of behavior, which dominated psychology and physiology in Merleau-Ponty's time. According to this theory, "the order in the reflex—that is, the adaptation of the response to the stimulus and the coordination of partial movements in the total gesture—is assured by pre-established connections from the sensible surface to the effector muscles" (SB 10/8, my emphasis). Thus the apparent unity and goal-directedness of behavior are mere illusions produced by the cunning design of the body-machine:

If behavior seems intentional, it is because it is regulated by certain pre-established nerve pathways in such a way that I in fact obtain satisfaction. The 'normal' activity of an organism is only the functioning of this apparatus constructed by nature; there are no genuine norms; there are only effects. (SB 9/7)

Consciousness, on this view, plays no causal role in human behavior; it can only be an illusion or an epiphenomenon—an incidental side-effect of the body's mechanical activity. If this conclusion is counter-intuitive then our intuitions must be at fault, for no alternative view is possible within the ontology of mechanistic science:

As soon as one ceases to place confidence in the immediate givens of consciousness and tries to construct a scientific representation of the organism, it seems that one is led to the classical theory of the reflex—that is, to decomposing the excitation and the reaction into a multitude of partial processes which are external to each other in time as well as in space. … The classical theory of the reflex and the methods of realistic analysis and causal explanation, of which the reflex theory is only an application, alone seem capable of constituting an objective and scientific representation of behavior. (SB 8-9)

The reflex theory of behavior suggested a clear research program: by systematically subjecting an animal body to one stimulus at a time and observing the response produced, one ought to be able to map the anatomical connections between the body's sensory surfaces and its muscles, and in so doing to generate a predictive theory of
that animal's behavior. However, when this research program was put into effect, it yielded quite unexpected results. The "elementary reflex" which was supposed to be the basic unit of behavior—a single stimulus that always produces the same simple response—turned out to be largely mythical. Such constant conjunctions were hardly to be found, even under the most artificial laboratory conditions. Instead, experiment revealed that the effect of a given "elementary stimulus" cannot be isolated from its context. That is, the animal's response depends on the other stimuli that are present at that moment; on the past stimuli to which the animal has been subjected; and on the prospective activity in which the animal is presently engaged. Despite these discoveries, most psychologists refused to abandon the reflex theory, instead attempting to prop it up with an ever-growing number of auxiliary hypotheses. To abandon the reflex theory would have meant abandoning the tried-and-true methods of realistic analysis and causal explanation, leaving psychology with no way to explain behavior: "If order cannot be based on pre-established anatomical structures, from where does the coherence of our reactions and their adaptation to the stimulus come?" (SB 32/32).

However, there were a few psychologists who argued that these experimental results demanded new concepts and new methods. In order to describe the observed facts of animal behavior, these psychologists introduced the concept of the Gestalt, which Merleau-Ponty translates as form (forme): a whole that cannot be decomposed into

---

8 "[T]he classical theory of nerve functioning is led by the force of things to burden itself with auxiliary hypotheses which are almost in contradiction with it, just as the Ptolemaic system revealed its inadequacy by the large number of ad hoc suppositions which became necessary in order to make it accord with the facts" (SB 16/15).
independent parts. The failure of the reflex theory demonstrated the existence of form at three levels: perception, behavior, and anatomy. If the effects of one stimulus cannot be isolated from those of another, then perception cannot be decomposed into a collection of elementary stimuli. If there are no elementary stimuli—i.e. no stimuli that invariably produce the same motor response—then behavior cannot be decomposed into a collection of elementary reflexes. And if there are no elementary reflexes, then the nervous system cannot be decomposed into a collection of autonomous circuits connecting sensors on the body's surface to effector's in the body's musculature. Thus the living body is not a machine that can be decomposed into real elements; and the order of behavior—its coherence and its adaptation to the body's situation—cannot be explained by pre-established anatomical structures.

If the decomposability of the machine implied a mechanical understanding of causality, then the existence of form implies a non-mechanical causality, which Merleau-Ponty calls "dialectical" (SB 160/174). The machine was defined by the exteriority—spatial, ontological, and causal—of its parts: each could come to be, change, or pass away without affecting the others. In a Gestalt, on the contrary, a change to one part alters every other part. "We will say that there is form whenever the properties of a system are modified by every change brought about in a single one of its parts and, on the contrary, are conserved when they all change while maintaining the same relationship among themselves" (SB 47/50). Thus the Gestalt exhibits a circular causality between part and whole: "The genesis of the whole by composition of the parts is fictitious. It arbitrarily breaks the chain of reciprocal determinations" (SB 50/53). In other words, a Gestalt must
be self-regulating or self-organizing (Morris 2006b). It is precisely this property that allows it to explain the order of behavior.

"What troubles and compromises anatomical conceptions is that they cannot easily introduce a regulation of the reflex either by the situation to which it responds or by its proper effects" (SB 35/35). For the reflex theory, the adaptiveness of behavior can only be explained in terms of blind mechanisms that are fixed in advance.

But from the point of view of this theory this adaptiveness is not a property of the actions themselves, but is instead a mere impression which they give to the onlooker. The actions are not determined in any way by the intrinsic nature of the situation, but altogether by these pre-existing bond-devices. The situation enters only as an agency which turns the key, presses the button, makes the machine go. But, like a true machine, the animal can only act according to the system of pre-established bonds, whether such an action be adequate to the circumstances or not. The relationship between situation and response is consequently purely contingent. (SB 35/35)

As a result, the reflex theory is incapable of explaining the flexibility and creativity of behavior—its capacity to respond dynamically and intelligently to situations it has never before encountered: "A machine is capable only of operations for which it has been constructed; the idea of a machine which would be capable of responding to an indefinite variety of stimuli is a contradictory one, since automation is obtained only by submitting the initiation of work to certain chosen conditions" (SB 87/96).

The dialectical view of causality, on the other hand, allows us to explain the intelligence of behavior without appealing to pre-established mechanisms, by showing how behavior is self-regulating. At the level of perception, we saw that the organism perceives its situation as a Gestalt: a whole in which a change to any one part affects every other part. Thus the organism's situation has a total significance for it, which is indivisible; and its behavior can only be understood as a response to this significance.
When we move from perception to behavior, we find that the significance of the organism's situation is indivisible from its own activity. Perception is not a passive reception, but an active exploration: the body does not wait to be stimulated by its surroundings, but "stimulates itself" by moving. Thus perception is both the cause and the effect of movement; there is a circular causality between "stimulus" and "response" which makes it impossible to treat them as independent events. "The facts suggest, on the contrary, that the sensorium and the motorium function as a single organ" (SB 36/36). At the level of anatomy, then, we can no longer understand the nervous system as a collection of independent mechanisms which would be triggered by individual stimuli. Instead, the function of the nervous system would be to connect the organism's muscles to its sensory fields in such a way that movement could be regulated by its own perceptual results:

[W]e should consider the afferent sector of the nervous system as a field of forces which express concurrently the intraorganic state and the influence of external agents; these forces tend to balance themselves according to certain modes of preferred distribution and to obtain movements from the mobile parts of the body which are proper to this effect. These movements, as they are executed, provoke modifications in the state of the afferent system which in turn evoke new movements. This dynamic and circular process would assure the flexible regulation which is needed in order to account for effective behavior. (SB 46/48-9)

The concept of a self-organizing whole allows us to explain the intelligence of behavior without appealing either to fixed mechanisms or to mental representations (SB 127/138). The phenomenon of animal behavior thus appears as an intermediary between matter and mind: it exhibits an intelligence that is not yet self-consciousness, and a
meaning that is not yet an idea. The problem with which the *Structure* began is thus transformed.\(^9\) The relation between consciousness and nature is split in two by the appearance of behavior as a mediating term, and revealed as two distinct relations: one between inanimate nature and living behavior, and one between behavior and consciousness. The introduction of this intermediate term transforms the terms it mediates: if the living body is not a machine, then neither is the inanimate nature from which it emerges; and if consciousness emerges from behavior, then the mind cannot be a disembodied region of pure self-presence. The study of behavior thus leads Merleau-Ponty to a *Gestalt* ontology in which nature is organized into three distinct levels of organization: the physical, the vital, and the human. The key to this ontology is the concept of *form*: "Equally applicable to the fields which have just been defined, it would integrate them as three types of structures by surpassing the antinomies of materialism and mentalism, of materialism and vitalism" (SB 131/141).

We find already within the physical order (i.e. inorganic nature) the existence of *form*. It is this which has led physicists to introduce the concept of the *field*:

The notion of form which was imposed upon us by the facts [of behavior] was defined like that of a physical system, that is, as an ensemble of forces in a state of equilibrium or of constant change such that no law is formulable for each part taken separately and such that each vector is determined in size and direction by all the others. Thus, each local change in a form will be translated by a redistribution of forces which assures the constancy of their relation; it is this internal circulation which is the system as a physical reality. And it is no more composed of parts which can be distinguished in it than a melody (always transposable) is made of the particular notes which are its momentary expression. (SB 137/147-8)

Thus it is not only in psychology and biology that we are forced to abandon the mechanical ontology; in physics too, we must move beyond the atomistic view that

---

\(^9\) Compare Toadvine 2009, p. 39.
nature can be decomposed into "elements or particles invested with absolute properties" (SB 138/148). If nature is not a machine, then we must abandon the notion of a clockwork universe whose every motion would be determined in advance by immutable natural laws.

It can happen that, submitted to external forces which increase and decrease in a continuous manner, the system, beyond a certain threshold, redistributes its own forces in a qualitatively different order which is nevertheless only another expression of its immanent law. Thus, with form, a principle of discontinuity is introduced and the conditions for a development by leaps or crises, for an event or for a history, are given. (SB 137/148)

If nature is self-organizing then the laws of nature are not fully given in advance. Rather, systems will generate their own laws, and there will be a circular relation between the laws that govern a physical system and the novel structures the system gives rise to:

[E]ach form constitutes a field of forces characterized by a law which has no meaning outside the limits of the dynamic structure considered, and which on the other hand assigns its properties to each internal point so much so that they will never be absolute properties, properties of this point. (SB 137-38/148)

The living body would be one such novel structure. Life emerges from inorganic nature as a new kind of system with a new form of organization. Behavior qua movement is a form spread out in time and space; but at the same time, behavior qua perception is the perception of form. In other words, behavior is a form that is sensitive to the forms around it. However, the organism's perceptions and behavior are not simply or fully determined by the physical form of its surroundings:

On the contrary, as we have seen, physical stimuli act upon the organism only by eliciting a global response which will vary qualitatively when the stimuli vary quantitatively; with respect to the organism they play the role of occasions rather than of cause; the reaction depends on their vital significance rather than on the material properties of the stimuli. Hence, between the variables upon which conduct actually depends and this conduct itself there appears a relation of
Thus behavior is a folding back of physical form upon itself which gives rise to a new kind of form: a pre-discursive, non-representational meaning that Merleau-Ponty will later call "motor significance" (PP 144/178).

It is out of this vital behavior and its pre-discursive significance that the mental or distinctively human order emerges. Just as perception and behavior were natural forms in which form itself became thematic, giving rise to a form of the second order, self-consciousness emerges when behavior begins to thematize its own meaning, generating a meaning of the second order—a symbolic meaning (SB 120-22/130-33). What distinguishes human behavior from that of other animals is our capacity to reflect on the meaning of our own behavior, and express this meaning in "symbolic behavior." Vital behavior already expresses the vital significance of the organism's situation. But it is only in symbolic behavior that this expression becomes the point of behavior—that behavior's expressive character becomes explicit for itself. Thus it is here that the question of truth—of the adequacy of expressive behavior to that which it is trying to express—arises for the first time (SB 122/133). The emergence of this new perspective transforms the human body's whole orientation toward its world. For the first time, the body is able to take up an outside perspective on its own situation; to see the world as a collection of independent objects, and its own body as one of these objects.

At the same time, however, self-consciousness remains a form of behavior—a way of moving and perceiving. As such, it is intrinsically visible: "The mental thus
understood is comprehensible from the outside" (SB 183/198). The mind is not a realm of pure interiority, shut up in itself and in principle inaccessible to others. On the contrary, what we call consciousness is an expressive, bodily activity that takes place in the world, and the meanings it expresses are there to be seen in the way that the human body acts and responds to its situation. However, this means that the mind is not transparent to itself: self-consciousness takes the form of a behavior that is oriented toward the meanings it itself is enacting in the world, but this reflection is always a work in progress—a project that can never be accomplished once and for all. It follows that we have no privileged introspective access to the meaning of our own actions. But this is not to say that we must reject introspection in favor of external observation. "There is no reason either to reject introspection or to make it the privileged means of access to a world of psychological facts" (SB 183/198). If I am not transparent to myself, than you and I are in the same position when it comes to interpreting my behavior:

The child who is supposed to say which colors appear similar to him and the monkey which has been trained to put all the slugs of the same color in a saucer are in the same situation. Nothing is changed when the subject is charged with interpreting his reactions himself, which is what is proper to introspection. (SB 183/198).

To know oneself is no easier than to know another. Indeed, it often happens that others understand better than I do the meaning of my own conduct. A mind that was completely transparent to itself would be completely opaque to others, and they to it; the price of total self-certainty is solipsism. Only a being that does not fully coincide with itself can be open to others, and to a world: that being is not consciousness, but the living body.
With this, Merleau-Ponty moves decisively beyond any understanding of phenomenology as an introspective method or a turn away from objective reality to first-person experience. Far from taking the distinction between subjective meanings and objective facts for granted, phenomenology demonstrates precisely that this distinction is unfounded—that the scientist cannot explain the facts of behavior without taking meaning into account. Thus the basic question of phenomenology, for Merleau-Ponty, is to understand how meaning arises within nature. This is not a matter of showing how meaning is always already given in advance, but rather of understanding how nature generates meaning: how sense arises from non-sense, or as Merleau-Ponty will later write, how nature can be "the autoproduction of a meaning" (N 3/19).

Unlike the Gestalt psychologists, who ultimately betrayed their own central insight by attempting to reduce mental and vital forms to physical ones, Merleau-Ponty insists on the originality of each ontological order with respect to that from which it emerges (Toadvine 2009, pp. 32-33). The emergence of life in nature or that of mind in life constitutes the appearance of a genuinely new kind of form, which operates according to its own distinctive logic and requires its own endogenous mode of explanation. "The advent of higher orders, to the extent that they are accomplished, eliminate the autonomy of the lower orders and give a new signification to the steps which constitute them" (SB 180/195). Thus Merleau-Ponty distinguishes his position from any kind of physicalism or reductive naturalism. At the same time, however, he is careful to distinguish his views from those of "critical" (i.e. Neo-Kantian) philosophy, which attempts to flatten out

---

10 This insight in The Structure of Behavior, which Merleau-Ponty gains from Gestalt psychology (Embree 1980), is crucial to his interpretation and appropriation of Husserl's phenomenological method in The Phenomenology of Perception (PP 57-60/84-87).
reality in the other direction by reducing the physical and vital orders to the mental.

"Lower" forms are never completely assimilated into "higher" ones: there remains always an element of opacity, an excess of the lower over the higher. The integration of lower forms into higher ones is always threatened by an inevitable dis-integration: mind is always disintegrating back into "mere" life through fatigue, illness, or injury; and life is constantly disintegrating back into "mere" matter in death.

T]he notions of soul and body must be relativized: there is the body as mass of chemical components in interaction, the body as dialectic of living being and its biological milieu, and the body as dialectic of social subject and his group; even all our habits are an impalpable body for the ego of each moment. Each of these degrees is soul with respect to the preceding one, body with respect to the following one. The body in general is an ensemble of paths already traced, of powers already constituted; the body is the acquired dialectical soil upon which a higher 'formation' is accomplished, and the soul is the meaning which is then established. (SB 210/227)

Although it is clear already in *The Structure of Behavior* that Merleau-Ponty wishes to avoid falling into the idealism of critical philosophy, he does not arrive at a decisive refutation of this position until his second book, the *Phenomenology of Perception*. In the *Structure*, realism (i.e. reductive physicalism) and idealism (i.e. reductive intellectualism) appear as two alternatives to be avoided. In the *Phenomenology*, however, Merleau-Ponty comes to the pivotal realization that these two seemingly opposed positions are in fact one and the same. Far from opposing the mechanical ontology of the realist, idealism simply takes it to its logical conclusions. If nature is a machine then it possesses no intrinsic unity and no endogenous meaning. Only the mind of the scientist can synthesize the disconnected parts of nature into an intelligible whole.
PREFACE

We began from a world in itself that acted upon our eyes in order to make itself seen by us; we have arrived now at a consciousness or a thought about the world, but the very nature of this world is unchanged. It is still defined by the absolute extiriority of its parts and is merely doubled across its extension by a thought that sustains it. We pass from an absolute objectivity to an absolute subjectivity, but this second idea is worth only as much as the first, and only finds support in contrast to the first, which is to say, through it. The kinship of intellectualism and empiricism is in this way much less visible and much more profound than is believed (PP 41/64).

Thus the refutation of mechanism is equally a refutation of transcendental idealism. In showing that the living body is not a machine, Merleau-Ponty has also demonstrated that the mind is not a constituting consciousness. "[B]ehavior is not a thing, but neither is it an idea. It is not the envelope of a pure consciousness and, as the witness of behavior, I am not a pure consciousness. It is precisely this which we wanted to say in stating that behavior is a form" (SB 127/138). The discovery of form thus opens up a new ontological field which is not reducible to the antinomies of Modern philosophy.

It must be admitted, however, that this new ontology remained elusive, despite the significant progress made by Merleau-Ponty before his untimely death. The concept of form remained for him more the index of a problem than a solution. This problem manifests itself in The Structure of Behavior as a problematic ambiguity in Merleau-Ponty's account of the relation between form and consciousness, or matter and mind (Toadvine 2009, pp. 24, 38; Morris 2006b). I noted above that Merleau-Ponty is careful to distinguish his own position from transcendental idealism by insisting that form does not require a consciousness to constitute it. However, he argues that physical form is a perceptual being, "conceivable only as an object of perception" (SB 144/156). This claim
must be understood in the context of Merleau-Ponty's argument that perception is a
*bodily* rather than an intellectual activity: to say that form can only be conceived as an
object of perception is *not* to take the idealist position that form only exists as an object of *thought*. Nevertheless, Merleau-Ponty's formulation risks reinscribing the logic of
transcendental idealism at the level of the living body. In place of a constituting
consciousness projecting meaning onto nature, we would have a living body that
constitutes its environment as meaningful through its own movements. But we must still
ask: is this meaning simply projected onto the world by the moving body, or do the
body's movements express a sense which is already nascent in inorganic nature, before
perception arrives on the scene? How can Merleau-Ponty claim both that life emerges
from inorganic form, and also that form exists only for perception?

This ambiguity is symptomatic of a fundamental problem: Merleau-Ponty's failure
to clarify the ontological status of form. The concept of form that Merleau-Ponty takes
over from Gestalt psychology defines it as a whole that cannot be reduced to the sum of
its parts. In a working note entitled "Gestalt" that Merleau-Ponty wrote near the end of
his life, he criticizes this as a "negative, external" definition: it says what form is *not*, but
does not succeed in explaining what it *is* (VI 204/255, translation modified). If form can
only be defined by contrast with the decomposable wholes of mechanistic ontology, then
it is only a criticism of this ontology, and not a genuine alternative. Merleau-Ponty never
abandoned the Gestalt ontology that he articulated in *The Structure of Behavior*, and he
never stopped searching for a positive account of form that would explain it on its own
terms. In his 1957-58 lecture course on *The Concept of Nature*, we find Merleau-Ponty
again posing the question: "How then are we to understand this relation of the totality to
its parts? What status must we give totality?" (N 145/194, translation modified). He calls this "a question which is at the center of this course on the idea of nature and maybe the whole of philosophy" (N 145/194).

In this course on Nature and its sequel the following year, Merleau-Ponty turns from psychology to biology in search of an answer to this question. As I noted above, Merleau-Ponty's criticism of the reflex theory in The Structure of Behavior already had implications for physiology and anatomy. The reflex theory sought to explain behavior in terms of pre-established anatomical structures; thus its failure suggested that the living body is not a machine whose anatomy could be decomposed into independent parts. In the Nature lectures, Merleau-Ponty takes up again the connection between behavior and anatomy, drawing on biologist G. E. Coghill's Anatomy and the Problem of Behavior. Coghill's studies of the axolotl salamander led him to argue that behavior is continuous with anatomical development: "if behavior is a mystery for a congealed anatomy, is because it can be understood only by a dynamic anatomy. …[T]he maturation of the organism and the emergence of behavior are one and the same thing" (N 144/192-3). The organism is not a machine, whose fixed architecture would determine its movements in advance. Rather, the organism is a dynamic form that is constantly being transformed by its own activities; thus growth is a slow behavior, and behavior a kind of second body, an extension of anatomy.

The problem of learning and behavior thus leads us back to the problem of growth or embryogenesis: how a complex and highly organized multi-cellular body grows from a single cell. The facts of embryogenesis, like those of behavior, pose a challenge to the mechanistic understanding of nature. Here, as in The Structure of Behavior, we discover
the existence of *Gestalts*: self-organizing wholes that cannot be decomposed into independent parts. Merleau-Ponty cites the embryological studies of biologist H. Driesch, who sought to determine the mechanisms behind the embryo's differential development: what causes one part of the embryo to develop into an arm and another part to grow into a leg, in such a way that the whole body develops into an organized, functional whole rather than a monstrous mishmash of body-parts? (N 230ff/293ff). What Driesch discovered was that the developmental destiny of any given cell in the growing embryo depends on its place within the whole. An embryo that is divided in two will not grow into two halves of an organism, but two smaller organisms, each fully formed (N 230/293). The growing body exhibits a *circular causality* in which each part is a possible cause of morphogenesis and differentiation for every other part. Thus the form of the adult body is not given in advance in any one of its parts, but arises only through the interaction of the parts—an interaction that is spread out over space and time.

Faced with the failure of mechanistic explanation, however, scientists and philosophers alike have often resorted to vitalism and teleology. Thus Driesch appeals to an unknown "factor E" (for "entelechy") which would explain the robust self-regulation of embryonic development. This move is parallel to the move from mechanism to idealism that Merleau-Ponty criticizes in his early works (Morris 2008): if the form of the organism cannot be explained away by dissolving it into independent parts, then it must be given in advance in some idea or *telos* which would organize the parts of the organism from the outside, thus ensuring the development of a species-typical form. This move explains nothing; it is merely a "retrospective illusion that makes us project what is yet to come into the past, or double the sensible world with an intellectual world, which adds
nothing to our understanding" (N 152/203, translation modified). Mechanism cannot explain the *robustness* of development—its apparent goal-directedness, the way it produces species-typical outcomes even in the face of external interference. But vitalism and finalism cannot explain the *contingency* of development—its dependence on external conditions and its capacity for bizarre "mistakes." The problem, once again, is to explain the relation of the whole to its parts without denying the existence of the whole on the one hand, or making the whole into a separate and pre-existing entity on the other. The organism is not located at a single point in space and time, but it is also not aspatial or atemporal. It is material without being mechanical, meaningful without being ideal.

Growth, like behavior, exhibits a kind of intelligence which is not that of a human mind; nature generates complex, organized forms in an entirely different way from that of human engineering. "Human teleology constructs by assembly of machines, whereas the organism does it by auto-differentiation…. There is not the work of an exterior form on a mechanical cause, but the work from within to without, by growth and differentiation" (N 182/238). The natural phenomenon of the living body demonstrates the insufficiency of both mechanism and finalism. However, the sciences of Merleau-Ponty's time did not yet have a concept of form that could explain the self-organization of the organism. Thus Merleau-Ponty was able to show the shape that a solution must take, and the errors that must be avoided; but his turn to biology did not ultimately yield the solution he was seeking.

The present study takes up Maurice Merleau-Ponty's unfinished project of developing an ontology of nature whose concepts are drawn from the phenomenon of
life, rather than from human *techne*. Merleau-Ponty is a phenomenologist, and his philosophical project is a phenomenological one. To take up his project for ourselves, then, cannot be merely to study the things he wrote. Rather, we must return to the phenomena themselves that Merleau-Ponty studied, and to which he constantly drew our attention. This is what I have attempted to do in this dissertation. Rather than offering a commentary on Merleau-Ponty's texts, I have tried to build on his insights by studying for myself the phenomena of behavior, perception, and development. It would be strange if in returning to these phenomena, we were to limit ourselves to the scientific evidence available during Merleau-Ponty's life. This would be to treat Merleau-Ponty's thought as a museum-piece, rather than a living project of thought. Merleau-Ponty took up the latest scientific developments of his time—both to learn from them and to criticize them—and we must do the same if we wish to follow his method. The science of complexity and self-organization has made significant progress since Merleau-Ponty's death, and it is to this science that I will turn in search of a new, positive account of form. At the same time, the mechanistic theories that Merleau-Ponty criticized—notably the reflex theory in psychology—have been supplemented and supplanted by new theories: Neo-Darwinism in biology and Cognitivism in psychology. Merleau-Ponty presciently noted in *The Structure of Behavior* that the refutation of mechanism can never be completed, "since behaviorism can always invent other mechanical models with regard to which the discussion will have to be recommenced" (SB 127/138). To take up Merleau-Ponty's project today, then, we must recommence this discussion and show that these new theories, too, are inadequate to the phenomena they seek to explain.
I also follow Merleau-Ponty's method in *The Structure of Behavior* of beginning from within the universe of science, rather than from an explicitly philosophical or transcendental standpoint. There are two compelling reasons for proceeding in this way. The first is that it allows us to criticize science on its own terms, rather than opposing ourselves to it from the outset. If Modern science has been held back by mechanistic prejudices that are not adequate to natural phenomena, then it should be possible to demonstrate this empirically, from within the scientific standpoint. The present gulf between science and philosophy is more than a simple division of labor; it is an artifact of the Modern oppositions between subject and object, materialism and idealism. In overcoming mechanistic ontology, we will also overcome the artificial division between science and philosophy—which means arriving not only at a more self-conscious science, but also at a more empirical philosophy. This brings us to the second reason for beginning from within the scientific standpoint, which is to learn from science what nature is and how to understand our place within it. It would be easy to simply begin from the standpoint of reflection, rather than demonstrating through long and painstaking inductive research that science must become self-reflective. If we make the transcendental turn prematurely, however, we will arrive straightaway at an idealism which is only the obverse of scientific realism. The mind that we "discover" by pure reflection is as much a product of our mechanistic prejudices as the material world discovered by the scientist. By proceeding empirically, however, we will learn that neither the body nor the mind are what we thought they were. It is not within ourselves but in nature that we will discover the truth of who we are.
My approach in this dissertation has been influenced significantly by the work of Renaud Barbaras and Evan Thompson. Both Barbaras and Thompson draw on Merleau-Ponty’s work in their own attempts to re-think the relation between consciousness, life, and nature. Both combine Merleau-Ponty’s ontology with some further account of the living body: Barbaras (2003, 2008, 2010) takes up Hans Jonas’ account of life as metabolism, while Thompson (2007) carries on the account of life as autopoiesis developed by his teacher, Francisco Varela. These thinkers have in common a project of understanding consciousness by studying the phenomenon of life. My hope in this dissertation is to make some contribution to this project. However, I differ somewhat from these thinkers in my approach to the phenomenon of life. Both Jonas and Varela attempt to answer the question: "What is life?" That is, they attempt to distinguish what is living from what is not. I share their interest in this question, but I do not believe that we are presently in a position to answer it directly. The question of life has become hopelessly obscured in the Modern era by the mechanistic view of nature. As a result, we cannot hope to understand the difference between life and non-life without first understanding the difference between the living body and the machine. When we move directly to the more originary question of life without first grappling with the question of mechanism, our thought remains entangled in mechanistic concepts and the essence of life eludes our grasp.

For Jonas, the defining characteristic of life is metabolism: "the process through which a form maintains itself as identical through a continuous renewal (replacement) of matter" (Barbaras 2003). The living body depends for its continued existence on the exchange of matter with its surroundings. But this means that the organism cannot be
identified with the matter that composes it at any particular time. On the contrary, the organism can only maintain its identity by changing the matter that composes it. "In the realm of the lifeless, form is no more than a changing composite state, an accident, of enduring matter" (Jonas 2001 [1966], p. 80). It is only in the emergence of life that the difference between matter and form becomes a concrete reality. Metabolism represents "the emancipation of form… from the immediate identity with matter" (Jonas 2001 [1966], p. 81). However, the organism's freedom from matter, its non-identity with the stuff that composes it, is also at the same time a dependence on matter. For the organism can be more than this material only by constantly taking in new material. Its unique power of changing its own matter is equally the necessity that it do so (Jonas 2001 [1966], pp. 83-4).

Jonas' study of the living body's way of being pushes the traditional concepts of form and matter to their breaking point—but never goes beyond them. He does not recognize that these categories are alien to the phenomenon of life, transplants from the realm of human techne. As a result, he never fully succeeds in refuting the reductionist view which holds that the form of the living body is not an efficacious reality, but merely an epiphenomenon of the underlying movements of atomic matter. Jonas argues that a mathematician God would be unable to recognize the phenomenon of life at all (2001 [1966], p. 87). Against this view, Jonas can only offer the scientist's own experience of herself as living. But the reductionist is perfectly happy to accept this contradiction of her own experience, and to explain the phenomenon of life by explaining it away. Evidently, it is not enough to show that the consequences of reductionism are absurd from the standpoint of lived experience. We must show that reductionism fails on its own terms—
that is, that it does not succeed at explaining the phenomena it sets out to explain, and
that even a mathematician God would be forced to admit into his mathematical
explanations the existence of *forms* or molar entities that cannot be located at a single
point in space and time. Otherwise, we remain mired in the dualism between atomic
matter and macroscopic form, mathematical science and lived experience.

We encounter similar difficulties in Thompson's *Mind in Life* (2007). My own
turn to the science of complexity was inspired in part by Thompson, who uses dynamic
systems theory to bolster Merleau-Ponty's claim that inorganic nature already exhibits a
non-mechanistic, dialectical causality which makes the emergence of life possible (2007,
Ch. 4). However, Thompson does not take the critique of atomism far enough: he argues
against compositional reductionism, but he offers no new ontology to take its place. As a
result, many of the old ontological assumptions remain operative. This is most apparent
in Thompson's account of behavior and perception or "sense-making." For Thompson, as
for Jonas, the self-organization of the organism brings forth not only an identity, but also
a perspective on the world: its concern for its own continued existence endows things and
events with a *valence* or *significance* (helpful or harmful, food or poison). "An organism's
environment is not equivalent to the world seen simply through the lenses of physics or
chemistry" (2007, p. 153). Following Varela, Thompson describes the difference between
the "physicochemical world" and the meaningful environment of the organism as a
"surplus of significance" (2007, p. 154).

We find here the same problematic ambiguity that we noted in *The Structure of
Behavior*. This is not surprising, since Thompson takes over Merleau-Ponty's ontology
without clarifying the ontological status of form. What is the relation between the so-
called "physicochemical world" and the meaningful environment of the organism? Is the
"surplus of meaning" simply a projection of the organism upon the indifferent world of
atomic matter? If so, then perception never really knows the natural world from which it
emerges. Because Thompson does not arrive at a new concept of form, we are left with
the old concept of matter. Mechanism sneaks back in, under the guise of
"physicochemical laws," and the dualism of mind and body reappears as a dualism of life
and matter.

In order to avoid falling into similar perplexities, I have chosen in the present
study to set aside the question of what distinguishes the living from the non-living—
along with that other most perplexing question, the difference between the human and the
animal—in favor of a less primordial but more urgent and tractable question: what is the
difference between a living body and a machine? This question will force us to face head-
on the mechanistic and atomistic prejudices that are so deeply entrenched in Modern
thought. And it will also lead us to a phenomenon that Thompson and Barbaras largely
ignore: the phenomenon of development. The autopoietic account of life takes the
*individual cell* as its paradigm; the multi-cellular organism is then understood by analogy
to the single cell, without any account of how the former develops out of the latter. By
studying the phenomenon of embryogenesis, the present study will bridge the gap
between these two levels in a phenomenological—which is to say, empirical—way.

Merleau-Ponty began his career by studying behavior, and arrived only later at the
study of embryogenesis. I have chosen, however, to follow the development of the living
body, rather than the development of Merleau-Ponty's thought. Thus I begin in Chapter
One by studying the growth of the living body, in order to demonstrate the inadequacy of
our current, techno-logical concepts of form and matter. In Chapter Two, I turn to the emerging science of complexity for a new concept of form that does not depend on the traditional opposition between form and matter, and which will allow us to explain the facts of embryogenesis for the first time. Finally, in Chapter Three, I extend this account of anatomical development to the development of movement and perception. Although I have not attempted to follow Merleau-Ponty's thought chronologically or to duplicate his results exactly, I often found myself echoing Merleau-Ponty's arguments, confirming his conclusions, and expanding on or nuancing some of his key claims. I have tried throughout to indicate these points with footnotes citing relevant passages in Merleau-Ponty's texts, though the reader familiar with Merleau-Ponty's work will no doubt spot even more parallels than I have marked.

This dissertation is the fruit of a double apprenticeship. I have attempted to follow in the footsteps of Maurice Merleau-Ponty by following the examples of two living philosophers: Kym Maclaren and David Morris. I would like to take this opportunity to thank Prof. Maclaren, who has been a teacher, role model, and friend to me for over a decade. I am also indebted to Prof. Maclaren for introducing me to John Russon, whose summer seminars I have been privileged to attend for many years. I would like to thank Prof. Russon and the members of these seminars for showing me what philosophy can be, and what can be achieved when people think and talk together about the things that matter.
It was through Prof. Russon's seminars that I first met David Morris. I was extremely fortunate to find myself living in the same city as Prof. Morris just at the time that I decided to write my dissertation on Merleau-Ponty. I am immensely grateful to him for taking me on as a student and advisee, and for his generosity and support over the last four years. The influence of his writing and teaching can be found on every page of this work.

This dissertation would not have been possible without the unflagging support of Jeffrey Bloechl. I am grateful to him for his encouragement and his confidence in my work. I am also grateful to my friends and colleagues in the Philosophy Department at Boston College, from whom I have learned a great deal.

Finally, none of this would have been possible without the love and support of my family: my parents, Bill Brender and Rosanne Moss; my sister, Emma Moss Brender; and my partner, Marianne Pelton, to whom this work is dedicated.

I would also like to acknowledge the financial support of the Social Sciences and Humanities Research Council of Canada, as well as the Graduate School of Arts and Sciences at Boston College.
INTRODUCTION: Returning to life

A. Descartes' legacy

We find ourselves today in a peculiar position with respect to the Cartesian legacy. We pride ourselves on having done away with Descartes' immaterial mind, replacing his incoherent dualism with a thoroughly consistent materialism. Advances in science since Descartes' time have allowed us to locate in the brain those functions which Descartes argued no mere body could perform. And yet, in our rejection of Cartesian dualism, we have preserved a great deal of Cartesianism—far more than is usually acknowledged. In order to see how and why this is so, we must examine what has changed in our view of the mind and the body since Descartes.

Our central innovation, the one that allowed us to unify mind and brain, is the identification of thinking with computation. I will refer to this as the doctrine of computationalism.\(^{11}\) Computationalism, and the very idea of computation in its contemporary form, preceded the invention of the digital computer.\(^{12}\) Computationalism and the digital computer in fact share a common origin in the new mathematical logic developed toward the end of the nineteenth century. Logicians were interested in giving a formal, scientific account of how we establish one truth on the basis of another. Their goal was to explain the mysterious act of inference by breaking it down into steps that

---

\(^{11}\) This doctrine is sometimes referred to in the Philosophy of Mind literature as "machine functionalism", and in Cognitive Science as "Cognitivism", although these terms may refer to broader positions than the one identified here.

\(^{12}\) See the illuminating account provided by Jean-Pierre Dupuy (2009) in his history of the cybernetics movement.
Introduction

require no thought or insight.\textsuperscript{13} To this end, they invented symbolic languages in which propositions could be expressed without the ambiguity of ordinary language.\textsuperscript{14} Logical inference could then be modeled in these languages by the derivation of one symbolic proposition from another according to a limited set of deduction rules. These rules were strictly formal or syntactic: they took account only of the symbolic form of propositions rather than their content. With these tools in place, logicians could model domains of (usually mathematical) knowledge as systems of propositions derivable from a given set of axioms. It then became possible to prove rigorous "metalogical" theorems concerning the provability and unprovability of propositions or sets of propositions within these axiomatic systems.\textsuperscript{15} Some logicians, identifying correct inference with thought itself, claimed to be giving exact mathematical descriptions of the "laws of thought."\textsuperscript{16} Thus, they came to understand thinking as a purely formal manipulation of symbols according to strictly syntactic rules. It was this conception of thinking as algorithm that made it possible for us to conceive—as Descartes could not have—that machines could think.

The idea for the digital computer came from a 1936 paper by the logician Alan Turing, in which he tried to set out a purely theoretical definition of provability (i.e. derivability within a formal system). In order to capture the purely syntactic, "mechanical" nature of inference in a formal system, Turing imagined a machine that

\textsuperscript{13} Notice the form that this explanation takes: it is an attempt to analyze thinking into its smallest parts, on the assumption that these parts will not themselves involve thought. In other words, it is an attempt to explain away thinking, to explain it in terms of something other than itself. We will encounter this style of explanation again and again in the coming chapters.

\textsuperscript{14} The first such language was Gottlob Frege's Begriffsschrift or "conceptual notation," which he described in his 1879 book of the same name. The subtitle of this book was "A formula language of pure thought modeled on that of arithmetic."

\textsuperscript{15} It should be noted that these metalogical proofs are not themselves expressed in a purely symbolic language or arrived at through formal rules of inference; they are ordinary mathematical proofs, established and verified through human insight and expressed in "natural" (albeit technical) languages.
could read and write symbols on an infinitely long tape. Moving up and down the tape, the machine's next step (erasing, writing, or moving) would be entirely determined by the symbol it was currently reading and the contents of its memory (a finite array of stored symbols). The initial state of the tape would constitute the machine's input or program, and the final state of the tape, if the machine ever came to a halt, would constitute its output. Prior to Turing's paper, the word "computation" referred to calculations performed by a human being; if such computation were sometimes described as "mechanical," this meant only that it could be performed by rote, with little or no thought or understanding. Turing's innovation was to realize that the "mechanical computations" of inference in a formal system could actually be performed by a machine. Though Turing's machine was purely theoretical (its infinitely long tape making it impossible to build in reality), it did not take long for scientists to begin conceiving and building real machines inspired by Turing's model.

The attempt to give a purely formal model of thought by breaking inference down into thoughtless, mechanical steps thus led to the invention of machines that could carry out these steps. And this radical reconception of thought as mechanical is what finally broke down the wall Descartes had erected between mind and body. If machines can think, then we no longer need to posit an immaterial mind distinct from the material body. Instead, we can locate the machinery of thought in the body itself, specifically in the brain. Like a digital computer, the brain receives input from outside of itself, and responds with output. In the brain's case, this input comes from the body's sense organs,
and output goes to the body's muscles.\textsuperscript{17} The brain draws inferences about the world from the data transmitted to it by the senses, and on the basis of these conclusions it instructs the body's muscles to move appropriately. Thus, both knowledge and action are a matter of drawing correct inferences from the data that the brain receives as input. These inferences amount to a symbolic representation of reality, a constantly updated map of the world that informs the brain's outputs to the body. The correctness of our beliefs and the effectiveness of our actions depend entirely on the accuracy of this map.\textsuperscript{18}

Undoubtedly, this constitutes a major revision of the Cartesian position. And yet, computationalism nevertheless preserves some of the most important elements of Descartes' ontology. First, we take for granted that the human body and brain are machines, \textit{partes extra partes}—like the machines we build ourselves, though far more complicated. The preservation of this Cartesian understanding of the body is every bit as crucial to the new monism as its rejection of Descartes' immaterial mind. Second, thought for us remains in principle disembodied, just as it was for Descartes. The living body is inessential to the work of computation, and my brain could in theory be transferred to a different body without altering my identity. Even this particular \textit{brain} is not essential to who I am: my "program," my distinctive set of algorithms for deriving output from input, could in principle be transferred to some other computing machine without changing anything essential about me. Finally, thought remains representational. Just as an idea, for Descartes, is a sort of mental image that may correspond more or less well to the physical

\begin{footnotesize}
\textsuperscript{17} This arrangement is not essential to the nature of thought: the brain must receive input from somewhere, in order to perform its computations, but this input need not come from a living body, as our own artificial computers amply demonstrate.

\textsuperscript{18} This account of the nervous system's role in perception and behavior will be challenged in Chapter Three.
\end{footnotesize}
world, so is thought now understood as a symbolic representation of reality, inferred from the data of our senses.

The appeal of computationalism lies primarily in its claim to have unified mind and body, and its promise of a thoroughly naturalistic and materialist explanation of human consciousness. However, there is reason to wonder whether computationalism delivers on this promise, whether it has truly freed itself—and us—from Descartes' dualistic ontology.

B. The problem of meaning

Despite computationalism's appeal and the confidence of its adherents, it has not been without its critics. One particular criticism will be our focus here: the so-called "Chinese Room" argument from John Searle's famous paper, "Minds, Brains, and Programs" (1980). Searle proposes the following thought-experiment (pp. 417-418): imagine that you have been placed in a room with a large pile of papers on which Chinese characters are printed. Through a slot in the wall, someone occasionally inserts some more pieces of paper with Chinese writing on them. You have been provided with a detailed set of rules (in English) for correlating one set of papers with the other, based only on the shape of the Chinese symbols. No prior knowledge of Chinese is required to follow these rules: you simply look up the symbols on the papers that come through the slot, choose the characters that the rule-book calls for from your stock-pile, and push these papers out of the room through the slot. Unbeknownst to you, the papers being

---

19 Assume for the purposes of this example that you do not speak or understand Chinese.
20 I have simplified Searle's description slightly, without altering the essentials.
introduced into the room are questions written by native Chinese speakers, and the papers you are pushing out of the room are answers to these questions. The set of rules you are following is so sophisticated that to those outside, the room (or whatever is in it) appears to be carrying on a perfectly fluent conversation in Chinese. You are equally unaware of the fact that those who designed the room and wrote the rules that you are following consider the papers that come in through the slot "input," the papers you push through the slot "output," and the rules you are following a "program."

It should now be clear that this "Chinese Room" is a computer, albeit a rather odd one: instead of magnetic memory and a CPU made of silicone transistors, it is built out of stacks of paper and a human being. Nevertheless, the room functions in the same way that a digital computer does: it manipulates and responds to symbolic input according to purely syntactic rules. The point of Searle's thought-experiment is that even though the room appears, from the outside, to be speaking fluent Chinese, the human operator inside it has no understanding of the language, or of the conversation she is participating in. It follows, according to Searle, that even computers that can mimic human abilities and intelligence do not possess anything like human understanding, and do not explain anything about human understanding (p. 418). If it be objected that it is the system as a whole that understands Chinese, and not the person inside it, we can imagine that the person has memorized the rules, so that room and rulebook are done away with: now there is only a human being carrying on a conversation in Chinese by following a set of

---

21 Searle was responding in particular to the work of Roger Schank, who wrote a computer program that could respond in a human-like way to questions about stories it was told. Advocates of computationalism (or what Searle calls "strong AI") claimed that the computer running Schank's program literally understood the stories it was told, and could be used to explain how human beings understand stories and answer questions about them (p. 417).
purely formal, syntactic rules. Here again, Searle argues, although the person in question may appear to be fluent in Chinese, she actually understands nothing of what she is reading or writing; its content is completely opaque to her (p. 419). Searle concludes that "whatever purely formal principles you put into the computer, they will not be sufficient for understanding, since a human will be able to follow the formal principles without understanding anything" (p. 418).

This argument has been met with a number of objections, the principal one being that Searle hopelessly muddies the issues by inserting a human homunculus into the workings of his Chinese-speaking computer. One might defend Searle by arguing that he is merely returning to the roots of computationalism, when "computation" referred precisely to operations performed by rote by a human "computer". The point then would be to contrast such rote operations with intelligent, insightful action, and to argue that the latter cannot be reduced to the former. For example, the algorithm for long division will allow you to compute the right answer without understanding anything about arithmetic, but the correctness of this answer—and hence of the algorithm itself—can only be verified by someone with a non-formal understanding of division and its meaning.

The computationalist may still reply, however, that what we call "non-formal understanding" is really nothing but a more sophisticated algorithm whose workings we are not privy to. Indeed, critics of Searle's argument have argued that it relies on vague, introspective intuitions about what and how we understand, intuitions that may or may not be empirically accurate. The only objectively valid criterion for understanding, according to the computationalist, is the results it produces. If a computer's performance of a given task is indistinguishable from that of a human being, then the computer
understands the task just as well as the human does, regardless of what it "feels like" for either one.22

The answer to this criticism, and the true strength of the Chinese Room argument, lies in a further point that Searle makes about the meaning of the computer's inputs and outputs: "the formal symbol manipulations by themselves... are quite meaningless; they aren't even symbol manipulations, since the symbols don't symbolize anything. In the linguistic jargon, they have only a syntax but no semantics" (1980, p. 422, emphasis in original). The computer only appears to be thinking—to know and understand things about the world—because its inputs and outputs are symbolic, and thus appear to have a meaningful content. But symbols have no content in themselves, for in themselves they are not symbols at all, but only things—ink on a page, or pixels on a screen. They are meaningful only for beings who can interpret them as symbols, and discover a meaning in them. Searle's crucial point about the computer is that it is not such a being. The Chinese Room's inputs and outputs appear to its human observers to be meaningful Chinese sentences, but they have no such meaning for the computer itself. The "correctness" of the computer's outputs, its apparent fluency in Chinese, lies entirely in the interpretation given to these outputs by its human interlocutors. The Chinese Room itself is utterly incapable of distinguishing between correct and incorrect outputs, since for it these outputs are nothing but causal effects of physical inputs, the end of a complex chain reaction.23

---

22 This of course is the famous “Turing Test” of machine intelligence (Turing 1950).
23 The problem Searle identifies here has become known in cognitive science as "The Symbol Grounding Problem," after Stevan Harnad's (1990) article of the same name.
Introduction

In attempting to give an objective definition of understanding, the computationalist ends up attributing to the computer properties that are only in the eye of the observer. This can be seen even more clearly if we imagine that, instead of "conversing" with human interlocutors, the Chinese Room exchanges inputs and outputs with another, identical Chinese Room. There should be no temptation, in this scenario, to say that Chinese is being spoken or understood. There is here only a mechanical exchange of inputs and outputs, one computer triggering an automated response in the other, in a closed feedback loop. This is not a conversation. The computationalist might insist, of course, that to a (human, Chinese-speaking) observer it is indistinguishable from a real conversation. But this would again be to import into the situation an outside observer for whom the signs exchanged are meaningful. The Chinese Rooms are completely incapable of generating meaning on their own. As Searle puts it, you cannot get semantics from syntax.

We see here a general problem with any attempt to give an "objective" account of meaning or subjectivity. The computationalist is right to criticize introspective accounts, which would reduce the meaning of my situation to the meaning I think it has. However, we are no better off if we exchange the introspective standpoint for a purely external one. We will then arrive only at a description of what the situation means to the observer, when the whole problem was to describe (objectively) what it means to the system being observed. If we take the observer's perspective for granted then we only postpone the problem we set out to solve, for the observer is also a thinking being, and her perspective must also be accounted for. The claim that computers are thinking can only be sustained by appealing illicitly to the perspective of an observer who is not a computer, whose
Introduction

thought is more than an algorithm. In its attempt to give an *objective* account of thought, computationalism ends by emptying thought of all meaning.

**C. The meaning of life**

This critique of computationalism immediately raises a larger problem: if meaning cannot be generated by formal symbol-manipulation, then where does meaning come from? How are human beings able, as computers are not, to endow symbols with meaning? More broadly, what makes human conduct meaningful in a way that the computer's activity is not? In what follows, I will attempt to answer these questions by returning to the phenomenon of the living body. This phenomenon calls for both scientific explanation and philosophical reflection—perhaps more than any other phenomenon. The living body is nature at its most complex, its most astounding: nowhere else in nature do we find such a diversity of parts and forms working together in such harmony and coordination. It is this complexity and harmony that call for scientific explanation. At the same time, the living body is *the kind of being that we ourselves are*. If we wish to understand ourselves and our relation to nature, as philosophy does, then we must begin with the phenomenon of life.

In the following chapters, I will argue that the failures of computationalism stem from its preservation of Descartes' mechanistic ontology: the assumption that nature in general—and the living body in particular—are complex machines composed of mutually external parts. As we will see, this attempt to explain the natural world—including ourselves—using concepts drawn from human *techne* renders the phenomenon of meaning incomprehensible. Rather than imposing on the living body an extrinsic set of
Introduction

cancepts drawn from a foreign domain, my goal is to allow life itself to show us the ontology that is appropriate to it, and to generate the concepts we need to understand it. We will find that the phenomenon of life undermines the mechanistic ontology that underlies both Cartesian dualism and computationalism. And we will see, further, that the phenomenon of life is the key to understanding the genesis of meaning, which this mechanistic ontology has so thoroughly obscured. In allowing life to show us what it means, we will discover the very meaning of showing and meaning.
CHAPTER ONE: Is the living body a machine?

"[W]e cannot think Nature without taking into account that our idea of Nature is impregnated with artifice" (N 86/120, translation modified.)
1. What are living bodies made of?

It seems natural to begin our discussion of the living body by asking what living bodies are made of. Thanks to centuries of painstaking work by physicians, anatomists, and biologists, we now have a number of answers to this question at our disposal: living bodies are made of organs and tissues, which are composed of living cells; these cells are built out of proteins and other organic molecules, which are in turn made of elements like carbon, hydrogen, and oxygen. Rather than proceeding directly to these answers, however, I propose that we dwell for a while on the question itself, to see if it is truly as natural as it first appears. What does it mean to ask what something is made of? What sort of answer are we looking for, and what are we assuming about the thing in question?

Fig. 1. An antique roll-top desk.

Consider a simple example: an old roll-top desk that I've had since I was a child (Fig. 1). What is it made of? If we could pose this question to the artisan who made it, we
1.1. *What are living bodies made of?*

would probably learn that she had built the desk out of wood and various metal fasteners. In other words, she would tell us about the *materials* from which it was constructed. When we ask what an artificial thing is made of, this is the sort of answer we expect: we want to know what materials we would need if we wanted to build a desk ourselves. We would be taken aback if the artisan replied that the desk was made of carbon and iron, for we know that she did not craft the desk from atomic elements, but from wood, nails and screws.

When we learn what materials went into the desk's construction, our question is answered. It would be strange if, upon learning that the desk was made of oak, we continued pestering the craftsman to tell us what the *oak* was made of. She would surely reply that she hadn't *made* the wood—it was cut from an oak tree. Similarly, the steel fasteners were forged out of iron ore, which was not *made* but rather mined from within the Earth. My desk, like all manufactured things, is made of materials that were not themselves manufactured. Human manufacturing depends on "raw" materials like wood and ore, which are natural formations rather than artificial products. When we ask what an artificial thing is made of, we are ultimately asking after these raw materials, from which every manufacturing process begins.

What would it mean to ask what these raw materials are made of? When we asked what the desk was made of, we were asking about the materials that its maker used in constructing it. The desk was made from the wood of the oak tree. But it makes no sense to ask what materials the oak tree's maker used in constructing it, since we know that the oak tree had no maker. No one made the oak tree out of anything, for the oak tree was not *made* at all. A tree can be cultivated, but it cannot be constructed or manufactured. Like
1.2. *What is the difference between manufacturing and growth?*

all living bodies, the tree is not built, but *grown*. What is the difference between manufacturing and growth? To answer this question, we will have to study these two ways of coming to be in more detail.

2. *What is the difference between manufacturing and growth?*

The goal of manufacturing is to furnish us with things that do not occur naturally, by constructing them out of raw materials that can be found in nature. If desks grew on trees, we would not need to manufacture them; since they do not, we have to build them ourselves. In the previous section, we considered the raw materials that my desk is made of. Now, let us consider *how* it was made. The construction of my desk began with a design: a plan in the designer's head or on paper that specified what the final product would look like. Because the desk is made from many parts, its design had to indicate the shape and dimensions of each part, how these parts would fit together, and what materials they would be made of. To make this design a reality, the artisan had to select some large pieces of wood cut from the trunk of an oak tree, and cut and shape these into the parts of the desk whose various forms were laid out in the plan. Once these parts were shaped to the plan's specifications, they could be assembled and fastened together to form the final product.

We can distinguish three stages in the making of my desk, which are characteristic of manufacturing in general:

1) formulation of a plan or design;

2) shaping of parts from raw materials; and

3) assembly of these parts into the final product.
1.2. What is the difference between manufacturing and growth?

Like my desk, most human artifacts are complex wholes made of various different parts. Thus manufacturing is almost always a process of assembly, in which a complex whole is put together out of several parts in such a way that they form a useful whole. This is why we tend to use the word "synthetic" as a synonym for "artificial": human manufacturing is, by and large, a process of synthesis.

More often than not, the parts from which the final product is assembled were themselves synthesized out of other parts. In the manufacture of a cotton shirt, for example, the raw cotton is harvested, processed, and spun into thread; this thread is then woven into cloth; and this cloth is finally cut into pieces and sewn together to make a shirt. Thus there may be several stages of "refinement" between the raw materials from which manufacturing begins and the material from which the parts of the final product are shaped. Within this process of refinement, "material" and "product" become relative terms: the same cloth that is a finished product for the textile manufacturer is a material for the clothing manufacturer. Another way of putting this is to say that human artifacts tend to have a hierarchical organization: they are wholes made of parts that are themselves made of further parts, just as the shirt is made of pieces of cloth that are themselves made out of thread. At the bottom of this hierarchy, however, we always find the raw materials from which the manufacturing process began, which were not themselves manufactured.

A complex whole like a desk or a shirt cannot be assembled out of any old collection of parts. In order for the parts of a complex artifact to work together in just the
1.2. What is the difference between manufacturing and growth?

right way, they must be crafted with foresight and intelligence. Both the production of the parts and their assembly into a useful whole must be guided by a careful plan which describes the form of each part and how they will fit together. Thus manufacturing is a conceptual or intellectual process: it can only be carried out by an intelligent being whose idea of the finished product precedes the existence of the product itself, and orders the creation and assembly of its parts.

Manufacturing is the method human beings have devised for producing complex, organized wholes. Nature, however, accomplishes this in quite a different way. Manufacturing is an intrinsically conceptual or intellectual process: it can only be carried out by an intelligent being whose idea of the finished product precedes the existence of the product itself, and orders the creation and assembly of its parts. The living body, on the other hand, must come to be without the foresight and direction of an intelligent being. It must somehow organize itself into a complex whole, of its own accord, without a plan or an idea of what this whole will look like.

Consider the oak tree, from whose trunk my desk was made. No designer sets its dimensions in advance, or plots out the forms its leaves and branches will take. No one knows exactly what the tree that grows from an acorn will look like, or even if that acorn will germinate at all. The tree begins its life not as an idea in someone's mind but as part

---

24 “Take a look at painters for instance, if you would, or housebuilders or shipwrights or any of the other craftsmen you like, and see how each one places what he does into a certain organization, and compels one thing to be suited for another and to fit it until the entire object is put together in an organized and orderly way” (Gorgias 503e-504a).

25 Of course, human beings can and do intervene in the growth of trees and other living things—cultivating those that are useful to us and inhibiting the growth of others. However, the work of cultivation is quite different from that of manufacturing; and most living beings are perfectly capable of reproducing without human intervention.
1.2. *What is the difference between manufacturing and growth?*

of another tree's living body—a part that breaks off to begin its own separate existence as a new organism. This new tree will resemble its parents in many ways, but it will also be shaped by the conditions it encounters over the course of its own life. It will be a new individual with its own unique history, and its body will bear the traces of that history.26

No builder crafts the parts of the oak tree. Its cells come to be by cell division, and they are all descended from a single cell. As the cells of the developing tree embryo divide and multiply, they differentiate themselves into various different cell-types. These specialized cells continue to reproduce and increase in number, organizing themselves into tissues and organs to form the familiar structures of a living oak: a hard, woody trunk that branches repeatedly into roots below the soil and branches and twigs above, from which sprout photosynthesizing leaves. The mature tree is a complex whole, within which we can distinguish many different parts. But the tree was not *assembled* out of these parts—it did not begin its life as a disconnected pile of leaves, branches and bark. Instead, it generated its own organs "on the fly" over the course of its development. As the tree grew from a shoot to sapling to mature tree, its trunk slowly divided into branches, its branches divided into twigs, and these twigs sprouted leaves.

The desk is built one piece at a time, gradually approaching completion, until finally every piece is in place. There are no baby desks—only incomplete ones. The tree, on the other hand, is never *unfinished*, but only *im mature*: at every moment of its existence, it is a living, functioning whole from which nothing is missing. The tree is

---

26 It has become commonplace today to say that the organism's genome constitutes a design or blueprint for its mature form. I will address this claim at length in section 1.4. For now, it will suffice to point out that even the proponents of this view admit that the living body is shaped by its environment as well as its genes. Thus the so-called "genetic blueprint," unlike the design of an artifact, would not *fully* determine in advance the form of the adult organism. This difference between the living body and the artifact is far from
1.2. What is the difference between manufacturing and growth?

never incomplete—but it is never complete, either. Unlike manufacturing, which comes to an end when the product matches its design, growth has no end-point. The processes that produce the adult tree also keep it alive once it has reached maturity, and they will stop only when the tree is dead. Manufacturing always precedes its own product: the manufacturing process comes to an end as soon as its product comes into existence. Growth, on the other hand, is contemporaneous with its product: it comes to an end only when its product ceases to exist.\(^{27}\)

No intellect commands the tree to grow, or supervises its development. The oak rises unbidden from the soil, its towering form emerging slowly and improbably from the tiny acorn. Somehow the little seed is able to produce the mature tree—or more accurately, the seed somehow has the power to become the tree, to transform itself into the fully-grown oak. Of course, this is not an unconditional power. The acorn can only become an oak if it has the rain, sun and air that it needs. Nevertheless, the sun and the rain do not build the tree; the acorn is not a passive material on which the environment acts. On the contrary, the tree is the agent of its own development.

This is the wondrous thing about the living body, a wonder we are so familiar with that we have almost forgotten it: the living body generates itself—it is both cause and effect of itself.\(^ {28}\) If this seems paradoxical to us, it is because we have come to think of all creation on the model of manufacturing. It is impossible for a being to synthesize itself, for synthesis always proceeds from part to whole: first the parts are formed, and

\(^{27}\) This is not to say that the living body increases in size until its death, but rather that its cells are constantly dying and being replaced by new cells; life is not a static property or fixed state, but a process.

\(^{28}\) See Kant's *Critique of Judgment* (§64).
then they are assembled into a finished product. This product cannot be responsible for the genesis of its own parts, or their assembly, since it does not exist until these tasks are complete. Thus manufacturing requires the intervention of some external agent, an intelligent being who can envision the final product and organize raw materials to realize this idea.

If the living body generates itself, it cannot be by manufacturing. But what is the alternative? Self-genesis still appears to be a paradox: in order to generate itself, to be its own cause, the living body would have to somehow precede itself—to exist before it exists. This paradox is resolved when we realize that the organism doing the generating both is and is not the organism that gets generated. The acorn, the sapling, and the mature tree are all the same organism—and yet they are not identical to one another. A being can be both cause and effect of itself only by changing over time. Again, if this is surprising to us, it is because we are accustomed to the perspective of manufacturing. The manufactured thing is a finished product, designed not to change once it is complete. Having no principle of change within itself, the artifact will only change if external forces act on it. It persists by a sort of inertia: all it must do to continue existing is not change. If external forces alter it too much from its original design, we say that it has ceased to exist, ceased to be the thing it was made to be. The living body, on the other hand, stays alive only by changing itself. If it does not change, it will die—it will cease to be itself. Thus the living body can remain itself only by changing itself. Its existence is an active project.

---

29 See Aristotle's *Physics* 192b.
1.2. What is the difference between manufacturing and growth?

What is the nature of this project? How must the living body change itself in order to stay alive? The organism keeps itself alive through the activity of metabolism, in which it takes in nutrients and energy from its environment, and expels waste products from its body (Jonas 2001, pp. 75-84). It is tempting to think of metabolism by analogy to the engines we build, as if food were fuel that the living body burns. But in fact, the living body's relation to its food is quite different from that of an engine to its fuel. A car needs fuel in order to run, but not in order to go on existing. The fuel passes through the car without becoming part of the car; aside from wear and tear, the parts of the car remain unchanged. The living body, on the other hand, needs food in order to survive. The nutrients it consumes do not simply pass through it, leaving its parts unchanged. On the contrary, these nutrients are incorporated into the living body, becoming part of it. It is by incorporating nutrients from its environment that the organism is able to grow, to repair parts of itself that have been damaged, and to replace the parts that it is constantly shedding.

The parts I find in my desk today are the same parts I would have found in it when I was a child, and the same parts that the desk was first constructed out of whenever it was made. By contrast, the cells you will find in my body today are not the same cells you would have found there when I was a child. First of all, there are a great deal more of them; second, most (if not all) of the cells in my childhood body have died and been replaced by new cells. Which is the real me—my body today, or the body I had as a child? They are both me, and yet I am not simply identical to either one of them. The living body is somehow more than the parts that can be found in it at any given instant. It is a being that exists in time, that takes time to be what it is. The organism can never be
1.2. *What is the difference between manufacturing and growth?*

located in a single instant, or a single configuration of atoms and molecules; to freeze it or try to hold it still in this way is to miss it, for it exists only in movement—it is not a *thing*, but an *activity*.

   Every part of the living body participates in this metabolic activity, and every part depends on this activity for its own continued existence. The leaves of the oak tree cannot live without the water absorbed by its roots, and the roots cannot live without the photosynthesis that takes place in the leaves. Every part of the living body depends on every other part. Thus a change to any single part will affect all the others. In this, the living body stands in stark contrast to the manufactured thing. Each part of my desk was crafted separately from the others, and each one is self-sufficient; having come to be independently, they are perfectly capable of existing apart. The builder may assemble these parts into a functioning whole, but the parts themselves are entirely indifferent to this whole, and to one another. The desk is a purely *extrinsic* whole, imposed upon indifferent parts, just as the shapes of those parts were imposed on indifferent materials. Thus manufacturing produces wholes in which a change to one part does *not* affect the others. It makes no difference to one drawer whether the other drawers are present or not: if one part of the desk is damaged or removed, the others will not compensate for this loss. The parts of the desk did not produce one another, and so they will not generate new parts to replace those that are lost. Nor will one part of the desk split off and grow into a new desk.\(^{30}\) Reproduction and self-repair, so commonplace in nature, are utterly beyond the capacities of even our most sophisticated machines. But these amazing powers are simply extensions of the power of *growth*. It is no more miraculous that a living body
should be able to re-generate a lost limb than that the body was able to generate this limb in the first place.

In looking at the different ways in which living bodies and human artifacts come to be, we have discovered that they also differ in their way of being. Indeed, we have discovered that growth is the living body's way of being—that how it comes to be and how it is are one and the same. The living body is a developmental being, a movement of becoming which is never fixed. If this is hard for us to conceive, it is at least in part because we have become accustomed to thinking of the living body, and nature in general, on the model of manufacturing and manufactured things. In the following sections, I will discuss the origins of this synthetic view of nature, and the conceptual difficulties it has engendered.

3. Atomism and the problem of form

The Ancient Greeks distinguished, as a matter of course, between "what is grown and what is put together" (Republic 533b): that is, between the products of nature and the products of human techne. Aristotle explains this distinction by saying that beings which come to be by nature are self-moving: natural beings have within themselves the source of their own growth and change, unlike things which come to be by techne, whose source lies outside of them in their human maker (Physics 192b). He goes on to say that it would be absurd to try to prove that nature exists, "for it is obvious that there are many things of this kind, and to prove what is obvious by what is not is the mark of a man who

---

30 Aristotle famously observes that if one plants a wooden bed in the ground, it will never sprout another bed; if anything were to grow from it, it would be a tree (Physics 193b).
1.3. Atomism and the problem of form

is unable to distinguish what is self-evident from what is not" (193a). And yet, what seemed self-evident to Aristotle no longer seems so to us today. The ancient distinction between growth and manufacturing has collapsed, and with it the distinction between nature and artifice. Of course we know that the oak desk was built by human hands, while the oak tree was not. But this now appears to us a merely superficial difference, which conceals a more fundamental unity: that desk and tree are both made of the same "basic building blocks of matter"—protons, neutrons, and electrons. If the tree is made of cells, molecules, and atoms in the same way that the desk is made of wood, then everything is synthetic: growth and manufacturing are just two different ways of putting things together.

The shift from Aristotle's view of nature to the one we hold today began in the seventeenth century (Arendt 1998; Lachterman 1989; Funkenstein 1986). It was part of the larger rejection of Ancient (and especially Aristotelian) ideas spearheaded by thinkers such as Galileo, Descartes, and Hobbes, which inaugurated the Modern era. The collapse of the nature/artifice distinction played a particularly important role in the emergence of Modern science, for it inaugurated a radical new approach to the study of nature. This new approach can be summed up in one simple but radical proposal, which I will call the "mechanical hypothesis." We saw above (1.1) that manufacturing always begins with natural formations that were not themselves manufactured. The mechanical hypothesis proposes that we treat these natural formations as if they too were artifacts. In other words, the mechanical hypothesis is that nature is a machine—a synthetic whole, like the products of human engineering. It follows that natural phenomena can be reverse-

---

31 The Ancient Greek word for nature, physis, comes from the verb to grow (phuo). The word for "what is
engineered, by posing to them the same questions we pose to manufactured things when we want to understand them: "What is this made of, and how was it made?" If nature could be reverse-engineered, then we could understand it as completely as we understand the things we make ourselves.  

Biologist Richard Dawkins (1996) articulates this method of investigating nature very clearly: "If we wish to understand how a machine or living body works, we look to its component parts and ask how they interact with each other. If there is a complex thing that we do not yet understand, we can come to understand it in terms of simpler parts that we do already understand" (11). This same principle was articulated much earlier by Descartes. The key assumption of this method is that the natural phenomena we want to explain are modular: that they can be divided into "simple parts," each of which can be understood in isolation from the others. Once we understand each of the parts individually, we can ask how they interact with one another; but we assume in advance that these interactions will not change anything essential in the parts themselves. In other words, we assume that the "simple parts" are indifferent to their interactions with one another—just like the parts of a machine. The parts determine the interactions, but the interactions do not in any way determine the parts. Causality and explanation are "bottom-up," moving from simple part to complex whole, and never in the other direction. Descartes captured this ontology of nature in his principle that all extended things—including space or "extension" itself—exist partes extra partes (parts outside of

---

32 See Kant's *Critique of Judgment* (§68): "[W]hen we study nature in terms of mechanism, we keep to what we can observe or experiment on in such a way that we could produce it as nature does, at least in terms of similar laws; for we have complete insight only into what we can ourselves make and accomplish according to concepts."
1.3. Atomism and the problem of form

parts). By this he meant not only that their parts are spatially outside of one another (i.e. spread out in space), but also that they are external to one another in their very being—that each part is what it is independent of all the others.

In order to explain the natural world in this way, we must determine in each case how to divide the natural phenomenon under investigation into its "component parts", and where to stop dividing. In the artificial thing, the answers to these questions are clear: the artifact was built out of parts, so these are the parts we want to divide it into; and we stop asking what these parts are made of when we reach raw materials that were not themselves manufactured. But a rock or a tree could be divided in countless different ways. If we split the rock in half, do we discover what the rock is made of? Or have we simply turned one rock into two? Of course the divisions we choose to make are often suggested by the natural being itself, in conjunction with our own interests. If we are interested in the functions of the living body, we will divide it into organs and tissues; if we are interested in how these organs develop, we will have to divide them into individual cells. But each of these divisions offers only a provisional answer to the question of what the living body is made of. It may satisfy a particular, limited interest; but the parts arrived at cannot offer a final explanation for the organism, since they are themselves still open to the further question: what are these parts made of? Whatever

---

33 Discourse on the Method, Part 2 (CSMK Vol. 1, p. 120; AT VI, p. 18).
34 "Modularity is the human mind's lever against complexity" (Victor 2012).
35 See Descartes' letter to Henry More (5 February 1649): "I call extended only what is imaginable as having parts outside of parts" ["ita illud solum quod e est imaginabile, ut habens partes extra partes, . . . dico esse extensum"] (CSMK Vol. 3, p. 362; AT V, p. 270). (I have modified the English translation, which for some reason renders "partes extra partes" as "parts within parts." ) Merleau-Ponty cites this passage in his lectures on The Incarnate Subject (IS 50/33).
36 "There is a hierarchy of subcomponents within components. We explain the behaviour of a component at any given level, in terms of interactions between sub- components whose own internal organization, for the moment, is taken for granted" (Dawkins 1996, p. 12).
1.3. Atomism and the problem of form

parts we divide the natural being into can usually be divided still further. Where should this process of division end? At what point can we claim to have arrived at the ultimate constituents of the phenomenon, which would make it totally intelligible to us?

Because it is committed to explaining complex wholes in terms of their simpler parts, mechanistic explanation is forced to keep on dividing the natural phenomenon until it reaches parts that cannot be divided further, and which can thus serve as ultimate terms of explanation. In other words, the dynamic of mechanistic explanation commits it to a doctrine of atomism, which says a) that there is a smallest unit or set of units into which everything that exists can be divided; b) that all beings are made out of these units, in the same way that an artifact is made out of raw materials; and c) that these indivisible units are the ultimate units of intelligibility, the ultimate terms in which all that exists must be explained. Notice that this is a methodological assumption of mechanistic science, not an empirical claim: the doctrine of atomism did not originate with the discovery of protons and electrons; nor did these discoveries constitute a proof of the doctrine. Conversely, one may dispute the claims of atomism without disputing the existence of quarks, or whatever particles physicists currently believe to be the most fundamental. Atomism is much more than the claim that beings in nature can be divided into tiny, indivisible particles; it is a claim about the relations of these particles to one another, and to the wholes in which they are found.

Having discovered what nature is made of by dividing it into its simplest components, mechanistic science is obliged to explain how these components are put together—how they interact to form the complex wholes that we perceive in nature.
1.3. Atomism and the problem of form

Since these atoms have been posited as completely indifferent to one another, they are not going to assemble themselves into complex wholes on their own. Like the parts of a machine, they require some external cause to bring this about. In manufacturing, this work is accomplished by the human craftsperson. In nature, according to the mechanical hypothesis, the work of shaping matter is accomplished by forces of nature, such as gravity or electromagnetism.\footnote{37} Atoms, by definition, cannot be changed, but they can be moved around with respect to one another, and this is what the forces of nature do. These forces operate according to the laws of nature, so called because they hold universally, without exception—they are the same everywhere and at all times.

Atomic matter and the forces of nature together explain what nature is made of and how it is made. The natural world we see today was produced by the constant action of the forces of nature on matter, in accordance with natural laws. Thus Kant wrote in 1755, "Give me the material, and I will build a world out of it! That is, give me the material, and I will show you how a world is to come into being out of it."\footnote{38} The positions of the atoms that make up the cosmos and the laws that govern the movements of those atoms together constitute a full account of the natural world. Not only do the laws of nature explain how the present universe arose, through the actions of force on matter, but they also predict how the universe will be configured in the future. This claim was famously articulated by the mathematician and scientist Pierre-Simon Laplace (1814/1951):

\footnote{37 The European founders of the mechanistic research program did not initially reject the Christian view that the world was created by God. Rather, they proposed that God had made the world in roughly the same way that humans made artifacts, and that the natural world could therefore be understood by asking how God had made the world and what He had made it out of. It was only later that mechanistic science came to}
1.3. Atomism and the problem of form

Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes. (p. 4, translation modified)

The mechanical hypothesis thus leads to a deterministic view of nature as a clockwork universe, in which everything is determined in advance.

Time, on this view, turns out to be a kind of illusion, produced by the limits of our own understanding. Our finite perspective on the world gives us the impression that events occur which are unprecedented, that nature produces phenomena which are truly new and could not have been predicted. But a greater intelligence, with a fuller understanding of nature than our own, would see that the whole history of the cosmos was already given at its inception. What appears to us as a new form of being is really just a different configuration of the same, unchanging matter. Form, like time, is in the eye of the beholder: it makes no difference to matter, or to the forces that move and shape that matter. The complex can always be explained in terms of the simple; thus complexity itself is a kind of illusion, an epiphenomenon that can be explained away, and which itself explains nothing.

The appeal of the mechanistic research program's lies in its promise to explain the ever-changing diversity of nature in terms of an underlying unity that does not change. It claims to bring predictability and certainty to an uncertain world. And with the promise

---

reject the hypothesis of a divine craftsman. However, this left science in the incoherent position of claiming that nature is an artifact without an artificer, a machine with no manufacturer.

1.3. Atomism and the problem of form

of prediction comes the possibility of control over nature.\textsuperscript{39} We have seen that human manufacturing is limited by its dependence on nature for the "raw materials" which it cannot produce itself, and from which it must always begin. If nature is itself an artifact, however, then there is in principle nothing that human engineering might not build. Once science has shown how a natural being is put together out of the basic building blocks of matter, engineering should be able to produce that thing out of these building blocks, and even modify it to better suit human needs. This dream is evident in our visions of future technological utopias, such as television's Star Trek, in which food is not grown but simply synthesized or "replicated" out of atomic elements. It applies not only to our natural environment, but also to our own bodies: since we are ourselves machines, there is no reason why we ought not to be able to repair ourselves the way we repair our cars, conquering illness, aging, and even death itself. Thus the mechanistic research program promises not only understanding of nature, but ultimately freedom from nature.

Can these promises be fulfilled? Is freedom from nature possible, or desirable? The mechanical hypothesis helped to usher in the age of Modern science, and it has produced astonishing results in a number of fields. Indeed, this approach has been so successful that it has become identified with science itself. What was once a hypothesis has become an ontology: we have gone from treating nature as if it were a machine, to taking for granted that nature actually is a machine. But the success of the mechanical

\textsuperscript{39} This has always been an explicit goal of the mechanistic research program. See e.g. Descartes' \textit{Discourse on the Method} (Part 6): "For my notions had made me see that it is possible to reach understandings which are extremely useful for life, and that instead of the speculative philosophy which is taught in the schools, we can find a \textit{practical} philosophy by which, through understanding the force and actions of fire, water, air, stars, heavens, and all the other bodies which surround us \textit{as distinctly as we understand the various crafts of our artisans}, we could use them in the same way for all applications for which they are appropriate and thus make ourselves, as it were, \textit{the masters and possessors of nature}" (CSMK Vol. 1, pp. 142-3; AT VI, p. 62).
1.4. The paradox of development

hypothesis does not license this conclusion. Many highly successful research programs have been founded on hypotheses that were later overturned—think of geocentric astronomy, or Newtonian physics. The mechanical hypothesis has been very useful, but that does not make it true. Furthermore, the usefulness of this hypothesis is running out; in many fields, it is already more of a hindrance than a help. Nowhere is this more clear than in our attempts to understand the development of the living body. In the next section, we will look at how biologists have tried to make sense of development mechanistically, and the difficulties that these attempts have encountered. I will argue that development cannot, in principle, be explained in synthetic or mechanical terms, and thus that the phenomenon of development demands a new ontology whose concepts are not drawn from machines or manufacturing.

4. The paradox of development

I argued above that the process of development sets the living body apart from the machine or artifact, which is a product of manufacturing. In biology, however, it has become commonplace to treat the living body as if it were a machine or synthetic whole. In this section, I will argue that this view of the body has arisen not in light of the phenomenon of development, but rather in spite of it. By reviewing the history of biology's attempts to understand development, I will show that the facts of development

---

40 Development is "the basic biological process. Development is what distinguishes biological systems from other sorts of systems" (Robert 2004, p. 34).
41 As we will see, however, there are some important exceptions to this consensus.
1.4. The paradox of development

cannot be explained in mechanical terms, and that it is only by systematically neglecting these facts that biologists have been able to maintain that the living body is a machine.\(^\text{42}\)

\textit{A. Epigenesis vs. preformation}

From the mechanistic point of view, the problem of development is a problem of form. The living body must ultimately be made up of the same matter as everything else; thus its special properties can only be explained by the complex way in which this matter has been arranged in the organism. The question, then, concerns the origin of this arrangement: how is matter shaped into the complex form of the living body?

The origin of form presents no puzzle in the case of manufactured things: their form is given to them by their manufacturers, according to some design. Manufactured things are products of intellect, and their form originates in an idea. For a long time, this seemed to be the only plausible explanation for living forms: namely, that they too were created by some (divine) intellect. The eighteenth-century theologian William Paley (1802) put this point very clearly: If we happen to run across a watch lying on the ground, and observe the way that its many parts fit together in a complex and delicate way to serve a single function, we are bound to assume that it was made by some intelligent being to perform this very function. Similarly, when we see how complex and how organized the living body is, how well all of its various parts fit together, we cannot help thinking that some intelligence must be responsible for its design. However, this theory of special creation only answers half the problem of organic form: it explains the origin

\(^{42}\text{Though we will see that some biologists are actively trying to bring the study of development back into mainstream biology. For a more in-depth but still accessible introduction to this history, see Fox Keller 2000. For a more detailed version of the argument I offer in this section, see Robert 2004.\)
1.4. The paradox of development

of species forms, but not how these forms are recreated in each generation. How is the distinctive form of each species passed on from parent to child? What guides the development of the individual body into the mature form characteristic of its species? This is the problem of embryogenesis: to explain how the matter of the developing embryo is shaped into the mature form of the adult body.

In response to this problem, late-seventeenth- and eighteenth-century naturalists argued that when God created the first living bodies, he placed within them the seeds of every future living being. Thus the seeds of each generation were encased within the seeds of their progenitors, and each of these seeds held within it the seeds of all of its descendants. Moreover, each of these seeds was itself a sort of miniature organism, imbued already with the form of the adult body it would one day become. Development, then, did not have to generate a new form, but merely to "unbox" the form already present in the seed since Creation. This may seem preposterous to us today. But supporters of "preformationism" were often driven to this view by the seemingly insurmountable difficulties facing the alternative, "epigenetic" view, which held that mechanical forces shaped a new organism out of unformed matter over the course of embryogenesis. It seemed impossible to explain how blind forces acting on inert matter could produce the incredibly complex form of the living body. According to the basic hypotheses of mechanistic science, matter is passive and indifferent to the forms it is given; it can be shaped, but it will not shape itself. Thus the process of development

---

43 This theory of "emboîtement" was first articulated fully by Malebranche. See Roe 2003, p. 5.
44 See Roe 2003, p. 9: "Matter was viewed by most as entirely passive, put into motion only through mechanical laws. But since these laws of motion are blind, that they could know how to form a living organism seemed out of the question. Both self-active matter and a God actively involved in each instance of generation were ruled out in the mechanistic universe of late-seventeenth-century thinkers. The theory of
could not be explained by the embryonic matter, but only by the forces acting on this matter. According to mechanistic assumptions, however, these forces are "blind", i.e. indifferent to the forms they create. Forces act on matter at the atomic level; the macroscopic forms they produce are mere side-effects. But the living body is so complex and well-organized, its myriad parts all fitting together so perfectly, that it seemed absurd to claim that its form was merely an accidental side-effect. The only alternative was to attribute the living body's complex organization to some intelligent designer. Form, according to mechanistic ontology, must be either an epiphenomenon or an idea.

B. Darwinism, genetics, and the Modern Synthesis in biology

No compelling alternative to the theory of intelligent design was proposed until 1859, when Charles Darwin published *The Origin of Species*. Darwin famously argued that the great variety of living forms, and their exquisite adaptedness to their natural environments, could be explained by "natural selection", a process analogous to the artificial selection carried out by human breeders.

The practice of breeding plants and animals is older than recorded history. Breeders rely upon two key properties of sexual reproduction. One is that offspring tend to resemble their parents. Thus, breeding together organisms with desirable qualities is likely to produce more organisms with these same qualities. The second is that this resemblance is neither perfect nor uniform: the various offspring of a given mating pair are rarely identical to one another. Thus, the breeder can select from the litter or the field those offspring with the most desirable traits, and breed them with other desirable
1.4. The paradox of development

specimens, to yield organisms that are better adapted to the breeder's aims. Darwin's insight was that a similar process could be at work in nature. Instead of human breeders selecting the organisms that best fit their needs, and breeding them together to produce offspring, competition between organisms for limited resources and mates could "select" the organisms best adapted to their environments and allow them to have the most offspring. In the same way that human selection over a few thousand years has produced a variety of new breeds of pigeon (Darwin 1985 [1859], pp. 81ff), so natural selection operating over millions of years in a variety of environments could have produced, from some common ancestor, all the various species alive today.

Darwin's theory of evolution by natural selection was taken as a major victory for the mechanistic research program. Biologists interpreted natural selection as a force, a "selection pressure" which slowly shaped the bodies of living beings into the complex forms we see today. However, the problem of embryogenesis remained. Darwin's theory presupposes that organisms vary from one another, and can pass on some of these variations to their offspring; it does not attempt to explain how they reproduce themselves, how a single fertilized egg-cell is able to develop into a complex organism that resembles its parents.

Although the theory of evolution did not resolve the outstanding problems of embryology, it did spur renewed interest in the field. Evolution and embryogenesis—the origin of species and the origin of individual organisms—were taken to be deeply connected, and it was expected that the study of each would shed light on the other.
1.4. The paradox of development

Similarities in embryogenesis across different animal species were one of the most important pieces of evidence for Darwin's claim that all life is descended from a common ancestor (Gilbert 2003, p. 468). As Darwin wrote in *The Origin of Species* (1985 [1859]), "Community of embryonic structure reveals community of descent" (p. 427). However, this evolutionary interest in the study of development did not last. In twentieth century evolutionary biology, embryology was supplanted by a new science: genetics, the study of inheritance.

The scientific study of heritability did not begin in earnest until 1900, when the pioneering work of Gregor Mendel was re-discovered. Before this, it was generally assumed that offspring inherited a blend of their parents' traits. But in his work with pea-plants, Mendel discovered traits that were inherited in a discrete or particulate fashion. For example, when pea plants with green seeds are crossed with plants that have yellow seeds, the seeds of the next generation are not a blend of green and yellow—their seeds are simply yellow, as if they had not inherited the green color at all. However, if these yellow seeds are planted, they do not produce another generation with all yellow seeds, as one might have expected. Instead, approximately one out of four plants in the third generation will have green seeds. In other words, while it seemed in the first generation that the green color had not been passed on, it had in fact skipped a generation, reappearing in the offspring of the yellow-seeded plants. In order to explain these results, Mendel postulated that each heritable trait is determined by two particulate "factors".

---

45 Darwin (1985 [1859]) also notes that the effects of human selection can be observed over much shorter time scales: "It is certain that several of our eminent breeders have, even within a single lifetime, modified to a large extent some breeds of cattle and sheep" (p. 90).

46 Mendel's work was published in 1866, but remained obscure for over thirty years until his results were independently reproduced by others. See Mendel et al 1950.
which the plant inherits from its parents. A given plant could have two green factors, two yellow factors, or one of each. In this last case, the yellow factor was "dominant" over the "recessive" green one, producing a plant with yellow seeds, as in the second generation above. In sexual reproduction, each plant passes on just one of its two factors to its offspring. Thus when plants with one yellow factor and one green factor are crossed together, there is a one in four chance that their offspring will receive two green factors, and so display the recessive trait of having green seeds.

From Mendel's humble work with pea plants, the science of genetics was born. Mendel's heritable "factors" become known as "genes," and the alternate forms (dominant and recessive) of a given gene were named "alleles." At first it was not clear to biologists how this new understanding of heredity fit with Darwin's theory of evolution, or even if the two were compatible at all. However, this changed with the rise of population genetics in the 1920s and 30s (Okasha 2012). This new field used Mendel's laws of heredity to generate mathematical models of how allele frequency would change over time in a given breeding population. For the first time, questions about the kinds of changes natural selection could produce, and how long this would take, could be answered mathematically. Thus Darwin's theory of evolution by natural selection was combined with Mendelian genetics to yield a new understanding of evolution as a change in gene frequencies over time. This fusion became known as the Modern Synthesis in biology.

The Modern Synthesis discouraged inquiry into developmental questions. Geneticists were focused on discrete traits that manifested themselves in the adult organism, and the discrete genes that were postulated as transmitting these traits from
1.4. The paradox of development

devlopment was treated as a "black box": the question of how exactly genotype led to phenotype could be answered later. The big picture, at least, seemed clear enough: one gene specifies a single trait; add up enough genes, and one has a *blueprint* for the whole organism. Tiny changes in this blueprint, brought about by random mutation, produce tiny changes in adult organisms which can then be acted upon by natural selection. Genes that produce fitter organisms multiply; genes that produce less well-adapted organisms disappear. Biologist Richard Dawkins has captured this understanding of evolution—which has become known as "Neo-Darwinism"—in his image of the "blind watchmaker," which is based on Paley's argument for intelligent design (see above). According to Dawkins (1996, p. 3), the living body is a complex machine, like Paley's watch. However, the "watch-maker" who designed the living body was not an intelligent being, but the blind forces of natural selection. Evolution is a blind watch-maker because it generates designs for working machines—but it accomplishes this without intelligence, purpose or foresight.

The Neo-Darwinist account of evolution is thus a kind of hybrid theory, which mixes an epigenetic account of species formation with a preformationist account of embryogenesis. Species forms are generated by blind, mechanical forces, but the development of individual organisms is simply the "unfolding" or "un-boxing" of a form or design that resides already in the zygote's genes. As an account of embryogenesis, this is far from satisfactory, as we will soon see in more detail. However, it seemed plausible in the context of the Modern Synthesis, where genetics reigned supreme. From its

47 "[D]evelopment is now standardly construed as the epigenesis of something preformed in DNA… However, if, as so many have argued, neither epigenesis nor preformationism is correct, then, to my mind, a monstrous hybrid should be no better" (Robert 2004, p. 35).
1.4. The paradox of development

inception, genetics aimed at isolating discrete, heritable traits in adult organisms. The focus on discrete traits led to a view of the whole organism as a synthetic whole—a mere collection of independent parts; and the focus on heritable traits led to the tacit assumption that everything about the organism is determined in advance by its genes, just as a machine is determined in advance by its design. This preformationist view of development could persist only as long as genetics remained ascendant in the study of evolution, and embryology remained on the sidelines. Ironically, it was discoveries in genetics—starting with the discovery of the double-helix—that eventually overturned this state of affairs, and brought developmental considerations back into the study of evolution.

C. DNA and the return of development

If the Modern Synthesis constituted a first revolution in genetics, then a second revolution was brought about by the discovery in 1953 of the structure of DNA. At first, this discovery was taken as confirming and extending the Neo-Darwinist theory of evolution set out in the Modern Synthesis—and indeed, this remains a common view today. However, this second revolution in fact led to a series of discoveries that ultimately undermined the foundations of Neo-Darwinism—and with them, the view of the body as a synthetic whole constructed according to a genetic blueprint.

The discovery of the double-helix structure of DNA promised answers to long-standing questions about the mechanism of inheritance. Up until this discovery, the gene had been a strictly hypothetical entity: a particle of inheritance postulated to explain the observed facts of heredity. The discovery of the double-helix finally allowed biologists to
identify the gene with a concrete anatomical structure. The irregular, linear sequence of bases in a strand of DNA suggested that, as Watson and Crick put it in 1953, "the precise sequence of the bases is the code which carries the genetical information." The idea of the gene as information encoded in DNA strongly supported the received view of the genome as a design for the adult organism. It implied that if we could "crack" this code, we would be able to "read off" an organism's design from its genes—and perhaps even learn to design new organisms ourselves.48

It had already been suggested, in the 1940s, that genes might specify heritable traits by producing proteins that control biochemical reactions—the one gene–one enzyme hypothesis. With the discovery of DNA's linear structure, this hypothesis appeared in a new light, suggesting that the sequence of nucleotides in a gene might correspond to the sequence of amino acids in a protein (Fox Keller 2000, pp. 51-54). In 1957, Francis Crick formulated this "sequence hypothesis" as follows: "the specificity of a piece of nucleic acid is expressed solely by the sequence of its bases, [and] this sequence is a (simple) code for the amino acid sequence of a particular protein." It took several years to decipher this code, but by 1966 biologists could say with confidence that each of the amino acids found in proteins is coded by a different triplet of nucleotides in DNA. It would be difficult to overstate the excitement that attended this discovery: it seemed that we had found the book of life and learned to read its language.

This simple and appealing picture continues to exert a powerful hold on our imaginations. However, subsequent discoveries in molecular biology have complicated

---

48 This is the explicit aim of the relatively new sub-field of biology called "Synthetic Biology." As I noted above (1.3), the prospect of reverse-engineering natural phenomena always carries with it the dream of re-designing these phenomena to better suit human needs.
and eventually undermined this simple picture. The first of these discoveries came just two years after Crick proposed his "sequence hypothesis." In 1959, F. Jacob and J. Monod showed that structural genes which "code" for proteins are not the only kind of gene; there exist also regulator genes that determine the rate at which structural genes are transcribed. Indeed, recent estimates suggest that only about three percent of the human genome codes for proteins—and this percentage may be even lower in other organisms (Fox Keller 2000, pp. 55-59). Further discoveries followed: we now know that messenger RNA does not passively "transcribe" DNA sequences into proteins, but actively edits these sequences by splicing together transcripts from different locations, and even replacing one base with another in the transcribed sequence (Fox Keller 2000, pp. 59-64). These discoveries severely undermined the one gene–one enzyme hypothesis: the vast majority of the genome does not "code" for protein at all, and those DNA sequences that are "transcribed" into proteins can generate a number—sometimes hundreds—of different proteins, depending on how they are edited in mRNA.49

These discoveries also refuted Crick's "Central Dogma" of molecular biology, which held that "information" could only flow from DNA and RNA into protein, and not the other way. This hypothesis was crucial to the understanding of DNA as a fixed "blueprint" for the adult organism. We now know that the transcription of DNA into RNA, and RNA into protein, is subject to the "complex regulatory dynamics of the cell as a whole" (Fox Keller 2000, p. 63). Rather than a one-way, linear causality in which DNA

49 Of course, the "one gene–one enzyme" hypothesis can be preserved by changing the meaning of "gene", as molecular biologists sometimes do, so that it refers not to sequences of chromosomal DNA but to mature mRNA transcripts that have already been spliced and edited. But this move comes at a high price: mRNA transcripts are short-lived entities, put together "on the fly"; unlike chromosomal DNA, they are neither
1.4. The paradox of development

would control the synthesis of proteins, we find in the cell a system of circular causality in which DNA's regulation of protein synthesis is itself regulated by the proteins synthesized—as well as by a host of other factors, such as proteins and hormones synthesized by other cells.\footnote{Moss 1992, p. 344.}

From a developmental perspective, the discovery of gene regulation was not at all surprising. Indeed, the biologist T. H. Morgan had anticipated Jacob and Monod's discovery by almost three decades, in a book titled "Embryology and Genetics": "The implication in most genetic interpretation is that all the genes are acting all the time in the same way. This would leave unexplained why some cells of the embryo develop in one way, some in another, if the genes are the only agents in the results. An alternate view would be to assume that different batteries of genes come into action as development proceeds."\footnote{Cited in Fox Keller 2000, p. 56.} The central problem of embryology is to explain how a single cell can give rise to the complex, multicellular body characteristic of its species.\footnote{Robert 2000, p. 1.}

As the cells of a developing embryo divide and multiply, they must differentiate themselves into hundreds of specialized cell types; and as these differentiated cells continue to reproduce, they must organize themselves in just the right way to form the many different organs and tissues of the growing body. All of these cells are genetically identical, being descended from the same fertilized ovum or zygote; and yet, something causes them to develop into stable nor heritable. On this definition of "gene", then, the gene is no longer the unit of inheritance (Fox Keller 2000, pp. 63-64).

\footnote{"Explorations of the mechanisms involved at the level of the DNA molecule itself, have not led to any privileged point of causal origins, but rather immediately refer back to the complex state of the cell/organism as a whole as the causal basis of the activity of the genes" (Moss 1992, p. 344).}

\footnote{The central problem of developmental biology is to understand how a relatively simple and homogeneous cellular mass can differentiate into a relatively complex and heterogeneous organism closely resembling its progenitor(s) in relevant respects" (Robert 2000, p. 1).}
1.4. The paradox of development

blood cells or bone cells, neurons or myocytes. Thus the DNA sequence cannot, on its own, explain the process of differential development that produces the adult body.

From the standpoint of development, the notion of the genome as a design or "blueprint" for the adult organism is absurd. A blueprint tells an intelligent builder what her final product should look like. From her standpoint outside of the thing she is building, she can compare what she has accomplished at each stage of construction to the goal set out in her plan. The cells in a growing body have no such external view of the body as a whole; each is responding only to its immediate surroundings. No cell is in charge of development; there is no central control center telling each cell what to do. And yet somehow, miraculously, these cells are able to organize themselves into a functioning body of incredible complexity. If DNA were a blueprint, then development would be like giving a trillion copies of the same blueprint to a trillion people, and asking them to build a city out of their own bodies.

Over the last thirty-five years, a movement has arisen within biology to bring developmental questions back into the study of evolution. We might trace the beginning of this "counter-revolution" to biologist S. J. Gould's *Ontogeny and Phylogeny* (1977); or the publication two years later of "The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist program" (1979), co-authored by Gould and R. C. Lewontin. This article criticized the Neo-Darwinist theory of evolution for dividing organisms into discrete "traits" and then trying to give an evolutionary account for each of these traits individually: "Organisms are integrated entities, not collections of discrete objects" (Gould & Lewontin 1979, p. 585). Gould and Lewontin argued that the form of an adult organism is not a collection of traits individually optimized by natural selection,
1.4. The paradox of development

but an integrated whole that is subject to many different constraints, of which adaptation is only one. The Neo-Darwinist conception of the genome as a design suggests that any design is possible, and that the only constraint on these designs is adaptive success. But the adult life forms we see around us are not only products of evolution; each one is also a product of embryogenesis and development, and these processes place their own (physical, chemical, and biological) constraints on the kinds of organisms we see in the world. Thus the answer to the question of why a given species looks the way it does is not only that this form has proved well-adapted to its environment, but also that evolution has happened upon developmental pathways that could generate adult organisms of this form. The powerful suggestion here is that what evolves are not discrete traits of adult bodies, nor designs encoded in DNA that specify these traits, but rather developmental processes that consistently and robustly generate the adult form of a given species from a single cell. On this view, evolution generates new forms of life not by tweaking a design for the adult organism, but by tweaking the self-regulating developmental pathways that generate the adult form. This understanding of evolution has given rise to a whole new sub-field of biology called "evolutionary developmental biology" or "evo-devo" (Gilbert, Opitz & Raff 1996).

Rather than thinking of the adult organism as a final product manufactured according to a design, the developmental standpoint prompts us to see the adult as a single moment in an on-going process. It is just one stage in the continuous life-cycle: reproduction–embryogenesis–maturation–reproduction (Thompson 2007, p. 188). In this

53 The form of a given species is "so constrained by phyletic heritage, pathways of development and general architecture that the constraints themselves become more interesting and more important than the selective force that may mediate change when it occurs" (Gould & Lewontin 1979, p. 581).
perspective, the phenomenon of heredity is not a resemblance between adult organisms and their parents, but between a life cycle and its parent life cycle—a resemblance of process rather than product. DNA plays a vital role in this process, but it is development, not DNA, that produces the adult organism.\(^5^4\)

D. The concept of a genetic program

The discovery of gene regulation and the renewed interest in development have led some biologists to re-examine the mechanistic and gene-centric assumptions of Neo-Darwinism.\(^5^5\) However, mainstream biology has largely tried to avoid this re-examination, by abandoning the model of the genome as blueprint in favor of a new model of the genome as a program for development: a set of instructions for each cell to follow, similar to a computer program. The concept of a genetic program is supposed to be able to account for the differential development of cells in embryogenesis, which the concept of a genetic blueprint could not explain. The idea is that the program contains many "modules" which can be turned on and off: one module contains the instructions that will lead a cell to develop into a white blood cell; another module will lead to the development of a muscle cell. Like a computer program, the genetic program is supposed to consist of a series of conditional statements: if protein \(x\) is present, start synthesizing protein \(y\); if the concentration of hormone \(z\) exceeds threshold \(r\), stop synthesizing protein \(q\); and so on. This understanding of the genome attempts to account for the complexity

---

\(^{5^4}\) We will return to the role of DNA in development at the end of Chapter Two.

\(^{5^5}\) See for example the work of Stephen Jay Gould, Richard Lewontin, Evelynn Fox Keller, and Scott Gilbert.
and flexibility of development, without giving up the claim that the process of
development is specified in advance in the organism's genetic code.

Although the conception of the genome as a program for development attempts to
overcome the difficulties facing the less sophisticated picture of the genome as a
blueprint for the adult organism, it does not go far enough. Development is characterized
by its robustness and flexibility—its capacity to resist and adapt to change in order to
generate consistent outcomes in a variety of different circumstances. Programs, on the
other hand—as we know from our experience with computers—are brittle and easily
disrupted. Because programs proceed in a linear, step-by-step fashion, a tiny error in a
single line of code can cause an entire program to crash. Furthermore, programs are
flexible only to the extent that they anticipate various possible situations in advance, and
have a pre-programmed response prepared for each one. When a program encounters a
situation that it does not have a pre-set response for, it has no way to adapt.\footnote{A machine is capable only of operations for which it has been constructed; the idea of a machine which
would be capable of responding to an indefinite variety of stimuli is a contradictory one, since automation
is obtained only by submitting the initiation of work to certain chosen conditions" (SB 87/96).}

Even if DNA can rightly be conceived in some sense as a type of "program," it
must be recalled that this program is being carried out simultaneously by billions or
trillions of cells, and that all of these cells are interacting with one another in real time.
Even if we had a legible copy of the genetic program, there would be no way of
predicting the behavior of a single cell without understanding the complex dynamics of
its environment: namely, the interactions of all the cells in the developing body, moving
and changing one another in space and time. The concept of a genetic program is often
used to maintain our focus on DNA, as if it alone could account for the process of
development. But even if DNA does contain "instructions" that can turn a totipotent cell into a neuron or a skin cell, something has to trigger the initiation of one module or another. The puzzle of development is not simply how a single totipotent cell gives rise to a variety of specialized cells, but how the development of all these different cells is coordinated—without central control or external guidance—so as to produce a functioning, three-dimensional body. This coordination cannot be explained at the cellular level; it can only be accounted for in terms of the complex, cascading interactions of the cells in space and time.\textsuperscript{57}

The notion of DNA as a genetic program—like the other preformationist accounts that preceded it—is an attempt to avoid what I will term the \textit{paradox of development}: zygotes consistently develop into the complex multi-cellular bodies typical of their species—but \textit{without this adult body being fully specified in advance anywhere in the zygote}. When we encounter a system as orderly and complex as the living body, we find it impossible to conceive of this order as having arisen spontaneously, without being planned in advance. We feel compelled to trace it back to some \textit{already-given} source of order and determinacy. Indeed, preformationism and epigenesis are both variations on this same theme. Preformationism explains the form of the adult body by referring it back to a design already given in the genes of the zygote; epigenesis explains it in terms of matter and forces of nature that are given in advance, and not subject to change.\textsuperscript{58} I have

\begin{flushright}
\textsuperscript{57} "[T]he developmentally specific information resides not in the genes but rather in the spatiotemporally delimited developing system, which is therefore the ontogenetically primary unit; accordingly, interaction [between the genome and its environment] is not limited to gene activation but rather implicates positive and negative feedback loops at a variety of levels within and without the developing system and which contribute to the very constitution of the organism" (Robert 2004, p. 63).

\textsuperscript{58} "Biologists for over 200 years have been grappling with the problem of explaining organismic development in the light of an understanding of matter that does not readily lend itself to the task, resulting
argued, however, that the facts of development resist all such attempts. Development is a creative process (Robert 2004, p. 43), which generates new forms that are not fully given in advance.

If this seems paradoxical to us, it is only because we insist on understanding the living body by analogy to the things we make. I have argued, however, that the living body is not a synthetic whole: development cannot be understood in mechanical terms as the external imposition of a form on passive materials, or the unfolding of a form that has already been completely determined in advance, like the design of a machine. In order to make sense of the phenomenon of development—which is to say, the phenomenon of life—this mechanical ontology must be abandoned.

5. Conclusion

The mechanistic conception of nature has evolved over time, and so has the mechanistic account of form. We can trace this evolution through four distinct stages. We encountered each of these stages already, one by one, over the course of Chapter One, but it will be helpful to review them all together now.

The simplest version of mechanism understands nature as the product of a divine craftsman, who shaped the natural world out of matter in the same way that a human artisan shapes an artifact out of raw materials. According to this teleological mechanism, form in nature originates in a divine plan, an idea in the mind of God. Just as in an oscillation between some form of preformationism that defers the problem and an epigeneticism which finds a need to supplement the properties of living matter with special vital forces" (Moss 346). 59 "Christian philosophy had long regarded God as the artifex maximus and his creations as analogous to human art" (Des Chene 2001, p. 4).
1.5. Conclusion

the form of a human artifact is set in advance by its designer, so are the forms we find in nature determined in advance by God.

This simple mechanism slowly gave way, as Modern science developed, to a non-teleological mechanism. Rather than supposing that nature was created by a divine artisan, this new mechanism held that nature is the product of blind forces acting on atomic matter. Nature, on this view, is an artifact without an artificer. Just as an artisan shapes raw materials into complex wholes, so do the forces of nature shape matter into the wholes we see in nature. However, the forces of nature are unlike an artisan in that they do not operate according to any plan or for the sake of any goal. Whereas for teleological mechanism form is the product of an idea, form on this account is a mere accident, a side-effect or epiphenomenon. Strictly speaking, in fact, there is no such thing as form in nature: it is only our minds which unite the atoms of nature into meaningful wholes. Nature in itself is composed of discrete, unchanging and indivisible parts, and it is these parts that the forces of nature act upon. The wholes that we perceive in nature have no causal power; they make no difference to the parts that compose them. Thus form is nothing but a subjective, anthropomorphic projection of meaning onto a meaningless, mechanical nature. If form is not an idea in the mind of God, then it must be an idea in the mind of the scientist.

The 20th Century saw the emergence of two new varieties of mechanism. The first of these, which we called Neo-Darwinism, arose out of the fusion of Darwin's theory of evolution with Mendelian genetics, as we saw in Section 1.4. This third version of mechanism is a blend of the first two: in order to account for the genesis of living forms, it reintroduces the teleological notion of design; however, it avoids the appeal to
1.5. Conclusion

teleology by making this design the product of blind forces rather than a divine architect. The "forces" of evolution ("selection pressure" and "genetic drift") produce genetic blueprints for working organisms by causing tiny, random changes in existing designs and selectively propagating the "best" (i.e. fittest) variants. Thus Neo-Darwinism is a kind of mechanical teleology which attempts to explain the harmony and apparent purposiveness of living forms without resorting to the hypothesis of a divine creator. While this mechanical teleology might seem to grant a certain reality and efficacy to form in nature, this is only an appearance. In fact, the description of the genetic code as a "design" must be regarded, on Neo-Darwinism's own terms, as pure anthropomorphism. The genetic "blueprint" can only be incorporated into a mechanical universe if it is understood, not as a genuine whole, but as a collection of mutually external parts. Moreover, each of these parts, strictly speaking, must be indifferent to the fact that it specifies a part of the adult organism, for mechanical causes are indifferent to their effects. Thus it is only in the mind of the scientist that the organism's genes offer a "blueprint" for its adult body. Form remains either an idea or an epiphenomenon.

The most recent version of mechanism is a variant of mechanical teleology which arose out of the invention of digital computers. Computers are the most powerful, sophisticated, autonomous machines that human beings have ever built, so it was natural that we would use them as models for those phenomena of nature that had thus far resisted mechanical explanation—in particular, cognition and development. As we saw in Section 1.4, the original Neo-Darwinian understanding of design as a static blueprint has been replaced by the more sophisticated concept of a genetic program, a modular set of sequential instructions analogous to a computer program. The theory of a genetic
program fits better with the facts of development (particularly its flexible and distributed character) than that of a genetic blueprint. This new computational mechanism allows us to attribute a certain amount of intelligence and agency to the cells of the developing embryo. In a similar vein, cognitive scientists have attempted to explain the intelligence and agency that animals exhibit by describing animal brains as computers whose "hardware" and "software" were "designed" by the blind forces of evolution. (We will address this computational theory of mind and behavior in Chapter Three.) Thus the cells of the developing embryo are understood as tiny computers whose evolved genetic programs produce, in higher animals, a much larger computer—the brain—with its own, even more sophisticated program. However, the same thing must be said of these so-called "programs" as was said of the "blueprints" above: as long as nature is regarded as a pure mechanism, there can be "programs" only from the human point of view, and not in nature as it is in itself.\textsuperscript{60}

The mechanical understanding of nature has evidently undergone significant changes in the course of its history. Nonetheless, these four different accounts of how form arises in nature share certain essential features which justify our treating them all as varieties of mechanism. They all take nature to be a synthetic whole composed of discrete, mutually external parts—like a product of human manufacturing. (See Section 1.2 above.) These discrete parts, like the raw materials used in human manufacturing, constitute a passive, indifferent substrate, devoid of causal agency. If they are to be moved and arranged into patterns or structured wholes of the sort that we find in nature,

\textsuperscript{60} "Drawing an analogy with computer programmes is unhelpful in specifying the nature of putative genetic programmes, for a computer programme is such only on account of its relation to the intentional programmer" (Robert 2004, p. 49).
1.5. Conclusion

this must be accomplished by some external cause acting on passive matter. Thus the
various versions of mechanism share a common conception of form as a shape imposed
from the outside upon a passively indifferent material. They differ in their accounts of
what precisely the external cause is that imposes form on matter. But crucially, they all
agree that the cause of form must, like the matter on which it acts, be something given in
advance, such that the forms of things have always already been determined in advance
of the process of formation.

Form is the problem that mechanism sets for itself. On the mechanical conception
of nature, form is that which calls for explanation, that for which some cause must be
found. And yet, the forms we observe in nature are precisely that which mechanism
cannot explain, cannot find an adequate cause for. We could call this the "Humpty-
Dumpty problem": mechanistic science cannot put the phenomena back together out of
the parts it has divided them into. We are like children who open up a favorite pet in
order to see how it works, only to realize that by dividing it into parts we have killed the
very thing we wanted to understand. By insisting on understanding the parts of a complex
whole in isolation from one another, we foreclose in advance the possibility that these
parts might come to be what they are only through their interactions with one another.
We thus turn self-organizing phenomena into artificial wholes that require some external
force to organize them, and are left unable to explain how this occurs. In particular, the
genesis of living forms—with their immense complexity and astonishing harmony of
parts—cannot be accounted for by mechanical explanation, in any of its incarnations.
Development cannot be understood by analogy to manufacturing, as the external
imposition of a pre-determined form upon a passive material. For as we have seen, the
living body is a dynamic, self-organizing whole which is generated by the interactions of its own parts.

The problem of embryogenesis is not simply an empirical problem, which could be resolved by collecting more facts. It is also a conceptual problem, which demands that we revise our techno-logical concepts of matter and form. In the next chapter, I will attempt to develop an organic ontology of nature that can make sense of development. I will argue that nature is not fully given in advance, but constantly generates order and form anew, where they did not exist before. And I will show, contra atomism, that form is both real and efficacious in nature, where it is neither an idea nor an epiphenomenon.
CHAPTER TWO. How does form arise in nature?

"Human teleology constructs by assembly of machines, whereas the organism does it by auto-differentiation" (N 182/238)
1. What is form?

When we contemplate the natural world, we cannot avoid being struck by two wonders. The first is the harmony and order of nature: the coordination of its myriad parts, of its causes with its effects. The natural world does not appear to us a haphazard chaos. On the contrary, it strikes us as an orderly, intelligible whole, as if it had been orchestrated by some vast intelligence. However, when we investigate nature more closely, trying to understand the principles upon which it is organized and the mechanisms by which its order is maintained, we are confronted with a second wonder. Nothing in nature works the way we expect it to—that is, the way it would work if we humans had designed it. If nature seems to exhibit a kind of intelligence, it is a profoundly alien intelligence, one radically different from our own. The harmony and order we find in the natural world arise in very different ways than the harmony and order of the machines and artifacts we build for ourselves.

In Chapter 1, we saw how the human way of creating complex, harmonious wholes (manufacturing) differs from nature's way of doing this (growth). And we also saw how our attempts to understand growth by analogy to manufacturing inevitably break down. In this chapter, we will try to develop a more adequate account of how nature generates complex forms—one which draws its concepts not from human manufacturing, but from natural phenomena themselves. The test of this account will be whether it can succeed where mechanical thinking failed, by making sense of development—the genesis of living form.

---

61 See the "Critique of Teleological Judgment" in Kant's *Critique of Judgment*. 

55
2.1. What is form?

As long as we define matter as parts outside of parts, form can only be an idea, and never a reality. For form describes a molar being, an ensemble of relations, a whole spread out in space and time (N 157/209). There can be no form in a point or at an instant; and a form which can be decomposed into points or instants is not a form at all. Thus mechanistic science must deny the reality of form even as it tries to understand it; it can only explain form by explaining it away. But these explanations are never satisfactory.

If, on the contrary, we were to admit that matter can be altered by its interactions with itself; that it can become entangled in systems which transform their own parts; and that the parts of such systems become so mutually implicated that they can no longer be separated or even fully distinguished from one another; then we would no longer need to deny the reality of form. We could allow ourselves to recognize the existence in nature of wholes that cannot be decomposed, of "total processes whose properties are not the sum of those which the isolated parts would possess" (SB 47/49). We would then have to pose the question of form in a new way: rather than searching for the forces that could impose form on indifferent matter, we would have to ask how matter organizes itself into the forms we see in nature. We would have to learn to see form, not as a static shape, but as a dynamic movement. Rather than searching for the origin of form—a past in which the determinacy of form would already be fully given—we would have to study the genesis\textsuperscript{62} of form: the ongoing process in which existing forms complicate and transform themselves into new and more complex ones.\textsuperscript{63}

\textsuperscript{62} On the difference between genesis and origin, see Foucault 1977.
\textsuperscript{63} The focus of this dissertation is the genesis of living form—specifically, the development of a complex multi-cellular organism out of a single cell, which while much simpler than the mature organism is already
2.1. What is form?

However, this new way of thinking poses considerable challenges. How are we to understand the process of self-organization? What can we say about a whole which resists analysis? We cannot be satisfied with a vacuous holism—a Parmenidean monism which can say only "It is," or a Romantic intuitionism which rejects discursive understanding (SB 43/45). Where are we to find new concepts that can give us some purchase on the phenomena of form and development?

The last fifty years have seen a new surge of scientific interest in the phenomena of complexity and self-organization. Innovations in mathematics have revealed the existence of order in systems that had previously seemed intractably complex or "chaotic." These innovations have led to the formation of a new subfield of applied mathematics called *dynamic systems theory*, which focuses on so-called "non-linear" systems that resist decomposition into independent parts. In this chapter, I will draw on dynamic systems theory to develop a new conceptual vocabulary for thinking about how form arises in nature—and in particular, how living bodies develop.

My goal is not to offer a mathematical theory of development. Rather, I want to draw from this theory certain qualitative concepts, certain new ways of looking at dynamic, self-organizing phenomena. Mathematics has always been a source of conceptual innovation for Modern science. Too often, however, we have confused our mathematical models with the phenomena we were modeling, and so mistaken mathematical description for an end in itself, rather than a means to insight and understanding. This mistake is as old as Modern science, as old as Galileo's (1623/1957)

---

quite complex and organized. I will not attempt to address the important and difficult question of how life arises from matter in the first place. (But see p. 63 below for a brief indication of the direction such an answer might take within my account.) For more on this question, see Thompson 2007, Ch. 5.
famous declaration that nature itself is mathematical, and that only the mathematical is real (pp. 237-8). More recent attempts to reduce mathematics to formal logic—to algorithms that a computer can carry out—have only exacerbated this error. More than ever, we fall into the trap of thinking that computation, rather than understanding, is the point; that the purpose of science is to arrive at the right answer through mechanical operations, to predict rather than to understand. My hope is that we can learn from mathematics and mathematical science without falling into this trap ourselves. After all, mathematics is a language, albeit a compact and often technical one; its concepts are no less valid than those coined by philosophers. What is more, mathematics excels at the precise description of structures and relations, which is precisely the task facing us now.\textsuperscript{65}

In section 2.2, I will introduce the mathematical concept of symmetry in order to propose a new view of form as asymmetry or difference. This will allow me to re-pose the question of how form arises in nature as the question of how difference arises out of indifference, or asymmetry from symmetry. In section 2.3, I will attempt to answer this question in a general way by studying the phenomenon of symmetry-breaking in some simple inorganic systems. By introducing some key concepts from dynamic systems theory—notably the concepts of nonlinearity and stability—I will offer a rough explanation of how and why asymmetry arises in nature. Finally, section 2.4 will apply these concepts to the phenomenon of embryogenesis or anatomical growth. I will offer an account of growth as a cascade of symmetry-breaking events, and show how it is possible for the complex form of the living body to develop in a robust and reliable way without
2.2. Form as asymmetry

this form being imposed on the body from the outside or determined in advance by a blueprint or program.

2. Form as asymmetry

The first mathematical concept I want to introduce is the concept of symmetry. I noted above (in Section 2.1) that it is the order in nature that seems to call for explanation, the harmonious coordination of its parts into complex, patterned wholes. Symmetry is a concept that mathematicians use to talk about order, pattern and form in a precise and rigorous way. In everyday language, too, we associate symmetry with order, and even with beauty. However, the connection between order and symmetry—in the technical, mathematical sense of this concept—may surprise you. In this rigorous sense, greater symmetry does not mean greater order. On the contrary, order and form arise from the loss or breaking of symmetry (Ball 2009a, pp. 20-25; Stewart and Golubitsky 1992, pp. 2-5).

![Fig. 2](image)

65 Indeed, it is ironic that mathematics has been so co-opted by positivist philosophies in the last century, since mathematics is the least positivist of all disciplines. There are no substances in mathematics; a mathematical entity can only be defined by its relation to other mathematical entities.
2.2. Form as asymmetry

Consider the three images in Fig. 2. Which is the most symmetrical, and which is the least? Mathematicians measure symmetry by the number of transformations which leave a thing unchanged. The star in the box on the left is symmetrical (which is to say invariant) under six rotations and six reflections (i.e. "flips" or "mirrorings" across a given axis), for a total of twelve symmetries. The circle in the middle box, on the other hand, is symmetrical under an infinite number of rotations around its center and reflections across axes passing through its center; thus it has a much higher degree of symmetry than the star on the left. And the uniform, white field in the box on the right is even more symmetrical than the circle; it is invariant under rotations and reflections around any axis, as well as translations (shifting or sliding in the plane) in any direction.

Contrary to what we might expect, then, greater symmetry does not imply greater order or structure. On the contrary, the greatest symmetry belongs to structureless uniformity (like the homogeneous field in Fig. 2). Form arises through the breaking-up of this uniformity, the introduction of differences which break one or more of its symmetries: the circle breaks the limitless symmetries of the white field by introducing a privileged point, the center, which is the axis of all symmetrical rotations, and through which the axes of all symmetrical reflections must pass; and the star on the left breaks these symmetries still further by introducing certain privileged axes which were not present in the circle.

One might expect that the problem of how nature generates pattern and form would be to explain how symmetry arises out of chaos and disorder. But in fact, disorder is much more symmetrical than order. If we take a beautiful bronze sculpture and melt it
2.2. Form as asymmetry

down into a uniform pool of liquid metal, its form and structure is lost—but it gains a
great deal of symmetry. Thus the question of the genesis of form is not how symmetry
arises out of disorder, but rather how the symmetry of disorder gets broken in determinate
ways to produce the characteristic *asymmetries* of the forms we find in nature.

We expect nature to be symmetrical. That is, we expect that things will be the
same in one place or time and another unless there is some *reason* for them to be
different. This is a deep presupposition of Modern science. How can we claim that
experiments performed here on Earth reveal *universal* laws of nature, which are true
everywhere and for all time? These claims are based on the *indifference* of our
experimental results to the time and place where they occur (Morris 2006a, pp. 50-51). If
an experiment assumes nothing about its location, and the same results can be obtained
again and again by different observers at different times and places, then we have reason
to believe that the results are universal, i.e. indifferent to where/when they occur. In other
words, we assume that nature is the same everywhere unless there is some reason for it to
differ, some difference that would make a difference to the phenomenon we are studying.

However, if nature were *perfectly* symmetrical, there would be no phenomena to
study—and no one to study them. A perfectly symmetrical nature would be perfectly
uniform, entirely devoid of differences. When we seek to understand the order we find in
nature, we are asking after the origin of nature's *differences*, of its *asymmetries*. We are
asking why things are different in one place than they are in another, or why they are
different now than they used to be. The basic question of Modern science, then, is not
"Why is there something rather than nothing," but "Why is there difference rather than
indifference or uniformity?"
2.2. Form as asymmetry

Because Modern science expects nature to be symmetrical, it is nature's asymmetries that call for scientific explanation. However, our expectation that nature is symmetrical leads us to assume that every difference must have a reason, must be explained by some cause. In other words, every difference must be the product of some prior difference; an asymmetrical effect must be the product of an equally asymmetrical cause. Or, stated conversely: symmetry is conserved; a symmetrical cause has an equally symmetrical effect.\(^{66}\) This is the implicit assumption of the mechanical explanations we examined in Chapter One, which can only explain form by showing that it was somehow given in advance. It is difference that calls for explanation; but mechanical explanation operates by transforming differences into discrete, self-identical things, and assuming that each thing must have its own discrete cause.\(^{67}\) For example, Neo-Darwinism attempts to explain the highly asymmetrical form of the living body by analyzing it into a collection of discrete traits, each of which would be determined in advance by a discrete gene. Thus the genome would already contain, prior to development, all the differences that constitute the adult form.

If it were true that symmetry is conserved in nature, form would be a constant in the universe, neither created nor destroyed. Every asymmetry would have its explanation in some prior asymmetry, all the way back to the beginning of time. But this is not, in fact, what science tells us about the history of the universe. According to our current best theories, the universe began at a single point from which it exploded outward in all

\(^{66}\) This principle was formulated explicitly by Pierre Curie (1894): "Lorsque certaines causes produisent certains effets, les éléments de symétrie des causes doivent se retrouver dans les effets produits. Lorsque certains effets révèlent une certaine dissymétrie, cette dissymétrie doit se retrouver dans les causes qui lui ont donné naissance." See also Stewart & Golubitsky 1992, pp. 7-8.
directions (the "Big Bang"). At its origin, then, the universe was highly symmetrical. If symmetry were conserved, one would expect to find that the universe today had perfect spherical symmetry, centered on its point of origin. Matter and energy would be distributed uniformly throughout the cosmos. There would be no galaxies, no stars, and no life. Of course, this is not the universe we live in. Since the Big Bang, the universe's initial symmetry has been broken again and again. As the universe expanded, matter became concentrated at certain points in space. Over time, these points of higher density developed through further losses of symmetry into galaxies, stars, and planets.

These cosmological asymmetries—in particular, the energy gradient generated by the Sun's nuclear fusion, and the Earth's location in that gradient—created the conditions for the emergence of life on Earth. But this emergence required its own further cascade of symmetry-breaking events: the formation of early proto-cells broke the symmetry of the "primordial soup" by enclosing their chemical contents and isolating them from their surroundings; much later, multicellular life formed when cells began to stick together, forming newly asymmetrical structures. While we cannot go back in time to witness the origin of life on Earth, we can watch a single cell transform itself into a multicellular body: this process takes place in the life of every multicellular organism as it develops from a zygote into a mature adult. This process of development—which biologists call morphogenesis—is another cascade of symmetry-breaking events, in which a single spheroid cell differentiates itself into a highly asymmetrical body with a front and a back, a top and a bottom, a left and a right.

67 "A mechanical action… is one in which the cause and the effect are decomposable into real elements which have a one-to-one correspondence" (SB 160-61/174).
2.3. Symmetry-breaking as the genesis of form

For atomism, the problem of form is a problem of synthesis: to explain how atoms are put together into the complex wholes that we observe in nature. But if we understand form as asymmetry, as I am proposing, then the problem of morphogenesis becomes a problem of differentiation rather than synthesis: the question is not how forms are put together out of already-differentiated parts, but rather how nature differentiates itself into parts in the first place. This allows us to pose the problem of embryogenesis, which we encountered above (1.4), in a new and more fruitful way: how does an embryo develop from a relatively symmetrical, spheroid zygote into the much less symmetrical form of a mature human body? We will return to this question in section 2.4. But in order to answer it, we must first understand the process of symmetry-breaking in more detail. How can we explain the differences we see in nature without appealing to pre-existing differences? What causes a system—be it a solar system or a living body—to become less symmetrical over time? How does form—now understood as asymmetry—arise in nature?

3. Symmetry-breaking as the genesis of form

We began this chapter with the question of how form—particularly the complex form of the living body—arises in nature. In the previous section, I proposed a new view of form as difference or asymmetry. On this view, the genesis of form in nature takes

---

68 This is a version of what Waldenfels (2000) calls "the paradox of expression" in Merleau-Ponty, which we might also call the problem of the genesis of sense. The paradox is that expression must be different from what it expresses, for otherwise it would be a simple repetition rather than a true expression; but at the same time, expression must be the same as what it expresses in order to be an expression of it. (See Lawlor 1998; Morris 2006b.) Here we see the same problem with respect to causal explanation: the cause must be the same as its effect in order to explain why the effect is the way it is; and yet, if the cause is the same as its effect then it does not explain it at all, but merely postpones explanation. (See Morris 2006a.)
2.3. *Symmetry-breaking as the genesis of form*

place through the *loss or breaking* of nature's symmetries—the genesis of difference from indifference. In this section, we will study the phenomenon of symmetry-breaking, in order to understand how and why asymmetries arise in nature. In section 2.3.A, we will look at some examples of symmetry-breaking in simple, physical systems, in order to arrive at a basic description of *how* symmetry-breaking occurs. This description will introduce three key concepts from dynamic systems theory: *nonlinearity, critical points,* and *symmetry-breaking bifurcations.* In sections 2.3.B and C, I will try to explain *why* symmetry-breaking occurs by introducing two more concepts from dynamic systems theory: *stability* and *instability.* Finally, section 2.3.D will draw out some of the ontological implications of these concepts.69

*A. Symmetry-breaking bifurcations in nonlinear dynamic systems*

We will begin our study of symmetry-breaking by looking at a few simple examples, in order to arrive at a basic description of *how* symmetry-breaking occurs in nature. Consider a simple physical system: a load-bearing, cylindrical beam standing on its end (*Fig. 3*). There are two antagonistic forces at work in this system: the load exerts a downward force on the beam, and the structural integrity of the beam resists this pressure by pushing back against the load, supporting its weight. Now, suppose that the weight of the load can be varied continuously in real time. We are now considering a *dynamic* system—a system whose state is changing over time. When the load is light, the structural integrity of the beam is sufficient to support its weight. What will happen if we

---

69 For the reader who wants a bit more mathematical detail, I have supplied a longer and slightly more technical introduction to dynamic systems theory in Appendix A.
2.3. *Symmetry-breaking as the genesis of form*

gradually increase the load? At first, nothing: the beam maintains its shape and position. At a certain point, however, the beam can no longer support the weight being placed on it, and it abruptly buckles under the pressure, assuming a new, curved shaped.

![Image of a load-bearing beam](image)

**Fig. 3. A load-bearing beam** (Strogatz 2001, p. 44)

Three features of this system's behavior deserve our attention. First, notice that the shape of the buckled beam is *less symmetrical* than its earlier, vertical state. Before it buckled, the cylindrical beam was *indifferent* to rotations around its vertical axis. In other words, there was no way to distinguish one horizontal direction from another in terms intrinsic to the system itself. We could try to label these directions if we wanted to; but there would be no way to distinguish our beam from one that had been rotated by 180°. Thus our labeling would necessarily be arbitrary—an external imposition that finds no purchase within the system itself. When the beam buckles, however, this indifference is lost—the system itself now distinguishes one direction from all the others. Thus the buckling of the beam is a *symmetry-breaking* event, in which our dynamic system loses some of its initial symmetry.
Second, notice how the global state of the system—namely, the shape of the beam—changed as we varied the load on the beam. We generally expect that a small variation in one part of a system will produce a small change in the system's global state—and similarly, that a large variation will produce a large result. In other words, we expect systems to behave in a proportional or linear way (Fig. 4). However, this is not what we observed in the case of the buckling beam. The relation between the load on the beam and the beam's shape was nonlinear: changes in one did not produce proportional changes in the other. Instead, we found that variations in the load produced no observable change in the system's global state as long as the load remained below a certain critical value. When this critical point was reached, however, a very small increase in the load suddenly produced a disproportionately large change in the system's state—the buckling of the beam (Fig. 4).^70 Thus a gradual, quantitative change in one parameter of the

---

^70 You can see why the study of systems that exhibit nonlinear behavior has sometimes been called "catastrophe theory."
2.3. *Symmetry-breaking as the genesis of form*

system, when it reached a certain critical point, produced an abrupt *qualitative* change in the system—a loss of symmetry.

![Diagram showing a system state splitting](image)

**Fig. 5. Linear behavior in a dynamic system**

Lastly, notice that when the beam buckles, it may assume one of several possible states—all of them less symmetrical than its initial, vertical shape. In dynamic systems theory, symmetry-breaking events such as the buckling of the beam are represented visually in what are called *bifurcation diagrams* (Fig. 5). In these diagrams, the line representing the state of the system splits when the "control parameter"—in this case, the load on the beam—reaches a critical point. Each branch of the line represents one of the new, less symmetrical states that the system can assume at this critical point. The term "bifurcation" is something of a misnomer, since there may be far more than two such possible states; nevertheless, these diagrams have taken their name from the simplest case, and symmetry-breaking events have become known in dynamic systems theory as *symmetry-breaking bifurcations*. As you can see in Fig. 5, a system may have several
2.3. *Symmetry-breaking as the genesis of form*

critical points, and thus may undergo a whole *cascade* of symmetry-breaking bifurcations. In these cases, the system's bifurcation diagram becomes a "bifurcation tree," which shows how the system in question can become progressively less symmetrical—which is to say, more complex—as the control parameter varies. Notice again how the gradual, quantitative variation of the control parameter produces a series of abrupt, nonlinear, qualitative changes in the state of the system.

![Bifurcation diagram](image)

**Fig. 6. Bifurcation diagram** (adapted from Mainzer 2005)

Precisely *which* new states are available to the system when it reaches a critical point depends on the system's initial symmetries (Stewart and Golubitsky 1992, Ch. 3). Since our beam has a circular cross-section, it can buckle in any direction. A beam with a square cross-section, on the other hand, could buckle in one of only four directions (Fig. 6). The set of all possible states that the beam can assume when it buckles has the same number of symmetries as the unbuckled beam. This follows from the fact that the
2.3. Symmetry-breaking as the genesis of form

Symmetries of a system represent *differences to which it is indifferent.* For the circular beam, one direction is as good as another; for the square beam, however, there are four privileged directions (corresponding to the four rotations that leave the square beam unchanged). The more symmetrical a system is, the greater the number of states it can assume at the critical point. If the system could somehow take on all of these states at once, it would remain just as symmetrical as it was before. But of course, this is impossible; at the critical point, only one of these states can be "chosen" by the system, which means that the system must become less symmetrical.

![Fig. 7. Directions in which a beam can buckle depend on its initial symmetries](image)

It is impossible to predict in advance, on the basis of the system's own internal characteristics, *which* of these possible states the system will "choose." Prior to the loss of symmetry, the system is precisely *indifferent* to the difference between one of these states and another—if it were not, then this symmetry would already have been broken. Thus there is no reason for the system to "prefer" one of these states over another; it is only *after* the loss of symmetry has occurred that we can even distinguish these states from one another with respect to the system itself. This poses a puzzle, however: if there
is no reason for the system to prefer one state over another, then how can it choose
between them? This brings us back to the problem we posed in the previous section: how
to explain the differences we see in nature without appealing to some pre-existing
difference? The phenomenon of symmetry-breaking—the genesis of difference from
indifference—presents us with a paradox. In order to become less symmetrical, a system
must become less indifferent by choosing one of several equivalent possibilities; but if
these possibilities are truly equivalent, then there is no way to choose between them.
Either the choice has already been made, or it cannot be made; if things are not already
different, then they can never become so. We will return to this symmetry-breaking
paradox below (in 2.3B). But first, let us look at another example of symmetry-breaking
in nature.

The buckling beam is an example of symmetry-breaking in *space*: when the beam
buckles, it assumes a less symmetrical *shape* then it had before. A system that underwent
a whole cascade of such symmetry-breaking bifurcations could develop shapes of great
complexity; indeed, this is precisely what we will see below when we study the
anatomical development of the living body (Section 2.4). For now, though, let us
consider another simple physical system. This system is an example of symmetry-
breaking in *time* rather than space. The system in question was discovered in the early
1950s by a Russian scientist named Boris Belousov who was trying to create a simple test
tube version of the Krebs cycle—a chemical reaction central to animal metabolism (Ball
2009a, pp. 110-119; Strogatz 2001, p. 134). Mixing a few simple reagents together with a
catalyst, Belousov found to his surprise that the yellow mixture faded to clear after about
2.3. *Symmetry-breaking as the genesis of form*

a minute, then regained its yellow color a minute later, and continued to oscillate regularly between these two states for an entire hour. Belousov had discovered the first *oscillating chemical reaction.* The spontaneous oscillation of chemical reactions is now a well-known phenomenon. It is interesting for us because it represents a loss of symmetry in time, rather than in space. As long as the concentration of certain catalysts remains below a certain critical value, the mixture remains in a steady state. This state is uniform with respect to time, and therefore highly symmetrical with respect to *time-transformations* such as shifting the system forward or backward in time (translation), or time-reversing the system's behavior (reflection). In its initial, steady state, all of these transformations leave the system unchanged. When the concentration of the catalysts reaches a certain critical point, however, a symmetry-breaking bifurcation occurs, and the concentrations of the reagents abruptly begin to fluctuate in a regular, periodic pattern (Fig. 8). This bifurcation breaks most of the system's initial temporal symmetries. Post-bifurcation, the only time-shifts that leave the state of the system unchanged are those that shift the system forward or backward by the exact *period* of the system's oscillation, i.e. the amount of time it takes for the oscillating system to return to the same state. In other words, rhythm or periodicity is a temporal symmetry; but a periodic system is much less time-symmetrical than a system that is not changing at all. In time, as in space, order, structure, and pattern emerge through the breaking of symmetry.

---

71 At the time, chemists believed that chemical reactions always move toward equilibrium in a linear way. Indeed, this belief was so entrenched that no one took Belousov's discovery seriously until years later, when another chemist named Zhabotinsky confirmed it.
2.3. *Symmetry-breaking as the genesis of form*

![Diagram of system state, old order, critical point, and new order over time.](image)

**Fig. 8. Dynamics of an oscillating chemical reaction before and after bifurcation**

Finally, let us look at a third physical system which undergoes a symmetry-breaking bifurcation in both time and space. Consider a thin layer of fluid being heated from below and cooled from above, like oil in a frying pan (Ball 2009a 24-25; Stewart & Golubitsky 1992, pp. 118-122; Kelso 1995, pp. 6-8). As long as the temperature difference between the top and bottom layers of the fluid remains below a certain critical value, there is no large-scale motion in the fluid. At the microscopic level, of course, individual molecules are moving around at random; but all of these microscopic movements average out to rest at the macroscopic level, since the individual molecules do not "favor" any particular direction of movement. This fluid is highly *disordered*, which means that it is highly *uniform*, highly *symmetrical*. It makes no difference where in the fluid you look: all of the molecules are doing the same thing, namely moving randomly in all directions. When the temperature difference between the top and the bottom of the fluid increases past a certain critical point, however, the system undergoes a symmetry-breaking bifurcation in which the random micromotions of the molecules are...
2.3. *Symmetry-breaking as the genesis of form*

replaced by large-scale, coordinated movement. The warmer fluid at the bottom rises, while the cooler layer at the top sinks, and these two movements together generate a pattern of convection rolls (Fig. 9). Each convection roll rotates in either the clockwise or counter-clockwise direction. Once again, it is impossible to predict which of these two states the system will fall into when it reaches the critical point. In either of these new states, the system is far less symmetrical than it was before the bifurcation. Whereas before, it made no difference to a given molecule where in the fluid it was located, the motions of individual molecules now depend a great deal on their location within the macroscopic convection pattern.

![Fig. 9. Convection rolls in a fluid heated from below](image)

This loss of symmetry represents a significant increase in the order of the system. Before the bifurcation, the parts of the system (namely the individual molecules of the fluid) were largely independent of one another, and of the system as a whole. Consequently, the system exhibited very little order or organization. When the bifurcation occurs, the parts of the system begin to interact with one another in a complex
2.3. *Symmetry-breaking as the genesis of form*

way, organizing themselves into a stable, system-wide pattern that coordinates the movements of each part. This system—known as Rayleigh-Bénard convection after two theorists who investigated it—is a classic example of *self-organization* in (inorganic) nature. Notice that the transition from disorder to order is driven by a simple temperature gradient, which cannot in any way be said to contain or pre-determine the orderly arrangement of the convection rolls. Rather, when the energy gradient breaks the symmetry of the uniform, unmoving liquid, the system falls into a new, less symmetrical state in which the parts of the system are no longer indifferent to one another. The order and coordination of the parts do not need to be imposed on them from the outside, or determined in advance by some design; they come "for free" when the symmetry of the system is broken.

**B. Stability and instability**

By studying symmetry-breaking phenomena in three simple physical systems, we have arrived at a basic description of how symmetry-breaking happens—and thus, of how form arises in nature. We observed that natural systems can behave in a *nonlinear* way: a change in one part of the system can produce a *disproportionate* change in the system as a whole. I introduced the concept of the *critical point* to describe the way that a gradual, quantitative variation in one part of a nonlinear system can produce an abrupt, qualitative change in the system's global state. We also observed that this change consists in the system's falling into one of several possible new states, all of which are less symmetrical than the system's previous state. To describe this phenomenon, I introduced the concept
2.3. Symmetry-breaking as the genesis of form

of a symmetry-breaking bifurcation, and the technique of drawing bifurcation diagrams to visualize the various new states that the system can assume when it reaches a critical point. However, these descriptions raised a number of new questions. What exactly happens to the system when it reaches a critical point? What causes the loss of symmetry? And how does the system "choose" which new, less symmetrical state to fall into, if the system is symmetrical (which is to say, indifferent) with respect to these states?

In order to answer these questions, I must introduce a new pair of concepts: stability and instability. Consider a simple dynamic system: a marble sitting at the bottom of a bowl. Left to itself, the marble will remain at rest. If some external force disturbs the marble, pushing it away from the bottom of the bowl, it will soon return to its original position. In dynamic systems theory, the marble's position at the bottom of the bowl is called a stable fixed point or a stable attractor. It takes energy to push the system away from this point, and the system will return to this point if left to itself. External forces can introduce fluctuations into the system, causing the marble to oscillate back and forth. However, these fluctuations are quickly damped out: the magnitude of the marble's oscillations rapidly decreases until it has once again returned to its resting position at the bottom of the bowl. No matter where the marble is placed in the bowl—that is, no matter what the system's initial conditions are—the marble will always end up resting in the same place. To say that this position is stable is to say that the system robustly produces and maintains it, regardless of its initial conditions, and in the face of fluctuations caused by external forces. This is not to say that the stable outcome is inevitable: external forces can keep the marble away from its stable attractor, or even send it flying out of the bowl
altogether. Within a certain range of initial conditions, however, and in the absence of external interference, the system will consistently return to its stable fixed point.

Now, imagine that the bowl could be inverted without disturbing the marble. Instead of resting at the bottom of a bowl, the marble is now sitting at the top of a peak. So long as no external force disturbs the marble, it can remain in this position indefinitely. However, this state is highly unstable: the slightest disturbance will send the marble tumbling away down the slope. The marble's position at the top of the inverted bowl is called an *unstable fixed point*: a system that begins in this state will remain there indefinitely in the absence of external forces, but the slightest fluctuation will send the system away from this state, and toward the nearest stable attractor.

This, in essence, is what happens to a dynamic system when it undergoes a symmetry-breaking bifurcation. When the control parameter reaches a critical point, the system's old state becomes unstable. Like the marble at the top of the bowl, the system can remain in this unstable state indefinitely so long as nothing disturbs it. The slightest perturbation, however, will send the system away from this unstable state toward one that is more stable—but less symmetrical.

![Potential landscape with stable fixed point at x = 0](image-url)

*Fig. 10. Potential landscape with stable fixed point at x = 0*
2.3. *Symmetry-breaking as the genesis of form*

This image of the marble in a bowl can be used to visualize the behavior of any dynamic system with a stable fixed point, using what dynamic systems theory calls a potential landscape diagram (Fig. 10). Here the \(x\)-axis represents the global state of the system, and the \(V\)-axis represents the potential energy of the system—which we can think of as the amount of work required to move the system from one state to another, or the amount of work that the system itself will do if allowed to move from one state to another on its own. To read this graph, think of it as a landscape, and think of the state of the system as a heavy ball rolling through this landscape (Fig. 11).

![Fig. 11](image)

The potential landscape allows us to visualize the behavior of a dynamic system. Movement along the \(x\)-axis represents a change in the system's global state. The slope of the landscape represents the work required to produce such a change: it takes work to push the system "up-hill"; moreover, the system will tend *of its own accord* to roll "down-hill" toward the lowest points in potential the landscape. Places where the landscape is flat correspond to the system's fixed points: they are places where the system...
2.3. Symmetry-breaking as the genesis of form

can come to rest. In the landscape above (Fig. 11), we can see that \( x = 0 \) is a stable fixed point of this system: if a fluctuation moves the system "up the slope" in either direction, it will roll back down again to rest at this fixed point. We can indicate where the stable fixed points in the potential landscape are with a black ball (Fig. 12).

A system with an unstable fixed point will have a different potential landscape (Fig. 13). Here we can see that \( x = 0 \) represents an unstable fixed point: the system is at rest in this state, but the slightest fluctuation will send the system "rolling down-hill" away from this point in one direction or another, toward the nearest stable attractor (not shown). We can indicate the unstable fixed points in the landscape with a white ball.
2.3. *Symmetry-breaking as the genesis of form*

![Diagram of potential landscape with one unstable fixed point](image1)

**Fig. 13. Potential landscape with one unstable fixed point**

These potential landscape diagrams become especially useful when we consider systems with *more than one* fixed point, which are known as *multistable systems*. For example, consider the system whose potential landscape is shown in Fig. 14. This system has five fixed points: three stable, and two unstable. The fixed point at \( x = 0 \) is the global minimum, meaning that it is the system's most stable fixed point. However, a large enough fluctuation in \( x \) could push the system over into the *basin of attraction* of one of the system's other two fixed points. Thus a multistable system is capable of *multiple stable behaviors*; and a sufficient "push" in the right direction can shift it from one stable state into another.

![Diagram of potential landscape of a multistable system with five fixed points](image2)

**Fig. 14. Potential landscape of a multistable system with five fixed points**
2.3. Symmetry-breaking as the genesis of form

The concepts of stability and instability help us to understand what happens to a dynamic system when it undergoes a symmetry-breaking bifurcation. In particular, these concepts help us to understand why some dynamic systems behave in a nonlinear way. When we change one part of a nonlinear dynamic system, this change may not produce an immediate effect in the system's global state. Instead, it may alter the system's potential landscape. Critical points are places where these variations produce a qualitative change in the potential landscape: causing a fixed point to appear or disappear, to change from stable to unstable, or vice versa. For example, consider a system that undergoes a simple symmetry-breaking bifurcation when one of its parameters exceeds some critical point (Fig. 15). At first the system rests at a stable fixed point, represented by a deep basin or well in the potential landscape (A). As the control parameter approaches the critical point, however, the stability of this initial state steadily decreases. In the system's potential landscape, the potential basin that the system is resting in becomes increasingly wide and flat (B). As a result, the system becomes increasingly sensitive to external forces: disturbances that would only have produced small changes in the system's state before now produce larger and larger fluctuations, and the system takes longer and longer to "damp them out" and return to its stable resting point. When the control parameter exceeds the critical point, the fixed point that was once at the bottom of a potential well becomes a local maximum—the top of a peak in the potential landscape (C, D). What had been a stable attractor of the system is now an unstable fixed point. The system can still remain in its formerly stable state indefinitely,
2.3. Symmetry-breaking as the genesis of form

so long as no external forces act upon it. However, the slightest fluctuation will now send the system "rolling downhill" away from this fixed point, toward some newly stable state.

![Changing potential landscape of a system undergoing a bifurcation](image)

**Fig. 15. Changing potential landscape of a system undergoing a bifurcation**

This symmetry-breaking bifurcation can also be represented in a bifurcation diagram (Fig. 16). Now that we understand the concepts of stability and instability, we can make our bifurcation diagram more precise. The stable fixed point that the system was in prior to the symmetry-breaking bifurcation does not disappear entirely; rather, it becomes an *unstable* fixed point, which we can represent with a dotted line. The
2.3. Symmetry-breaking as the genesis of form

bifurcation of the solid line represents the appearance, at the critical point, of two new stable fixed points—each of which is less symmetrical than the system's stable state pre-bifurcation. We now know that this abrupt bifurcation is the result of gradual changes in the system's potential landscape (Fig. 15). Because of its shape, this simple bifurcation is called a *pitchfork bifurcation*.

![Diagram of a pitchfork bifurcation](image)

**Fig. 16. "Pitchfork" bifurcation**

Not only do the concepts of stability and instability explain the surprising phenomenon of nonlinearity, but they also resolve the other puzzle we posed above, which I called the "symmetry-breaking paradox": how does a dynamic system "choose" one new state over another when it undergoes a symmetry-breaking bifurcation, given that the system pre-bifurcation is constitutively indifferent to the differences between these various possible states? (If it were not, then it would already have lost the symmetry in question.) The answer is that the *system itself does not choose*; indeed, left to itself, the system would remain in the same (unstable) state indefinitely, never losing its initial
2.3. Symmetry-breaking as the genesis of form

Symmetry. However, systems in nature are never isolated; they are always subject to disturbances from outside forces, however small. It is these external forces that cause the newly unstable system to fall into a less symmetrical (but more stable) state, and it is they—not the system itself—that decide which newly stable state the system will fall into. In the case of the buckling beam, for example, when the load on the beam reaches a critical point, the structure of the beam becomes unstable. In the absence of external forces, the beam could retain this unstable, vertical state indefinitely. However, the slightest perturbation will now cause the beam to buckle—falling out of its unstable state into a more stable but less symmetrical shape. Which way the beam buckles depends on the external forces that "knock" the system out of its unstable state. Notice, however, that these external forces are not responsible for the order that arises from the system's loss of symmetry. They determine which new state the system falls into, but all of these states are equally asymmetrical—which is to say, equally orderly. The stability of these less symmetrical states—the fact that the system is attracted to them—is not due to external forces, but to the system's endogenous dynamics.

C. Limit cycles and Hopf bifurcations

What about our other example, the oscillating chemical reaction? Potential landscapes are useful for visualizing a system's fixed points—places where the system comes to rest—but not so good at showing the dynamics of a system in motion. To do this, we can turn to a different tool: the phase portrait.72 Think of the phase portrait as a potential landscape seen from above, instead of in relief; a kind of contour map that

---

72 For more on phase portraits, see Appendix A, pp. 217ff.
2.3. *Symmetry-breaking as the genesis of form*

shows the paths a system's state will follow as it moves through the system's potential landscape. For a system with a single stable fixed point (Fig. 17A), the phase portrait might look something like Fig. 17B. The state of the system is represented by two variables, \( x \) and \( y \). (In the case of the oscillating chemical reaction, \( x \) and \( y \) would represent the concentrations of two different reagents.) Every point in the phase portrait represents a possible state of the system, and the arrows represent changes in the state of the system over time. In this phase portrait, we can see that all the system's trajectories spiral into a point. This point is a stable attractor for the system: no matter what the system's initial state is, it will eventually end up resting at this point; and perturbations that push the system away from this point will be damped out. This phase portrait describes the state of the chemical oscillator *before* its bifurcation: when the concentration of the catalyst is below the critical value, the system rests at a stable fixed point, and it will damp out fluctuations in the reagent concentrations.
2.3. *Symmetry-breaking as the genesis of form*

**Fig. 17. Potential landscape (A) and phase portrait (B) for a system with a single, stable fixed point*73

When the catalyst exceeds a critical concentration, however, a bifurcation occurs. The system's stable fixed point becomes unstable, and a new stable attractor appears. However, this attractor is not another fixed point, but rather a periodic oscillation, illustrated by the phase portrait in Fig. 18. The circular trajectory represents a periodic behavior in which the system cycles repeatedly through the same series of states, without ever coming to rest in any one state. A closed trajectory like this is called a *limit cycle*. Like a fixed point, a limit cycle can be stable or unstable. You can see from its phase portrait that this system's limit cycle is stable: all the system's trajectories spiral into it, indicating that the system will consistently fall into this pattern, regardless of its initial conditions, and that it will damp out perturbations that push it away from this trajectory.

For comparison, Fig. 19 shows the phase portrait of a system with an *unstable* limit cycle. The system can stay in this cycle indefinitely, but any perturbation will send it spiraling away in one direction or another toward a more stable state.

---

73 See Appendix A, pp. 222-3.
2.3. Symmetry-breaking as the genesis of form

Fig. 18. Phase portrait for a system with a stable limit cycle

Fig. 19. Phase portrait for a system with an unstable limit cycle
2.3. Symmetry-breaking as the genesis of form

The bifurcation that the chemical oscillator undergoes—in which a stable fixed point becomes unstable and spawns a stable limit cycle—is called a Hopf bifurcation. We can draw a three-dimensional bifurcation diagram to visualize this bifurcation (Fig. 20). When the control parameter (in this case the concentration of a catalyst) is below a critical value, the system rests at a stable fixed point, indicated by a solid line with a trajectory spiraling into it, (as in a phase portrait). At the critical point this fixed point becomes unstable, and a new attractor appears: a stable limit cycle, indicated by a closed orbit with a trajectory spiraling into it and away from the now-unstable fixed point.

Fig. 20. Hopf bifurcation

D. Nonlinearity, indecomposability, and unpredictability

The concept of nonlinearity lets us see how continuity can give rise to discontinuity, how quantitative variation in a system's parameters can give rise to qualitative novelty in the system's behaviors. This is of special interest for our present study, since our whole project is to understand how nature generates new forms of order
2.3. *Symmetry-breaking as the genesis of form*

which are both continuous and discontinuous with the older forms on which they are founded.\(^7^4\) We saw in the previous chapter (1.3) that the mechanistic view of nature makes time into an illusion: since matter and the forces of nature that shape it are immutable, the changes we observe in nature have all been determined in advance. What we experience as the coming to be and passing away of natural phenomena are nothing but the movements of unchanging atoms, which are entirely indifferent to the macroscopic forms their movements create. It is worth quoting once again the statement of Laplace (1814/1951) which sums up this view so well:

> Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes. (p. 4, translation modified)

This view of nature as entirely decomposable into simple parts whose movements can be completely predicted in advance presupposes that nature is *linear*. Recall that in a linear system, effects are always proportional to their causes. As a result, causes and effects in a linear system are *additive*: the effect of two causes put together is simply the sum of the effects each would have on its own. This means that complex phenomena in a linear system can be decomposed into simple causes and effects, which makes linear systems highly predictable. In a nonlinear system, on the other hand, causes can have disproportionate effects: a small change can make a big difference, and a big change can make no difference at all. Thus causes and effects in a nonlinear system are *not additive*:

\(^{7^4}\) "It can happen that, submitted to external forces which increase and decrease in a continuous manner, the system, beyond a certain threshold, redistributes its own forces in a qualitatively different order which is nevertheless only another expression of its immanent law. Thus, with form, a principle of discontinuity is
2.3. *Symmetry-breaking as the genesis of form*

one cannot predict what a combination of causes will do by looking at the effects of each cause in isolation. This means that complex phenomena in a nonlinear system *cannot be decomposed into simpler parts*; and this makes the behavior of nonlinear systems inherently unpredictable.\(^7^5\)

Our measurements of natural phenomena are never perfectly accurate; they are always approximations. In a linear system, this is not a problem, since small errors in measurement will only lead to small errors in the predictions of our model. If we want our predictions to be more accurate, all we have to do is increase the accuracy of our measurements. In a nonlinear system, however, a small measurement error can produce a *disproportionately large* error in our model's predictions.\(^7^6\) Moreover, the problem cannot necessarily be solved by improving the accuracy of our measurements, since this improvement may not produce a proportional improvement in the accuracy of our predictions.

---

\(^7^5\) "Why are nonlinear systems so much harder to analyze than linear ones? The essential difference is that linear systems can be broken down into parts. Then each part can be solved separately and finally recombined to get the answer. This idea allows a fantastic simplification of complex problems, and underlies such methods as normal modes, Laplace transforms, superposition arguments, and Fourier analysis. In this sense, a linear system is precisely equal to the sum of its parts. But many things in nature don't act this way. Whenever parts of a system interfere, or cooperate, or compete, there are nonlinear interactions going on" (Strogatz 1994, pp. 8-9). For more on this point, see Appendix A.

\(^7^6\) "If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment. But, even if it were the case that the natural laws had no longer any secret for us, we could still only know the initial situation *approximately*. If that enabled us to predict the succeeding situation *with the same approximation*, that is all we require, and we should say that the phenomenon had been predicted, that it is governed by laws. But it is not always so; it may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have the fortuitous phenomenon" (Poincaré 194, p. 68). This property of nonlinear systems is commonly described as "sensitivity to initial conditions." The famous image of the butterfly flapping its wings and causing a hurricane on the other side of the planet is an illustration of this sensitivity—weather patterns being one of the most well-known examples of nonlinear systems.
2.3. *Symmetry-breaking as the genesis of form*

Scientists tend to focus on problems they can solve, and areas of investigation where existing methods prove fruitful; this is part of why Modern science has been so successful. Thus nonlinear systems were largely ignored for hundreds of years, in favor of more tractable phenomena. Unfortunately, this neglect led to a general ignorance of the phenomenon of nonlinearity, and a widespread assumption that all natural systems are linear. In fact, quite the opposite is true: the vast majority of natural systems exhibit nonlinear behavior. This is bad news from the point of view of prediction, but good news from the point of view of life; for it is nonlinearity that makes possible the genesis of novelty and form in nature—and in particular, the genesis of living form. In the next section, we will see how the concepts introduced in this section allow us to resolve the paradox of development—to understand how embryogenesis can reliably generate species-typical bodies without the forms of these bodies being predetermined by a design or program.

The concept of stability gives us insight into how self-organizing systems can generate and maintain complex, orderly patterns of behavior. We have seen how a nonlinear system such as the oscillating chemical reaction can generate stable, periodic behavior (*a limit cycle*). Since all the system's trajectories are attracted to the limit cycle, the system will reliably generate this behavior *regardless of its initial conditions*, and also *maintain* it in the face of fluctuations and perturbations. Notice that this stable behavior is not "pre-programmed" anywhere in the system. There is no central agent measuring the difference between the system's present state and some "target" state. Rather, the stable behavior arises "naturally" or "spontaneously" out of the interactions of the system's parts, and is maintained in the same way. This concept of stability will be
2.4. The genesis of living form

key to understanding the robustness of development: its capacity to remain the same despite changes in its environment.

The concept of multistability in dynamic systems helps us to understand how a self-organizing system can exhibit a variety of stable, orderly behaviors, and how it can switch between these different behaviors in response to environmental perturbations—again, without there being any central control unit switching these behaviors on and off. This concept will be key to understanding the flexibility of development: its capacity to change and adapt in response to changes in its environment.

Finally, the concept of symmetry-breaking bifurcations at critical points will help us to understand how the embryo develops from a single, relatively symmetrical cell into a highly asymmetrical, multi-cellular organism.

4. The genesis of living form

The problem of development is to understand how a complex, multicellular organism can develop—reliably and robustly—from a single cell. In Chapter One, I criticized mechanistic approaches to this problem, which assume that the form of the adult organism must already be determined prior to development by a genetic design or program. In this section, I will offer an alternative view of development which explains how the order and determinacy of the adult organism can arise dynamically through the process of development, rather than being fully given in advance. Applying the concepts developed earlier in this chapter, I will argue that the process of development must be understood as a cascade of symmetry-breaking bifurcations that transform the
2.4. *The genesis of living form*

relatively symmetrical zygote into the much less symmetrical form of the mature organism.\(^7^8\)

Every multicellular organism begins its life as a single cell, and grows by repeated cellular reproduction. Individual cells reproduce by *division*: the "mother" cell splits into two "daughter" cells, each of which inherits a copy of its mother's DNA and half of the mother cell's cytoplasmic contents. Daughter cells are usually identical to their mother, and to each other. Thus we would expect the repeated cell division of a spheroid cell to produce a roughly spheroid mass of identical cells. The mystery of embryogenesis is that it produces, instead, a highly *asymmetrical* body of cells which are highly *differentiated* from one another. As they divide and increase in number, the cells of the growing body become differentiated into various different cell types which differ greatly in size, shape and function. The "totipotent" zygote (which can generate any cell type) gives rise by cell division to "pluripotent" cells (capable of generating a more limited range of cells), and these in turn produce cells of progressively more limited potency, culminating in the various specialized cells of the mature body (which are generally "unipotent", capable only of reproducing cells of the same type). As the growing body becomes larger and more internally differentiated, it also takes on a series of different macroscopic shapes which eventually culminates in the organism's mature form. Development can thus be divided into three interactive processes, taking place at different spatiotemporal scales: *division, differentiation*, and *morphogenesis*.

\(^7^7\) The empirical details of this account will be drawn largely from Forgacs & Newman 2005.
2.4. The genesis of living form

A. Division

Our study of embryogenesis begins with a single cell: the fertilized ovum or zygote. A eukaryotic cell consists of a semi-permeable lipid membrane surrounding a partly organized, partly liquid medium (the cytoplasm), in which various organelles (such as the mitochondria and the nucleus containing the cell's DNA) are embedded. The membrane separates the interior of the cell from the surrounding environment, allowing only certain molecules to enter and leave. Thus the cell can be thought of as a bounded region, within which various chemical reactions are taking place. Certain reagents can cross this boundary; which ones, and at what rates, depends on both the cell's internal state and the state of its immediate environment. We can think of the biochemical contents of the cell as a self-regulating dynamic system made up of various chemical reagents whose concentrations are changing over time at different rates (Forgacs & Newman 2005, Ch. 3).

The first major step in development is cleavage, in which the unicellular zygote divides repeatedly to form a multicellular aggregate called the blastula (Forgacs & Newman 2005, Ch. 2). All cells reproduce by a process of division in which the cell membrane pinches closed down the middle, dividing the cytoplasmic contents of the "mother" cell into two separate "daughter" cells. This division is usually preceded by a replication of the cell's chromosomes, such that each daughter cell not only inherits half

---

78 "Viewed from the perspective of physics, early development, during which the organism acquires its final shape, is a series of symmetry-breaking events starting from a highly symmetrical spheroidal egg and arriving at a body with a much lower degree of symmetry" (Forgacs & Newman 2005, p. 180).
79 Eukaryotes are cells with nuclei, as opposed to prokaryotes such as the unicellular bacteria, which have no membrane-bound organelles.
80 A more complicated model could also take into account the spatial aspect of the cell's biochemical state, since chemical concentrations may not be uniform throughout the cell. However, the simpler model is often sufficient.
2.4. The genesis of living form

of the mother cell's cytoplasmic contents, but also a complete copy of the mother's DNA. The timing of DNA replication and cell division is regulated by the dynamics of the cell's biochemical contents (i.e. the changing concentrations of chemical reagents within the cell). Ordinarily, the mother cell doubles in mass prior to division, so that each daughter cell is roughly the same size as its mother. In cleavage, however the zygote (whose diameter is roughly ten times that of most other cells in the human body) divides repeatedly without increasing in mass. As a result, the cells of the developing blastula become smaller in size as they increase in number.

Experimental evidence indicates that the rhythm of cell division during cleavage is regulated by a biochemical oscillator (Forgacs & Newman 2005, Ch. 3): a chemical reaction that spontaneously generates a stable limit cycle, in which the concentrations of certain reagents vary periodically (see Section 2.3 above). Experiments with frog embryos have shown that the timing of cell division during cleavage is regulated by periodic oscillations in the concentration of a regulatory factor called cyclin B. Fertilization of the ovum triggers the beginning of cleavage by generating a Hopf bifurcation (Novak and Tyson 1993): the concentration of cyclin B within the ovum goes from a stable steady state to a stable limit cycle—a pattern of regular oscillation. This

---

81 "Molecular clocks that regulate entry into DNA synthesis and mitosis are based on temporal oscillations of the concentrations of members of the cyclin family of proteins. Such oscillations are the physical consequences of positive and negative feedback effects in dynamical systems, such as that represented by the cell's biochemistry" (Forgacs & Newman 2005, p. 27).
82 Novak and Tyson (1993) have modeled this process as a nonlinear dynamic system whose state depends on the changing concentrations of nine regulatory proteins that together govern the concentration of cyclin B.
2.4. The genesis of living form

limit cycle is then passed on from mother cell to daughter cells, so that the cells of the growing embryo continue to divide at the same rate.\(^{83}\)

When a unicellular organism divides, the two daughter cells simply split apart and go their separate ways. When the zygote divides in two, however, its daughter cells adhere to each other, forming a new, multi-cellular body. This division separates the cytoplasmic contents of the zygote into two distinct regions, bounded by a double-layer of semi-permeable membrane. This boundary will only allow some chemicals to cross from one cell to the other; and the rate of diffusion across the membrane will be significantly slower than that within either cell. Thus each cell now has its own distinct biochemical state, and can follow its own unique destiny. It is this division of the unicellular zygote into two, (and then four, and so on), that makes possible the growing body's differentiation into a variety of different cell types, tissues, and organs. Thus the initial symmetry-breaking bifurcation that sets the process of cleavage in motion (breaking the time-symmetry of the dormant ovum) makes possible further losses of symmetry through the differentiation of one cell from another.

Of course, the division of the zygote does not separate the two daughter cells completely: certain biochemicals can still travel slowly across the membrane barrier from one cell to another. Thus the biochemical states of the two cells can continue to influence each other. This mutual influence turns out to be every bit as crucial to development as the two cell's separation. Consider what happens to the growing embryo if, immediately

\(^{83}\) Interestingly, the cyclin B oscillation will arise even in the absence of a nucleus, which means that the rhythm of cell division during cleavage (at least in its early stages) is not regulated by the frog's DNA. Daughter cells inherit the cycle of periodic division not through their genes, but by inheriting half of their mother's cytoplasm, which contains the reagents that generate the cyclin B oscillation. However, we will soon see how models such as this one can also be used to explain the role of DNA in development.
2.4. *The genesis of living form*

after the zygote's initial division, the two embryonic cells are carefully pulled apart. In normal development, each of these cells would ordinarily give rise through cellular reproduction to roughly half of the mature organism. One might expect, then, that if the embryo is divided in two at this stage of development, each cell will develop into one half of a normal adult body. Instead, experiment shows that this procedure yields two small but fully formed organisms. In other words, both cells of the two-celled embryo are *totipotent*: each is capable of developing on its own into a complete multi-cellular organism. And yet, if these two cells are allowed to adhere to each other they will not develop into two small organisms but into a single larger one. Evidently, there is something about the relation between the two cells that shapes how they each develop; the cells of the developing embryo *mutually regulate* one another, and it is this mutual regulation which explains how the process of development robustly produces species-typical outcomes without these outcomes being fully given in advance like the design of an artifact.

If each cell is a self-regulating dynamic system in its own right, then the division of one cell into two mutually regulating cells generates a second-order dynamic system: a self-regulating system of self-regulating systems. The number of variables involved in such a system is rather daunting; however, it is possible to simplify matters somewhat. We can distinguish intra-cell dynamics from inter-cell dynamics by the different *scales* on which they take place. Since chemicals can diffuse more quickly within a cell than across cell membranes, chemical reactions happen more quickly within cells than between them. At small time-scales, then, we can treat the biochemical contents of a single cell as an isolated system, ignoring its interactions with neighboring cells. Once we
2.4. The genesis of living form

have a grasp on these local dynamics, we can ask what happens at progressively larger space- and time-scale when these local dynamics interact with the local dynamics of other cells.

B. Differentiation and the role of DNA in development

We have seen that an individual cell can be regarded as a nonlinear dynamic system, thanks to the complex interactions between the many biochemical reagents enclosed within the cell's lipid membrane. We have also seen that cells can regulate one another's biochemical dynamics by exchanging chemicals across their membrane boundaries, and that this mutual regulation plays a key role in development. In particular, these inter-cell dynamics within the growing embryo are what regulates the differentiation of embryonic cells into specialized cell types which will eventually form the various tissues and organs of the mature body. In order to understand this process, we need to understand the role that DNA plays in the biochemical dynamics of each cell.

We can distinguish one type of cell from another by the different set of proteins each is capable of producing (Forgacs & Newman, Ch. 3). This is determined by the physical state of the chromatin in a cell's nucleus—a complex of DNA and protein.

Recall from Chapter One that certain genes (i.e. stretches of an organism's DNA) specify particular proteins (as sequences of amino acids). The cell can generate the protein specified by such a sequence through a process called "transcription." However, not all of the genes that specify proteins are "open" or ready to be transcribed; in any given cell, only some stretches of DNA will be transcription-competent (which requires that they be packaged by chromatin-proteins in the right way). A cell's type is determined by the open
2.4. The genesis of living form

or closed states of a relatively small number of genes. (In humans, these genes number more than a hundred and possibly fewer than a thousand, and they specify around 250 different cell types).

The open or closed state of a given gene, as well as the rate of its transcription, are regulated by proteins called transcription factors, which work by binding to specific sequences of DNA. These transcription factors may be activators, which increase the rate of a gene's transcription, or repressors, which decrease it (Forgacs & Newman 2005, Ch. 3). Transcription factors are themselves proteins specified by genes and produced by transcription; thus there can be circuits or networks of transcription factors that mutually regulate one another's expression. The best way to gain insight into such autoregulatory networks is, once again, to use the tools of dynamic systems theory. Keller (1995) modeled six different autoregulatory transcription networks as nonlinear dynamic systems. Each system consisted of various different transcription factors whose concentrations vary over time and depend on one another in complex ways. These systems were found to be multistable: each transcription network had multiple attracting fixed points, corresponding to multiple stable biochemical states. These stable states were sufficiently different from one another to offer a possible explanation for differentiation: if each biochemical state corresponded to a different cell type, then differentiation would take place through a cell's shifting from one stable state into another (Fig. 21). This shift would require a fluctuation in the system's biochemical state (i.e. in the concentration of various transcription factors within the cell) large enough to push the system out of one
2.4. The genesis of living form

basin of attraction and into another. This fluctuation could be caused by changes in the cell's chemical environment.\textsuperscript{84}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure21.jpg}
\caption{Schematic representation of differentiation in the Keller model: a fluctuation in the concentration of transcription factor X shifts the system out of its initial stable state into a new basin of attraction, corresponding to a different cell type.}
\end{figure}

In the developing embryo, a cell's immediate environment usually consists of other, similar cells. The embryo thus consists of \textit{multiple} autoregulatory transcription networks, which interact and mutually constrain one another dynamically over time. In order to understand differentiation during development, then, we need a model of a \textit{second-order autoregulatory network}—an autoregulatory network of autoregulatory networks. Kaneko and Yomo (1999) have constructed such a model, representing the changing biochemical states of multiple cells interacting dynamically with one another by

\textsuperscript{84} It could also be brought about in other ways, such as asymmetrical cell division, in which transcription
2.4. The genesis of living form

exchanging chemicals. This model starts out with a small population of identical cells that can replicate themselves by division. These model cells were found to spontaneously develop stable, periodic behaviors akin to the biochemical oscillations which regulate the reproductive cycles of real cells. (See 2.4.B above.) As the population of model cells grew by division, all the cells initially exhibited the same biochemical behavior (corresponding to a single, shared cell type). However, when the number of cells exceeded a certain critical threshold, this uniform state became unstable and the multicellular system underwent a symmetry-breaking bifurcation: the population of cells split into distinct clusters of cells, each cluster exhibiting a different stable biochemical state. These new clusters remained stable as they grew by reproduction, with daughter cells inheriting the "cell type" (i.e. the biochemical state) of their parent. As they grew, some clusters underwent new bifurcations, dividing into further clusters with new stable cell states. As in a real body, different types of cell in the model arose through different developmental lineages, which could all be traced back to a single original cell type. (These qualitative outcomes were robust under both variations in initial conditions and fluctuations in parameters over time, suggesting that this may be a realistic model of differentiation and development in real multi-cellular bodies.)

This model demonstrates how a group of initially identical, self-reproducing systems could—simply by interacting with one another over time—generate a diverse population of systems with distinct, stable behaviors.85 It also supports and deepens the claims I made above in Section 1.4 about the role of DNA in development. In the Kaneko-Yomo model, DNA is neither a blueprint nor a program for development.

---

85 factors are unequally distributed between the two daughter cells (Forgacs & Newman 2005, p. 69).
2.4. The genesis of living form

Rather, DNA corresponds in this model to the parameters of the dynamic subsystems (the individual cells) whose interactions generate the higher-order dynamic system of the multicellular body. As we saw in Section 2.3, the parameters of a nonlinear dynamic system are not a blueprint or program for that system's behavior. To say that a system is nonlinear is precisely to say that there is no part-for-part correspondence between changes in the system's parameters and changes in its behavior. In a nonlinear system such as the Kaneko-Yomo model, a small change in the parameters of the original cells may produce a disproportionately large change in later outcomes—or it may produce no change at all. This is an accurate reflection of the relation between DNA and developmental outcomes: a large mutation may have negligible effects, while a small mutation may have enormous effects. Evolution, then, does not occur through tiny changes in a blueprint or program for development, but through small variations in the parameters of development as a dynamic, self-regulating process.  

Kaneko and Yomo (1999) have termed this process "isologous diversification." Notice that in Keller's model, the role the environment plays in differentiation is to push the cell out of one stable biochemical state and into another; whereas in the Kaneko-Yomo model, the cell's environment changes which of its biochemical states are stable. In dynamic systems terms, the environment in Keller's model is only a source of fluctuations, whereas in the Kaneko-Yomo model the environment can also alter the stability of the system, resulting in symmetry-breaking bifurcations (qualitative shifts) in the system's behavior. As a result, the biochemical states that are stable for cells in the multicellular system are not necessarily stable for the same cells taken in isolation. In other words, the behavior of the multicellular system cannot be predicted on the basis of the behavior of isolated cells; it is a non-additive whole. Here we see how the concepts of dynamic systems theory can give us insight into the relation between parts and wholes in nature. Two dynamic systems interacting with each other can destabilize each other, settling into a new state which would not be stable for either system in isolation. The dynamics of the individual systems are parameters with respect to the larger system; a change to the parameters of either subsystem, if it alters that system's dynamics, will alter the parameters of the larger system, which may alter its dynamics.
2.4. The genesis of living form

C. Morphogenesis

The Kaneko-Yomo model assumes that cells interact only with their immediate neighbors, exchanging chemicals by osmosis. (This is known as "juxtacrine signaling"). However, cells can also interact with one another over larger scales ("paracrine signaling") through chemicals that spread by diffusion (Forgacs & Newman 2005, Ch. 7). Such long-range interactions play an important role in both differentiation and morphogenesis. To conclude our discussion of embryogenesis, let us look at the role of these interactions in an important morphogenetic process: the symmetry-breaking formation of the embryo's body axes. Recall that the embryo starts out as a single cell that is more or less spherically symmetrical. Over the course of development, the embryo must somehow lose this initial symmetry in order to develop a front and a back, a top and a bottom, a left and a right.

It has long been known that certain groups of cells, known as organizers, play a key role in this process (Forgacs & Newman 2005, Ch. 7). For example, the Spemann-Mangold organizer in amphibians is a group of cells that induces the formation of the notochord, (the precursor of the spinal column). Other classes of vertebrates have similar organizers. Organizers work by secreting morphogens: chemicals that travel across the embryo by diffusion and alter the biochemistry of other cells. The process of diffusion generates a chemical gradient in which morphogen concentration decreases with distance from the organizer. If the effects of the morphogen on other cells depend on the morphogen concentration, as is often the case, then a single morphogen gradient can generate a variety of location-specific effects. This is one way in which symmetry-breaking patterns are formed during development.
2.4. The genesis of living form

Transplant experiments in frogs and other model organisms have revealed several interesting features of the "organizer" phenomenon. If the Spemann-Mangold organizer is transplanted into another embryo at the right stage of development, the host embryo will develop two different body axes, resulting in conjoined twins. Even more surprising, however, is what happens in the embryo from which the organizer cells were removed: cells adjacent to the excised organizer differentiate into organizer cells and assume its role (Forgacs & Newman 2005, Ch. 7). Thus the formation of the organizer, and hence of the body axes, is a robust, self-organizing feature of the developing embryo—a stable attractor of this dynamic system.

How does organizer-formation work? The fact that a new organizer forms when the old one is removed suggests the existence in this region of an activator: a morphogen that causes cells to differentiate into organizer cells. However, the fact that this does not happen if an organizer is already present suggests that the cells of the organizer secrete another morphogen which inhibits organizer-formation in nearby cells. Meinhardt (2001) has constructed a reaction-diffusion model of organizer-formation as a nonlinear dynamic system. Recall the Keller model above, in which different proteins in a single cell regulated one another's transcription rates, forming an autoregulatory system; and the Kaneko-Yomo model, in which multiple cells of this sort could interact with one another over short distances. Meinhardt's model adds another layer of complexity by allowing these cells to interact over large distances through diffusion, with different morphogens diffusing at different rates. In Meinhardt's model, a slow-diffusing activator is positively

---

87 The pattern-forming possibilities of such "reaction-diffusion" systems was first observed by Alan Turing (1952)—the same Turing who originated the idea of the modern computer. (See Ball 2009, 154-158; Stewart & Golubitsky 1992, Ch. 7.)
2.4. The genesis of living form

self-regulating, i.e. it increases its own transcription rate. But this same morphogen also increases the transcription rate of a fast-diffusing inhibitor, which negatively regulates transcription of the activator. The result is a self-enhancing feedback loop acting over a short range, which competes with an inhibitory reaction acting over a longer range. Starting from a uniform distribution of the two morphogens, this system will undergo a symmetry-breaking bifurcation in which the activator becomes concentrated in a small "hot spot" surrounded by a cloud of inhibitor. By adding additional morphogens to his model, Meinhardt showed that the formation of one organizer could induce the formation of other organizers at different locations. This cascade of symmetry-breaking bifurcations offers a powerful and realistic model of morphogenesis in the developing embryo.

D. Summary

The process of development which appears so puzzling—even paradoxical—from a mechanistic standpoint becomes intelligible when we regard the developing embryo as a self-organizing dynamic system undergoing a cascade of symmetry-breaking bifurcations. The concepts introduced in this chapter allow us to understand how it is possible for the complex form of the organism to develop reliably without this form being determined in advance by a design or plan. Development reliably generates viable organisms not because it is guided by a prospective view of this outcome, but because each stage of development—each symmetry-breaking bifurcation—is a stable, robust behavior of the stage that precedes it.

Fertilization breaks the time-symmetry of the dormant ovum by pushing its biochemical state out of its stable fixed point and into a periodic oscillation. This
biochemical oscillation generates a cycle of periodic cell divisions that transforms the single-celled zygote into an aggregate of many cells (the blastula). The division of the zygote into many smaller cells makes possible a further series of symmetry-breaking bifurcations, in which different cells assume different stable biochemical states.

The biochemical dynamics of each cell involve complex autoregulatory networks of gene transcription factors—proteins which both regulate and are produced by differential rates of DNA transcription. Some of these factors can pass through cell membranes, allowing the autoregulatory networks within different cells to regulate one another. Thus the growing body forms a self-regulating dynamic system of self-regulating dynamic subsystems, in which individual cells both generate and are constrained by the multicellular whole. If the division of the zygote into many cells is what makes differentiation possible, then it is the ongoing interactions between these cells that actually cause differentiation to occur.

Cells can exchange chemicals both locally, over small scales, and globally over larger scales. Global interactions are responsible for macroscopic symmetry-breaking bifurcations in development, such as the formation of body axes in the previously spheroid embryo. These bifurcations take place through the appearance of "organizers"—self-organizing chemical "hot-spots" that regulate the differentiation of surrounding cells and the orientation of large-scale structures such as the spinal cord.

5. Conclusion

The living body poses a puzzle for us. How can something so orderly and so complex—a work of engineering more impressive than any human accomplishment—
2.5. Conclusion

come to be without planning, without supervision, without intelligence? This perplexes us because we try to understand all creation, all coming to be, on the model of human manufacturing. In this chapter, I have tried to show that nature has its own ways of generating order and form, which are quite different from those of human artifice.

In order to understand the genesis of form in nature, we have to replace our mechanistic concepts with dynamic ones. Mechanistic thinking conceives of wholes as constructed out of parts which are given in advance of construction. This static conception of wholes must be replaced by a dynamic one, in which larger systems arise out of the ongoing, reciprocal interactions of smaller ones.

The concept of fixity must be replaced by the concept of stability. The products of human artifice are fixed: they remain the same so long as nothing changes them. Fixity is a static concept; stability is a dynamic concept. That which is stable does not merely persist by inertia in a given state, but actively resists change—not rigidly, the way a steel rod resists being bent, but elastically, the way a spring returns to equilibrium, or a marble returns to the bottom of a bowl. The stable system can be disturbed, but it takes work to shift the system away from the stable state; left to itself, it will soon damp out perturbations.

Closely related to the concept of stability is the concept of robustness, which replaces the notion of outcomes guaranteed in advance. The robust outcome is insensitive to dynamic fluctuations and variations in initial conditions—not because these make no difference to the system, but because the system can produce the same outcome in a variety of different ways. A system with a unique stable state can arrive at that state by an infinite number of different trajectories—but they all end in the same place. A robust...
outcome is not guaranteed: nothing determines it in advance, and no one is watching to make sure the system gets there. But a robust outcome happens more often than not, because there are more ways of getting there than there are ways of avoiding it.

To be robust, a system must be flexible: it must adapt to and accommodate change. The mechanistic version of this is the program, in which different preordained responses are set to be triggered by different anticipated situations. But true flexibility is dynamic; it cannot be arrived at by multiplying static responses. A dynamic system can be flexible in several ways. One is multistability, in which a large fluctuation in the state of the system can push it out of one stable state and into another. Another kind of flexibility can be seen in the phenomenon of bifurcations, in which a change in the system's environment provokes a qualitative change in the system, resulting in new stable behaviors.

The human experience of order is that it requires work on our part, both to create it and to maintain it. In my garden, for example, it takes work to get the flowers and herbs I like growing where I want them to grow, to protect them from squirrels and insects, and to keep the weeds from taking over the soil. If I left the garden to itself, it would soon be an overgrown mess. Things will not simply order themselves; and if the order we've made is left to itself, it will soon revert to disorder. This is certainly true of the kinds of order that human beings seek to establish; but it is not true of the kind of order that we find in nature—of which our own bodies are one example. Nature's garden has its own

88 "A machine is capable only of operations for which it has been constructed; the idea of a machine which would be capable of responding to an indefinite variety of stimuli is a contradictory one, since automation is obtained only by submitting the initiation of work to certain chosen conditions" (SB 87/96).
2.5. Conclusion

order, different from that which we try to impose on it; and this natural order is prior to
the order we seek to create. In nature, order is not a positive quantity which must be
created and maintained by special efforts, and which therefore requires special
explanation when it is discovered. On the contrary, order in nature arises by subtraction,
by the breakdown of nature's initial symmetry. Natural order is not positive, but
negative—a lack. It arises through the loss of disorder. Thus nature does not require the
constant action of God or natural forces to prevent matter from reverting to a disordered
state. Order and form arise in nature because, contrary to our expectations, it takes less
work to maintain them than it would to maintain a more disordered state. Thus form can
be more stable than disorder because it is a lower-energy state.\textsuperscript{89}

In the next chapter, we will apply this dynamic perspective to the study of
behavior, and in particular to learning: the genesis of new behavior. We will see that
learning is a process of development, continuous with the anatomical development that
we have studied in this chapter. And we will see, further, that anatomical development
can be understood as a slow behavior.

\textsuperscript{89} In the global long-term, this is not the case. The Second Law of Thermodynamics tells us that in a closed
system—such as the universe itself—the lowest energy state, in which the system will eventually settle, is
one of maximum disorder and uniformity: the so-called "Heat-Death of the universe". However, in local,
thermodynamically open systems driven by external energy inputs—such as our planet, or an individual
living body—order and form can arise as stable, lower-energy states.
CHAPTER THREE: How do living bodies make sense?
1. Introduction

In Chapter One, I argued that the living body is not a machine. I distinguished growth from manufacturing, and traced the conflation of these two processes to the mechanistic research program in Modern science. By studying the facts of development, we saw that growth cannot be understood by analogy to manufacturing as the imposition of a pre-determined form on passive materials. However, this left us with a problem which I called the "paradox of development": how do living bodies consistently develop from a single cell into a highly complex, species-typical adult form without this form being specified anywhere in advance? In Chapter Two, I attempted to resolve this paradox by turning to dynamic systems theory, in search of a non-synthetic account of how form arises in nature. There we found a new definition of form as asymmetry, which allowed us to re-frame the problem of morphogenesis as the question of how differences arise in nature: why are things different in one place than another, and why are they different now than they used to be? The mechanistic research program assumes that asymmetrical causes have equally asymmetrical effects. Thus differences in nature can only be explained in terms of pre-existing differences, and complex forms must have equally complex causes. This amounts to a demand that natural forms be determined in advance by their causes in the same way that artificial forms are determined in advance by their human designers. However, the study of nonlinear dynamic systems shows that effects can be more asymmetrical than their causes. Quantitative variations in the parameters of a nonlinear dynamic system can give rise to symmetry-breaking bifurcations in which the old configuration of the system is destabilized and a new stable configuration arises which is less symmetrical than its predecessor. The concepts of
symmetry-breaking and stability explain how order and organization can arise in nature without being planned in advance or directed by an external intelligence. Human beings create complex wholes by assembling them out of pre-existing parts according to a pre-established design or plan. In nature, on the other hand, complex wholes organize themselves by division rather than addition, articulating themselves into parts by progressively losing their initial symmetries.

At the end of Chapter Two, we looked at how the growing embryo becomes increasingly asymmetrical over time through the self-organizing dynamics of nested "autoregulatory networks" at multiple spatiotemporal scales. Within each cell, the genome both regulates and is regulated by its own transcription factors; and these intracellular dynamics also regulate and are regulated by those of other cells, both local and distant. Each stage of development—each symmetry-breaking bifurcation—is a robust outcome of the preceding stage's dynamics. Thus the endogenous dynamics of the growing embryo robustly generate a species-typical adult body without being guided by any design or plan that would specify the form of this body in advance. In this chapter, we will extend our account of development and morphogenesis from anatomy to behavior. If the living body's anatomy is not fixed in advance by its genes, then neither is its behavior. Thus learning, like growth, must be understood as the dynamic emergence of increasingly asymmetrical forms. The form we sought to explain in the previous chapters was a form in space: the anatomical form of the adult organism. What we learned, however, was that this form is only one moment in an ongoing process that is never finished: the growth and development of the living body. Thus the spatial form of the adult's anatomy is only a cross-section, so to speak, of a larger form, which is spread
3.1. Introduction

out in time as well as space.\textsuperscript{90} We saw in Chapter Two how this growth arises from the repeated division of one cell into two, and the subsequent differentiation of these once-identical cells into different types, tissues, and organs. Thus growth takes place through the articulation of the body into parts that can oppose one other, pushing one other in different developmental directions. In animal bodies such as ours, the movement of growth eventually generates a new kind of articulation which makes possible a new kind of movement: the articulation of limbs into joints. The joint allows the body to oppose itself in a new way, using muscle tension to move one part of the body by pushing off of another.\textsuperscript{91} This self-opposition allows the body to push off of its surroundings as well, moving across immobile surfaces and manipulating movable objects. It is these gross bodily movements that psychologists study under the rubric of "behavior."

We saw in Chapter One that what calls for explanation in the anatomical form of the living body is the internal coordination of its many parts into a functional whole, and the adaptation of this whole to its environment. Similarly, the challenge for any theory of behavior is to explain the internal coordination of the living body's movements and the adaptation of these movements to the organism's immediate situation (SB 10/8). Living bodies do not move in a chaotic or haphazard way; their movements appear organized and purposeful. They seem to respond in intelligent and goal-directed ways to the various and changing features of their environments: avoiding predators and other dangers, and seeking out food, shelter, and mates. The goal of a theory of behavior is to explain how the complex coordination of behavior comes about: how the various parts of the living

\textsuperscript{90} See SB 38/38: "]natomy should be considered as a stage \textit{[coupe]} in the development of physiology."

\textsuperscript{91} See Aristotle's \textit{Movement of Animals}, Ch. 8, and \textit{Progression of Animals}, Ch. 3. For helpful discussion of these texts, see Sanday 2008 and Beith 2011.
body's movements come together in a harmonious way to solve the problems presented by its environment. As in the previous chapters, I will argue that this question must be studied developmentally. Just as the anatomy of the adult body is not present from conception, but has to develop over time out of a much simpler form, so too do the coordinated and adaptive movements of the mature organism have to develop over time from much simpler behaviors.

As in Chapter One, I will begin by considering the best attempts to explain the organization of behavior mechanistically (3.1): Reflex theory, Cognitivism, and Connectionism. We will see that these theories face the same conceptual difficulties as the mechanistic theories of anatomical development that we studied in Chapter One. They distinguish absolutely the parts to be coordinated, determining them completely in advance of their interactions with one another. This then makes it necessary to find some external agent that would be responsible for coordinating these mutually indifferent and external parts—an agent that would also have to be given in advance of the coordination it effects. By framing the problem in this way, mechanistic theories set an impossible task for themselves. No cause given in advance can explain the flexibility and contingency of development, its adaptability and its capacity for bizarre "mistakes." As we will see, the facts of behavior, no less than those of embryology, call for a genuinely developmental account.

Mechanistic theories assume that the form of the body's movements is determined by the form of its nervous system. Thus motor development must be caused by changes in the anatomy of the brain. These anatomical changes, in turn, must be determined in advance by a genetic program. We have already undermined the beginning of this causal
story, by showing that the growth of the body—including the nervous system—is not determined in advance by the body's genes. In section 3.2, we will refute the rest of this story. We will see that the relation between anatomical development and behavior is a circular one: the body's movements are shaped by its anatomy, but also shape this anatomy in return. And we will see that the form of bodily movement is not contained in advance in its nervous system: the moving body is a self-organizing dynamic system whose behavior is more asymmetrical than its causes.

In section 3.3, we will turn to the phenomenon of perception, in order to refute the mechanistic claim that the form of behavior is pre-determined by the situation it is responding to. We will see that perception is not the passive reception of discrete sensations, but the active discovery of environmental differences through movement. The living body learns to move by exploring its surroundings, and the asymmetries it discovers in its environment lead to the development of increasingly asymmetrical movements. Thus learning is not a process of synthesis or association, as mechanistic theories would have it, but rather one of differentiation, in which the body's increasingly coordinated movements allow it to distinguish increasingly subtle differences in its environment. The environment that behavior responds to is therefore not given in advance, but develops in tandem with the organism's own exploratory movements.

2. Mechanistic theories of behavior

A. The reflex theory of behavior

Mechanistic explanations of behavior are as old as the mechanistic conception of the living body, originating in the seventeenth century with Descartes' mechanistic
3.2. Mechanistic theories of behavior

Descartes noted the existence of involuntary reactions in human beings—actions that we do not perform deliberately, and that we are even powerless to prevent. For example, when someone pretends to strike me in the face, I find it impossible not to blink, even if I know that they do not really plan to hit me.\(^9\) Descartes argued that these actions must be produced by automatic mechanisms in the body, without the intervention of mind or will. The body's sense organs must be connected to its muscles in such a way that certain changes or movements in the sense organs bring about, through a chain of mechanical causes and effects, a contraction of certain muscles. Descartes thus held a dualistic theory of human behavior: involuntary behaviors are caused by bodily mechanisms, while voluntary, rational actions are caused by the immaterial mind acting on the body. However, Descartes also held that human beings are unique among animals in having minds and being capable of rational, voluntary action. He thus argued that non-human animals are automatons whose behavior can be explained entirely in terms of bodily mechanisms.

Nineteenth century physiologists took Descartes' mechanical hypothesis even further: they argued that human beings are no different from other animals, and that our behavior too can be explained in entirely mechanistic terms, without appealing to an immaterial mind. Thomas Huxley, the famous champion of Darwinism, argued that consciousness is not the cause of behavior, as Descartes believed, but a mere side-effect of the bodily mechanisms that produce behavior (Huxley 1874). Anatomical research since Descartes' time had clarified how nerves transmit changes from one part of the body to another, and also revealed the existence of two different classes of nerve cells:

\(^9\) Passions of the Soul, Article 13 (CSMK Vol. 1, p. 333; AT XI, pp. 338-9).
"afferent" nerves, which connect the sense organs to the central nervous system (CNS), and "efferent" nerves which connect the CNS to the body's muscles. These discoveries led to a simple mechanistic theory of animal behavior. Each afferent nerve ending at the surface of the body is connected via the CNS to an efferent nerve controlling a single muscle or muscle group. When things in the organism's environment stimulate nerve endings in the body's sense organs, these stimuli trigger contractions in the body's muscles. These contractions together produce the gross bodily movements we call "behavior." The connection from nerve ending to CNS to muscle became known as the "reflex arc", probably because the CNS "reflects" the stimuli it receives into muscle contractions, so that the animal's behavior is a straightforward mechanical reflection of its situation.\footnote{For a history of the reflex theory's development, see (Phillips 1971).}

![Reflex Arc and Behavior Diagram]

**Fig. 22. Schematic of the reflex theory of behavior**

The underlying logic of this theory should be familiar to us from the mechanistic theories of development we discussed in Ch. 1. Recall how geneticists attempted to explain the coordination of the living body's anatomy by dividing the body into a
3.2. Mechanistic theories of behavior

collection of discrete traits, and positing the existence of a single, discrete gene as the cause of each trait. Similarly, the reflex theory attempts to explain the coordination between behavior and environment by analyzing the body's movements into discrete muscle contractions and positing the existence of a single, discrete stimulus as the cause of each muscular response. Just as genetics attempted to show how the form of the living body is determined in advance by finding a one-to-one correlation between gene and phenotypical trait, the reflex theory aims to determine behavior in advance by finding a one-to-one correlation between stimulus and response.

The relation between these two theories is more than simply analogical: they share an underlying conception of the living body as a machine, a collection of discrete parts and mechanisms assembled according to a clever plan. The reflex theory explains the adaptation of behavior to its environment—the fact that stimuli trigger appropriate, useful responses in the organism—by appealing to a fixed anatomy optimized by natural selection. Like a well-designed machine, the living body is provided with causal mechanisms that detect relevant features of its environment and produce responses which are advantageous to the organism. In human-made machines, this design is crafted by an intelligent being who anticipates the situations to which the machine will need to respond, and builds mechanisms into the device that will produce appropriate responses. In the living body, according to Huxley and his successors, this human foresight is replaced by the blind trial and error of evolution to achieve the same result: a blueprint for a machine whose automatic responses are well-suited to its environment.

Because the reflex theory understands the living body as a machine—a collection of mutually external parts—it can only conceive of behavior as a series of mutually
external events (SB Ch. 1). The internal coordination of behavior—the way that the organism's various movements seem to form an organized and purposeful whole—is only apparent, a unity attributed to behavior by the observer (SB 9/7). In reality, the living body is simply a collection of blind mechanisms which give the illusion of goal-directedness thanks to the cunning of their assembly. All of the unity and foresight exhibited by the living body lies in its design, and none of it in the actual performance of bodily movement. If the parts of the organism's behavior are coordinated with one another, this can only be due to the orderly arrangement of the individual stimuli that caused them. The organism is like a musical keyboard (SB 12-13/10-11); if it produces a harmonious melody, rather than a mere cacaphony, this is not due to any intrinsic harmony in the machine itself, but only to the harmonious order in which the keys were depressed. The form and organization of behavior must be present already in its causes; it is inconceivable that the living body should spontaneously produce order where none was given in advance, that the movements of the living body should be less symmetrical than their antecedents.

Given the common conceptual logic of the reflex theory of behavior and the "blueprint" theory of genetics, it should come as no surprise that the two theories face similar empirical difficulties. We saw in Ch. 1 that the blueprint theory ran into trouble when one-to-one correlations between gene and trait—upon which the whole theory was based—could not be found. It turns out that genes are not trait-makers, but difference-makers (Fox Keller 2010). If all its other genes are held constant, a variation in one gene may produce a corresponding variation in the anatomy of the adult organism. But these effects are not additive: we cannot predict the effects of varying multiple genes by first
3.2. Mechanistic theories of behavior

varying each one "in isolation." Indeed, the very notion of an isolated gene effect turns out to be incoherent. Individual features of the adult organism do not have a single cause; they are products of the total process of development. But this means that the adult organism cannot be divided into a collection of discrete traits, and that genes are not a "blueprint" for development.

A similar surprise awaited the psychologists and physiologists who tried to work out the empirical details of the reflex theory. They discovered that constant correlations between a single stimulus and a single response—the "reflex arc" which was supposed to constitute the basic unit of behavior—were virtually nonexistent in animals (SB Ch. 1; Goldstein 1995 [1939], Ch. 2). The "reflexes" they thought they had discovered turned out to be largely artificial phenomena, produced in the laboratory by holding constant (and then ignoring) the rest of the organism and its situation. Like gene actions, the effects of individual stimuli on the living body are not additive: one cannot predict the organism's response to multiple stimuli by testing each one "in isolation." But this calls into question the whole picture of the living body as a collection of discrete causal mechanisms, and the very possibility of dividing the organism's behavior into discrete responses to discrete stimuli.

B. Cognitivism

We saw in Ch. 1 that the empirical shortcomings of the "blueprint" theory of genetics have led to its replacement by a new model of the genome as a "program" or algorithm for development. In response to the shortcomings of the reflex theory of
behavior, a parallel move has taken place in psychology with the rise of Cognitivism.\textsuperscript{94} According to the Cognitivist, behavior is caused by a program or algorithm executed by the brain, which is a biological version of our own digital computers. This new theory still analyzes the organism's perceptions and behavior into discrete stimuli and individual muscle contractions. However, it offers a more complex picture of what takes place between stimulus and response.\textsuperscript{95} Instead of simply connecting afferent nerve endings to muscle cells, as in the reflex theory, the CNS is now held to construct a \textit{representation} of the external world on the basis of the input it receives from the senses. The CNS then uses this internal model of the body's surroundings to formulate a \textit{motor program} or plan for moving the body through its environment, which it transmits to the body's muscles. This computational theory is supposed to remedy the shortcomings of the reflex theory, whose simple "clockwork" picture of behavior required that the organism's movements correspond, part for part, to the stimuli it receives from the world. The new picture of the CNS as "information processor" purports to explain the complex, nonlinear relation between behavior's "inputs" (perception) and its "outputs" (movement). Since perception and movement are mediated by an algorithm that represents the world and formulates plans for motor action, the organism's movements do not need to resemble its sensations. Individual stimuli are no longer taken to be the causes of individual muscle contractions;

\textsuperscript{94} The reflex theory was already coming under fire by the end of the 19\textsuperscript{th} C. (see e.g. Dewey 1896), but the rise of Cognitivism had to wait for the invention of the digital computer in the middle of the 20\textsuperscript{th} C. In the interim, the reflex theory was propped up by a series of auxiliary hypotheses (which Merleau-Ponty catalogues in Ch. 1 of SB).

\textsuperscript{95} "Thus, although cognitive psychologists continue to insist that their task is to explain what "goes on" between the stimulus and response, they fail to notice that they do not thereby reject stimulus–response psychology but remain trapped within its limitations" (Costall 2004).
3.2. Mechanistic theories of behavior

instead, they contribute to the CNS' internal model of its environment, which then informs its motor plans.

![Diagram](image.jpg)

**Fig. 23. Schematic of the Cognitivist theory of behavior**

Notice that for Cognitivism, as for the reflex theory, things in the world are the cause of perception, and perception is the cause of behavior. Once again, it is inconceivable that behavior should be less symmetrical than its causes. Whatever order and organization is present in behavior must have been present already, either in the environmental causes of behavior or in the design of its anatomy—in this case, the design of the algorithm which is embodied in the structure of the CNS. For the reflex theory, the form of things in the world is transposed into the stimuli they produce in the organism, and from there into the organism's behavior. For Cognitivism, this transposition is less simple: the form of things in the external world is indeed transmitted to the CNS through their effects on the body's sense organs, but as a sort of "coded message" which the CNS must decipher. The sensory input received by the CNS contains all the information it needs to construct an accurate model of its environment, but the sense-data must be *processed* in order to extract this information from it. Notice, however, that if the CNS does its job correctly—if it processes the information it receives in the right way—the end result is an internal representation whose form *mirrors* that of the external world.
3.2. Mechanistic theories of behavior

Thus the purpose of all this information processing is not to introduce any new order into the system, but merely to replicate the form already present in the body's surroundings.\textsuperscript{96} For Cognitivism, as for the reflex theory of behavior, the form and organization of behavior must be present already in its causes. It remains inconceivable that the movements of the living body could be less symmetrical than their antecedents, introducing new order into the world which was not already there.

Developmentally, this new theory is no more plausible than its predecessor. The reflex theory attempted to show that the living body's behavior is predetermined by its anatomy, assuming that this anatomy is in turn specified in advance by some sort of genetic blueprint. For Cognitivism, behavior is the product of an algorithm, rather than a clockwork mechanism; but this algorithm must still be embodied in the organism's anatomy, which must again be specified in advance by some sort of plan. If there is no "genetic blueprint" for the adult brain, as I argued in Ch. 1, then the Cognitivist must explain how the algorithm for behavior develops organically over the organism's lifetime, starting from the single-celled zygote. It does no good here to appeal to so-called "learning algorithms" which alter their own code over time, for such algorithms do not develop "from scratch"—their "learning process" begins with an already-complex program written by a human programmer.

\textsuperscript{96} "[S]uch an explanation would permit us to understand how physical things are duplicated in behavior by a representation of themselves. Thus behavior remains defined, according to the simplest of schemata, as an imitation of things." (SB 135/145).
"We began from a world in itself that acted upon our eyes in order to make itself seen by us; we have arrived now at a consciousness or a thought about the world, but the very nature of this world is unchanged. It is still defined by the absolute exteriority of its parts and is merely doubled across its extension by a thought that sustains it" (PP 41/64).
3.2. Mechanistic theories of behavior

Our thinking about the Cognitivist picture will become clearer if we set aside for the moment the fact that the brain's algorithm is supposed to represent the external world and generate actions on the basis of this representation. An algorithm is essentially an input-output device: it specifies a sequence of mechanical operations which, when performed on data of the right kind, produces a useful result. The Cognitivist's algorithm takes sensory stimuli as input and produces muscle contractions as output. But notice that this algorithm is imbedded in a moving, living body, which is itself imbedded in a changing world. The muscle contractions that the algorithm produces as output cause the body to move, exposing its sense organs to new stimuli, which are fed back into the algorithm where they are transformed into new movements. Thus the algorithm's output shapes its subsequent input; what looks like a linear input-output device when considered in isolation turns out, when we put it back into its bodily context, to be part of a circular feedback loop—a self-regulating system that includes the nervous system, the moving body, and the environment it moves through. Moreover, we know that the nervous system is not static, but changes over time as a result of the body's interactions with its surroundings. Thus if the structure of the nervous system instantiates an algorithm, then that algorithm must be constantly changing as a result of its own activity. As we saw in the context of embryogenesis (Ch. 1), programs are more flexible than blueprints, but they still represent an attempt by the scientist to specify the course of development in advance. In the case of growth, the notion of a "genetic program" is an attempt to specify the behavior of each of the growing body's many cells in advance of their interaction with one another. In the case of behavior, the notion of a program in the brain is an attempt to specify the organism's behavior in advance of its interactions with its environment. But
even if the Cognitivist's description of the CNS' activity as algorithmic information-processing were correct, knowing this algorithm would not allow us to predict the organism's behavior. For one thing, the algorithm would not be the "sole author" of behavior, but would determine behavior only in collaboration or "conversation" with the world. For another, the algorithm's dialogue with the world would have to be such as to continuously modify the algorithm itself. And since the body's interactions with the world cannot be predicted in advance, every organism would have its own unique algorithm which could not be known in advance of its development. Thus if we want to understand behavior, it will not suffice to study the CNS in isolation. We will have to study the whole self-organizing system of the nervous system in its body in the world, following the endogenous dynamics of this system to see how it develops over time.

I suggested above that we set aside, temporarily, the Cognitivist's claim that the CNS constructs an internal representation of the external world. Would it alter our conclusions if this claim were brought back into play? In fact, Cognitivism's own mechanistic ontology demands that this claim remain suspended. From a strictly mechanistic standpoint, the Cognitivist's algorithm would be a representation of the external world only for us as observers, not in itself. We saw in Ch. 1 that the concepts of a "genetic blueprint" or a "genetic program", though seemingly mechanistic, are in fact anthropocentric and therefore impermissible within a strictly mechanistic ontology. Something can be a program or a design only for an intelligence that plans ahead and anticipates the product of its designs. In a mechanistic universe there are no purposes or plans, only causes which are indifferent to their effects. Similarly, the CNS is on Cognitivism's own account a machine—a collection of mutually external parts operating
3.2. Mechanistic theories of behavior

according to causal laws. There is no room in this picture for anything to be a representation of anything else. The mechanical actions of the world on the body produce movements in the "circuits" of the brain, which in turn produce movements in the rest of the body. There are no representations here, but only the unbroken chains of cause and effect that make up the clockwork universe, some of which happen to pass through the living body. For a human being designing an algorithm, it might be indispensable to think of certain variables as representing certain states of the external world, and to distinguish software—the algorithm considered as an abstract set of instructions and procedures—from the hardware in which this algorithm will be instantiated. But for the computer itself, there is no hardware/software distinction. There are only transistors clicking open and closed, magnetic bits switching from one state to another in a linear series of mechanical actions. Without form, there can be no representations. If form has no reality in nature—if it is only an anthropomorphic projection, an idea in the mind of the observer—then no part of nature can represent any other part. One atom cannot represent another; one arrangement of atoms may resemble another, but this macroscopic resemblance has no causal reality—it exists only in the mind of the observer who compares them. Yet when it comes time to study this mind scientifically, mechanistic science must insist that the form does not exist there either! Even the scientist's own perceptions of form, however indispensable they may be to the practice of science, must ultimately be explained away.
3.2. Mechanistic theories of behavior

C. Connectionism

Cognitivist models treat the brain as an abstract symbol-processing device, formally equivalent to our own digital computers. Thus they assume that the algorithm instantiated in the brain can be reverse-engineered without knowing anything about the "hardware" it is running on. The computer engineer's distinction between hardware and software, when applied to the brain, provides a new justification for the old dualism between mind and body: if mind is an algorithm, then it can be studied in abstraction from the neuro-anatomy in which it is instantiated. Concerns over the biological implausibility of Cognitivism's disembodied algorithms gave rise in the 1980s to a new "Connectionist" approach, which aimed to devise a more biologically realistic model of the brain (Thompson 2007, pp. 8ff). The result was the "neural network" model, which treats the brain as an ensemble of idealized neurons linked together by numerically-weighted connections which can be strengthened or weakened by the network's own activity. The behavior of such artificial neural networks can be studied by simulating them on a digital computer. Instead of being programmed in advance by a human designer, the simulated neural network is "trained" to carry out a given task by undergoing repeated trials. As a result, the operations of the network are opaque to its creators: they know that it works, without knowing how it works. The goal of the neural network is not to generate a representation of the world, but rather to carry out some concrete task—usually one of pattern recognition or categorization. The most successful neural networks seem to replicate the generality that is one of the strengths of organic behavior: having been trained on an appropriate data-set, the network can respond appropriately to new cases that it has never encountered before.
3.2. Mechanistic theories of behavior

In some ways, the Connectionist model seems to move beyond the mechanistic prejudices of its predecessors. The complex web of connections between neurons generates complex feedback loops within the network, giving rise to non-linear, self-regulating patterns of system-wide activity. Thus, although they are simulated on digital computers, neural networks are not themselves computers (i.e. finite-state automatons), but self-organizing dynamic systems whose architecture is modified over time by their own activity. Instead of being designed in advance to carry out a given task, neural networks grow over time, developing increasingly successful responses through repeated exposure to different situations. Simulated neural networks thus offer a more realistic model of how real nervous systems develop over time in living bodies. However, the Connectionist approach has so far been limited by its retention of a number of Cognitivist assumptions. Connectionists still tend to think of neural networks as "parallel processing" variants of our own digital computers. That is, they treat the neural network as an information-processing device whose function is to turn inputs into outputs, and the "training" process as a way of generating new algorithms which human programmers might not have come up with on their own. Thus Connectionism remains all too disembodied: its models try to capture the salient features of real biological nervous systems, but it makes no attempt to place these models within a realistic body and environment. Instead, neural networks are fed artificial inputs and trained to generate pre-determined outputs. Instead of being allowed to interact with a rich environment in an open-ended way, as organic, embodied nervous systems do, artificial neural networks are restricted to solving predefined problems whose solutions have been determined in advance by human designers.
3.3. Learning to move

Simulated neural networks could be a powerful tool for understanding how living bodies learn to move and perceive. In order to realize this potential, however, we will have to begin placing these networks in bodies—simulated or robotic—and allowing these bodies to explore realistic situations. Only then will artificial neural networks begin to help us understand how intelligent behaviors develop over time through the living body's interaction with its environment. For now, however, let us turn away from artificial models of the brain, and return to the phenomenon of the living body. We have now seen that mechanistic theories, which assume that behavior is determined in advance by its causes, are unable to explain the empirical facts of behavior and its development over time. Let us now study these facts in greater detail, to see how the living body's behavior, like its anatomy, organizes itself into qualitatively novel forms which are nowhere fully given in advance.

3. Learning to move

A. The view from above and the view from below

Biologists have always been struck by the resemblance between organisms of a given species, and especially between parents and their offspring. This focus on intra-species similarity led to the view that there is a single, fixed species-form which all the members of a species share, and which is transmitted from parents to offspring. Darwin's genius was to stop thinking like a taxonomist and to start thinking like a breeder (Darwin 1985 [1859]). The breeder wants to promote certain characteristics in the breeding

---

97 The discussion in this section is greatly indebted to Morris 2004, Ch. 2.
3.3. Learning to move

population and discourage others. Thus what she notices is not the similarities between siblings and conspecifics, but rather the differences between them.98

Psychologists Esther Thelen and Linda Smith (1994) propose a similar conceptual move with respect to development. From afar—what Thelen and Smith call "the view from above"—what strikes us about development is how similar it is from one individual to another within a given species. For example, early studies of human motor development found that all children pass through the same series of developmental stages. In one landmark study, M. B. McGraw described how infants learning to walk go through seven phases of erect locomotion (Fig. 24): reflex stepping (A), static phase (B), transition phase (C), deliberate stepping (D), independent stepping (E), heel-toe progression (F), and finally integrated walking (G) (McGraw 1945; Thelen & Smith 1994, p. 3). However, when we look more closely—what Thelen and Smith call "the view from below"—we see enormous variability in how different infants develop. No two individuals learn to walk in exactly the same way; and each person develops her own unique gait. Neither the path that development takes nor its outcomes are exactly identical from one individual to another.

98 "Not one man in a thousand has accuracy of eye and judgment sufficient to become an eminent breeder" (Darwin (1985 [1859], p. 91); "Nor let it be thought that some great deviation of structure would be necessary to catch the fancier's eye; he perceives extremely small differences" (Darwin (1985 [1859], p. 97).
3.3. Learning to move

Fig. 24. McGraw's seven stages of erect locomotion (McGraw 1945).

The classification of development into *stages* made the flexible, dynamic, and individual process of development appear fixed, rigid, and universal. This in turn suggested that development has a single cause which determines its timetable in advance, and which is indifferent to environmental factors. Psychologists located this "biological" cause of development in the process of brain maturation, whose timing they assumed was genetically pre-determined. Thus the division of development into distinct stages, which
3.3. Learning to move

began as a description of the phenomenon, became reified as its cause by projecting these stages backwards—first into the brain and then into the genes (Thelen & Smith 1994, p. 7). We see here once again the tendency to avoid the question of how new forms are generated, by postulating that they have been fixed in advance. This approach simply avoids all of the interesting questions about development.\(^9\) What propels the infant from one "stage" of motor development to the next, and why does this happen when it does? How do infants with different bodies, growing up in a variety of different environments, reliably develop into independent walkers? And how do some infants fail to do so?

In Chapter Two, we saw how the dynamical concepts of stability and symmetry-breaking can explain how complex systems robustly and reliably generate similar, functional outcomes without these outcomes being somehow determined in advance. Having seen how these concepts can help us to understand anatomical development, let us now apply them to motor development. We will see that the moving body, like the growing body, is a self-organizing dynamical system. Moreover, we will discover that growth and learning are not two separate activities, but a single process of development seen at two different temporal scales.

B. The moving body as a dynamic system

McGraw's description of the phases of erect locomotor development in infants presents us with a puzzle (Thelen & Smith 1994, Ch. 1). When newborns are held erect, they lift and lower their legs in alternating movements which strongly resemble the

\(^9\) Indeed, the belief among psychologists that the details of motor development were a problem for biologists led to a "nearly universal disinterest in motor development studies for nearly 40 years" (Thelen & Smith 1994, p. 7).
3.3. Learning to move

stepping motions of walking. (McGraw calls this "reflex stepping.") This is a surprisingly precocious behavior to find in newborns, who will not walk on their own for another year. The real puzzle, however, is that at about two months of age this behavior disappears—infants no longer perform these stepping movements when held erect. (McGraw calls this the "static phase.") Stepping behavior does not reappear until eight to ten months of age, when infants become capable of bearing their own weight on their feet. This loss of a coordinated and well-articulated movement challenges traditional views of development as a linear, progressive process. If kicking is a "reflex" in newborns, what could account for the loss or suppression of this reflex at two months, and its return as part of functional locomotive behavior a few months later? Careful experimental work by Thelen and colleagues has provided answers to these questions. But these answers demand that we revise traditional assumptions about the nature of movement and development.

Study of the kinematics (patterns of movement in space) and muscle activation patterns of newborn stepping revealed that it is nearly identical to another common newborn behavior: the infant's kicking while lying on its back. Kicking and stepping had always been seen as two distinct behaviors; but this new evidence suggested that they were in fact the same movement performed in two different postures. This discovery was especially significant because these two behaviors—stepping when erect and kicking when supine—have different developmental trajectories. Whereas erect stepping stops at around two months, infants continue to kick throughout their first year when supine, prone, or sitting. Thus what psychologists since McGraw had taken to be a fixed developmental stage (the "static phase") caused by genetically predetermined changes in
the infant's nervous system turned out to be a flexible behavior that varies with the infant's posture. Further experiments confirmed that the stepping behavior was much more labile than McGraw's account suggested: when infants in the "static phase" were held erect with their legs submerged in water, *the stepping behavior magically reappeared*. Conversely, stepping could be *suppressed* in infants under two months by attaching small weights to their legs. The lability of newborn stepping and its sensitivity to context posed a challenge to traditional accounts which posit that motor behavior has a single cause which is located in the central nervous system. Thus Thelen and colleagues were led by their studies of newborn stepping to offer an alternative, dynamic account of motor behavior.

What is striking about newborn stepping is its precocious *coordination*: hip, knee and ankle joints simultaneously trace a smooth and rapid path, first toward and then away from the body; in each leg, kicks follow one another in rhythmic succession, usually alternating with kicks in the other leg (Thelen & Smith 1994, Ch. 4). Traditional accounts assume that such highly coordinated movements must be heavily regulated by the CNS. Thus they hold that the timing of the legs' oscillations is controlled by some "central pattern generator," which produces the alternating flexion and extension of each leg by alternating contractions of each leg's flexor and extensor muscles. However, Thelen and colleagues' careful study of muscle activation patterns during infant kicking revealed that no such precisely timed alternation was taking place (*Ibid.*). Instead, the entire coordinated movement is produced by *a single, simultaneous contraction of both flexors and extensors* at the beginning of each kick. Because the flexor muscles are more powerful than the extensors, the net effect of this simultaneous contraction is to draw the
3.3. Learning to move

infant's feet up toward its body. Once the muscles have relaxed, the springiness of the leg causes it to rebound away from the infant's body, extending the leg. This same springiness finally draws the leg back to its initial resting position. Thus the highly coordinated kicking movement is not produced, as expected, by a highly coordinated pattern of muscle contractions, but by a single uncoordinated burst of activity in all the muscles of the infant's leg. It turns out that the rhythm and coordination of infant kicking and stepping is not due to CNS regulation at all, but to the intrinsic "mechanical" properties of the legs themselves. Each leg acts like a mass on a spring: when energy is pumped into this system by muscle contractions, the forces of gravity and spring tension combine to produce a rhythmic oscillation of the leg, (like that of the simple harmonic oscillators we considered in Ch. 2). Thus the coordination of infant stepping—its cyclic trajectory and rhythm, as well as the timing of each phase of the movement—turns out to be a dynamically self-organizing phenomenon: a form in space and time that is more complex (i.e. less symmetrical) than its causes.

These observations call into question the traditional, mechanistic assumption that the form of behavior—the coordination in space and time of the body's various moving parts—must be caused by a corresponding form in the CNS. Indeed, a bit of reflection will show that as a general principle there can be no one-to-one correlation between

---

100 We are running into a terminological difficulty here: the scientific study of motion has traditionally been known as "mechanics," and the properties of bodies that affect their displacement and deformation when forces are applied to them (e.g. mass, elasticity, viscosity) are accordingly known as "mechanical properties." Thus in describing the parameters of the moving body as a dynamic system, I find myself referring to the "mechanical" properties of the body, even though my aim in describing these properties is to show that the body is not a machine. See Appendix A, where I argue that even a simple "mechanical" system such as a driven harmonic oscillator can exhibit nonlinear dynamics which violate our expectations of how a "machine" should behave.
3.3. Learning to move

patterns of CNS or muscle activity and patterns of bodily movement. Try shaking your hand vigorously at the wrist, or raising your arm above your head and then letting it go limp. Notice that the movements of your hand and arm in these cases are only partly determined by the contraction of your muscles—they also depend on the force of gravity and the mechanical properties (mass, elasticity, etc.) of your limbs and joints. If you were upside-down, or under water, or holding something heavy in your hand, the exact same muscle contractions would produce quite different movements. In general, the consequences of a given muscle contraction always depend on the positions of your limbs with respect to one another and their orientation in the Earth's gravitational field. Thus the moving body is a nonlinear dynamic system whose movements are caused by a whole field of forces—a field that is constantly changing in response to these very movements. Thus the nervous system could not possibly coordinate bodily movement by rigidly determining the position of every joint and limb at every instant (Morris 1999; Morris 2004, Ch. 2). Rather, it sets the body in motion by pumping energy into the system and then modulating the dynamic patterns of coordination that emerge.

It follows that the idea of a "motor program" in the CNS that would fully determine the body's movements in advance must be abandoned. The CNS cannot regulate movement by sending "instructions" (in the form of muscle contractions) to the body, because the relation between muscle contractions and bodily movement is nonlinear: as we saw in the case of infant stepping, the same muscle contractions that produce kicking when the body is supine may produce no movement at all when the body is erect. Since the effect of any given muscle contraction is so variable, the nervous

---

101 This point was first made by the Russian physiologist Nikolas Bernstein. See Berstein (1967); Thelen &
system can only regulate bodily movement if it receives *constant feedback* from the moving body. But this means that the nervous system cannot regulate bodily movement *without also being regulated by this same movement*. In other words, these two dynamic systems—the nervous system and the moving body in which it is imbedded—must be coupled to each other in a circular feedback loop. It follows that the mechanistic view of the CNS as a central control unit that could be understood apart from what it controls—a cause that is indifferent to its effects—is untenable. It is the ongoing, circular *interaction* between the nervous system, the moving body, and the world that gives rise to behavior in real time (Fig. 25).

![Fig. 25. The self-organization of behavior.](diagram)

Mechanistic theories of behavior seek a single cause that would contain already all of the complexity and coordination of the body's movement. But this search is in vain, for no such cause exists. Instead, complex, coordinated forms emerge dynamically out of the interactions between the CNS, the body's own biomechanics, and its surrounding context—without being determined in advance by any *one* of these. This picture of motor

---

3.3. Learning to move

behavior as self-organizing dynamically in real time allows us to make sense of the flexibility of behavior, the living body's capacity to improvise new movements on the fly as its situation changes. When one or more legs are amputated from a dung beetle, it will immediately resume walking by reorganizing its gait, spontaneously inventing a new mode of locomotion (SB 39-40/39-40). Similarly, when my own foot begins to hurt, my body of its own accord finds a new way of walking that takes some of the weight off that foot. On a mechanistic account, behaviors such as these must be stored up in advance as programs or reflexes, which are merely triggered by changes in the body or its circumstances (ibid.). On a dynamic account, on the contrary, motor behaviors such as walking are self-organizing forms—a kind of stable equilibrium between the body and the world. When either body or world changes, the equilibrium between them will naturally shift as well, without needing any mechanism in the body to detect the change and switch over to a different pre-programmed response.

Mechanistic theories of behavior can only explain the coordination of movement by dividing it into discrete, pre-determined units. The question, for these theories, is how the CNS assembles complex movements out of individual muscle contractions. But we have now seen that this question is ill-posed. The relation between the body's muscle activity and its movements is nonlinear: there is no one-to-one correspondence between individual muscle contractions and individual movements, and the effects of muscle contractions on bodily movement are not additive. Thus the living body's movements are not composed of discrete muscle contractions. Its behavior, like its anatomy, can only be understood as a self-organizing form—a whole that articulates itself into parts. The organism does not build new behaviors out of pre-existing parts, like machines, but
3.3. Learning to move

grows them in the same way that it grows new organs: through a cascade of symmetry-breaking bifurcations. We now turn our attention to this development.

C. Motor development as morphogenesis

Not only does the dynamic account of behavior help us to make sense of how behavior adapts to change in real time, but it can also help us to understand how and why behavior changes over the course of development. Mechanistic theories of behavior must posit not only a fixed "motor program" behind each of the body's movements, but also a further "developmental program" to explain why these motor programs gradually change over the course of development. Because mechanistic theories appeal to static entities for explanation, change is always a problem for them. For our dynamic account, on the contrary, change need not be grounded in static structures. Instead, stasis is explained in terms of change: what appear to be static structures are in fact temporarily stable forms arising out of underlying dynamic processes. Thus the very same processes that generate stable, self-organizing behaviors also generate instability and the growth of new, more complex behaviors. We learned in previous chapters that nothing about the living body is fixed: its anatomy is also in movement, though the growth of anatomy is much slower, and therefore less apparent to us, than the movements of behavior. To understand motor development, we must understand how these two movements—growth and behavior—are related.

We saw above that motor behavior arises in real time out of the ongoing interaction between the nervous system, the moving body, and the world, and that this explains the real-time flexibility of behavior: when the body or the world changes,
3.3. Learning to move

behavior naturally changes along with them. These same dynamics also explain how behavior changes over the course of development: slow changes in the body's anatomy (including its neuro-anatomy) alter the stability landscape of the moving body, giving rise to new self-organizing patterns of movement. We have already encountered one example of this: the disappearance of newborn stepping. Thelen and colleagues discovered that the disappearance of this behavior was due to the increasing weight of growing infants' legs. Kicking remained a stable pattern when infants were supine, because these kicks did not need to lift the legs against gravity. When the infants were held erect, however, the same muscle contractions were no longer sufficient to set the legs into motion. This example illustrates the continuity between behavioral development and the anatomical development we studied in previous chapters: the growth of the body (including the nervous system) alters the parameters of bodily movement, giving rise to new motor behaviors (and the disappearance of old ones).

However, this continuity goes both ways: growth changes behavior, but behavior also shapes anatomy. This takes place in two ways, both familiar to us from our own lives. The first is well known to us under the name of "exercise": the way we use our bodies slowly changes our musculo-skeletal anatomy over time. Our own movements generate biomechanical and biochemical changes in our muscles and bones, which grow larger, stronger, denser, and more efficient the more we use them. The second is what we call "learning": the way we move our bodies changes the anatomy of our nervous system. The athlete or the musician who practices the same movements over and over again is not only strengthening her muscles and bones, but also reorganizing the synaptic connections in her brain. Mechanistic prejudices have led scientists to look for a fixed architecture in
3.3. Learning to move

the brain, in which each function would have its own pre-determined location. But in fact, the brain's organization is quite flexible—especially in children, whose brains are still growing (Kelso 1995, Ch. 9).

![Diagram showing the feedback between behavior and anatomy](image)

**Fig. 26. Motor development arises from the feedback between behavior and anatomy**

I have attempted to depict these relationships between behavior and anatomy schematically in Fig. 26. The central figure in the diagram represents the feedback loop between the limb dynamics of the moving body and the dynamics of the body's nervous system. Behavior arises in real time from the coupling of these two dynamic systems. At behavioral time scales—i.e. seconds or minutes—the anatomy of the body (including the nervous system) is relatively invariant, and can be treated as a set of fixed parameters for the moving body as a dynamic system. Over developmental time scales, however, the tiny anatomical changes wrought by each real-time movement add up, altering the stability landscapes of the moving body and its nervous system. These changes eventually give rise to *symmetry-breaking bifurcations* in the dynamics of these coupled dynamic systems, which generate qualitatively novel motor behaviors.
3.3. Learning to move

I have used the familiar terms "learning" and "exercise" to distinguish between changes in musculo-skeletal anatomy and changes in neuro-anatomy. However, we must not be fooled by this distinction into thinking that learning is something that takes place only in the brain, or something the brain could do on its own without the body. On the contrary, as Fig. 26 aims to show, the nervous system's anatomical development, like its real-time dynamics, is tightly coupled to the development and dynamics of the rest of the body. However, they tend to develop at different rates: neuro-anatomy typically changes more quickly than musculo-skeletal anatomy. Thus one can learn to move in a new way in just a few minutes or hours, with the relevant anatomical changes taking place almost entirely within the nervous system. This is the process that I have labeled "learning" above; but this is somewhat misleading, since larger developmental changes, such as learning to walk, will involve changes not only in the nervous system but also in the rest of the body.

Mechanistic theories of development try to locate the cause of motor development in anatomical growth, which they take to be determined in advance of, and therefore indifferent to, its effects on behavior. But we have now seen that anatomy is both cause and effect of behavior. Anatomy conditions behavior in real-time; but over the course of development these real-time behaviors slowly alter their own conditions, giving rise to new behaviors. Indeed, we can now see that behavior and development are not two

---

102 For the sake of simplicity, this diagram omits other aspects of the body's anatomy such as the endocrine system and the cardio-pulmonary system. These can be viewed as dynamic systems in their own right, and they are also coupled to both the moving body and the nervous system. Thus movement and neuro-dynamics are coupled in other ways not shown here.

103 Even talking about the moving body and the nervous system as two distinct dynamic systems can be misleading, given how tightly coupled they are. It may be useful for heuristic purposes to think about the intrinsic dynamics of each system in isolation; but the stability landscapes of the two coupled systems are quite different than they would be if the systems were isolated.
3.3. Learning to move

separate phenomena, but the same process observed at two different time-scales. Think again of the athlete or the musician practicing a given movement. The aspiring pianist plays the same scales again and again; slowly, imperceptibly, this repetition leads to improvement. Each repetition is slightly different from the one before, though this difference is vanishingly small. It is only over weeks and months that we can hear how the student's playing has changed. But there are not two separate movements here, one of playing the scale and the other of developing greater skill. Rather, the playing and the development are one and the same movement, observed at two different scales. When the student says, "I'm practicing," she means both, "I am playing the piano," and "I am getting better at playing the piano." A practice is an activity that we engage in precisely in order to change ourselves, to grow. As we change, our practice changes too, opening up new possibilities for action and new avenues for growth.

We can think of the relation between behavior and anatomy as analogous to the relation between a river and its riverbed. The direction of water's flow at any given time is determined by the shape of the landscape; but over time, the water's movement reshapes the land by erosion, changing the direction of its own flow. Thus there is a circular relation between moving water and apparently static landscape: the land shapes the river, but the river also—more gradually—shapes the land. The banks of the river are also in motion, though their movement is too slow to be seen at normal human time-scales. If we want to explain why the water flows the way it does, we cannot simply appeal to the shape of the land as the cause of the river's path, as if the landscape were fixed and indifferent to the water that flows through it. The shape of the land is a product of the water's flow; the river carves its own path through the landscape over time.
3.3. Learning to move

Similarly, we cannot appeal to anatomy as the cause of behavior, as if the body's anatomy were fixed and indifferent to its movements. The shape of the body is a product of its own activity, just as the shape of the land is a product of the water's flow.

Fig. 27. Ontogenetic Landscape for Locomotion (Thelen & Smith 1994)
3.3. Learning to move

Thelen and Smith (1994) have tried to depict this complex process of development in an "ontogenetic landscape" diagram. Thelen and Smith (1994) have tried to depict this complex process of development in an "ontogenetic landscape" diagram. \(^{104}\) Fig. 27 shows one such diagram, which depicts the development of human locomotion. The diagram has three dimensions. Each horizontal line represents the stability landscape of the moving body as a dynamic system at a single moment in time. As with the stability landscapes we studied in Ch. 2, the curvature of the line represents the stability of the system: "basins" represent stable attractors of the system's dynamics; the deeper the basin, the more stable the attractor. The vertical dimension of the diagram represents time, moving forward in time as we move from the top of the diagram to the bottom. Thus changes in the curvature of the horizontal lines as we move from top to bottom represent changes in the stability landscape of the moving body as a dynamic system. As you can see, most of the horizontal lines have multiple basins, meaning that the moving body is a multistable system which can shift between multiple stable behaviors. A change in the body's situation can push it out of one basin and into another. For example, a change in the slope of terrain can shift the body from walking to climbing; a change in the speed of locomotion can shift it from walking to running (Kelso 1995, Ch. 3). Over time, the number of stable behaviors available to the growing child increases through a series of symmetry-breaking bifurcations, depicted by the transformation of one basin into two. Just as the movement of the river reshapes the landscape it flows through, the activity of the moving body reshapes its own stability landscape. The ontogenetic landscape diagram thus depicts, in a qualitative way, how a behavior such as walking can be a robust outcome of human motor development without being specified anywhere in advance.

\(^{104}\) The concept of an ontogenetic landscape is adapted from biologist C.H. Waddington's "epigenetic
Mechanistic theories of behavior try to locate the cause of behavior in the body's anatomy, and the cause of this anatomy in a genetic plan. This is a description drawn from manufacturing: the machine's movements are determined in advance by its parts and their arrangement, which are in turn pre-determined by its design. But this is a poor description of the living body, whose anatomy is never fixed, cannot be determined in advance, and is continuously altered by the very movements it generates. Aristotle famously observed, with respect to human life, that our actions shape our character: we become what we do (Nicomachean Ethics II.1). We have now seen that this is the structure of all life: what the living body is at any given moment (its anatomy) is a product of what it has done in the past (its behavior). Neither the form of the living body nor the form of its movements can be determined in advance, for both are constantly developing "on the fly" through the body's creative engagement with its surroundings.

From a mechanistic standpoint, learning how to move means assembling new behaviors out of pre-existing units. But we have now learned that the body's movements are not built out of discrete muscle contractions. Development does not begin with a set of parts that must be assembled into a whole, but with a whole that must articulate itself into parts. Thus the body grows new behaviors in the same way that it grows new organs: through a cascade of symmetry-breaking bifurcations. We saw in our study of newborn kicking that the infant's lower body initially acts like one big muscle: the muscles of the legs can only contract all at once. In order for the infant to develop more complex, coordinated behaviors, this single muscle must articulate itself into many smaller ones. That is, the pattern of muscle activity generated by the infant's nervous system must
become less symmetrical. This process can be observed in the development of supine kicking over the infant's first year. We noted above that supine kicking, unlike erect stepping, does not disappear at two months; but neither does it simply persist unchanged. On the contrary, the alternating kicks of the newborn are gradually replaced by other, less symmetrical patterns: first single-leg kicking, often favoring one of the two legs; and later kicking with both legs simultaneously (Thelen & Smith 1994, p. 93).

This loss of symmetry is driven by the growing infant's interactions with its environment. We noted above that the nervous system and the moving body are coupled dynamic systems, with the nervous system receiving constant feedback from the moving body. This feedback occurs in two ways: through nerve endings in the muscles themselves, and through the body's sense organs. Thus the regulation of bodily movement circulates through the world; the growing body learns how to move in a more coordinated way through the perceptual feedback it receives from its own movements. This perceptual feedback also accounts for the increasing responsiveness of the child's movements to its surroundings. In the next section, we will study perception in more detail, in order to understand the role it plays in behavior and development.

4. Learning to perceive

A. Perception and the self-organization of behavior

Have you ever tried to walk around your own home in the dark without bumping into anything? It is surprisingly hard for a sighted person to navigate even the most familiar space without visual guidance. Indeed, even standing in one place becomes difficult for the sighted person deprived of her usual visual reference points. Try to keep
3.4. *Learning to perceive*

your balance while standing on one leg; now, see how much harder it becomes when you close your eyes. Harder still is to stand on a leg that has gone numb—we cannot balance on a foot that does not feel the floor beneath it pushing back against our body's weight. In sum, even the most practiced and familiar movements become difficult or impossible without constant perceptual feedback. If the purpose of perception were to construct an internal representation of the world, as Cognitivism proposes, then these phenomena would be difficult to explain. Once I have constructed a reasonably accurate internal model of my environment, this model should suffice to guide my movements without further perceptual feedback, as long as my surroundings stay relatively fixed. The necessity of constant perceptual feedback—even in static environments—suggests that the purpose of perception is not to generate an internal representation of the world, but rather to regulate the body's movements in real time. It is a longstanding philosophical prejudice that we perceive in order to *know*, and that knowing consists in representing.\(^{105}\) But these ideas have no basis in biology. The living body perceives not in order to know, but in order to *act*. It needs no internal model of the world to guide its movements, for it receives constant perceptual feedback from the world itself (Brooks 1991).

---

\(^{105}\) These prejudices have been identified and criticized at length in Heidegger's *Being and Time* (1962 [1927]).
3.4. Learning to perceive

Mechanistic theories of behavior can only explain the perceptual regulation of bodily movement by positing that the organism's movements are *caused* by things in its environment (Fig. 28). The problem of behavior, on this view, is to explain how things in the world affect the body's nervous system, and how these effects give rise to the body's movements. Mechanistic theories try to answer these questions by dividing world, perception, and behavior into discrete parts. They assume that the effects of the world on the nervous system can be analyzed into a collection of individual stimuli, and that bodily movement can be analyzed into a collection of individual muscle contractions. Thus the project of Behaviorism and Cognitivism alike is to reverse-engineer the nervous system in order to understand how different stimuli cause different muscle contractions.

On this view, the relation between the nervous system and motor behavior is one of linear causation: each part of movement has its own discrete cause within the nervous system, and these causes are indifferent to their effects. Thus the form of bodily movement is present in advance in the nervous system's activity. The CNS issues "commands" to the body, and these commands taken together constitute a "plan" or "program" for how the body will move.

---

106 Machines are not self-moving; their natural state is to be at rest. Thus if the living body is a machine, the causes of its movement must ultimately lie outside of it.
However, we learned above (3.2) that the causal relation between the nervous system and the moving body is not linear, but circular: the moving body and its nervous system are coupled, nonlinear dynamic systems (Fig. 29). Thus the effects of muscle contractions on bodily movement are neither constant nor additive: there is no one-to-one relation between muscle activity and bodily movement, which means that the nervous system cannot issue "commands" to the moving body, but only modulate its intrinsic dynamics. Since the effects of muscle contractions on movement are nonlinear and therefore highly variable, the nervous system cannot be indifferent to these effects, but must receive constant feedback from the movements it helps to generate. Perception is the body's way of producing this feedback: every movement of my eyes generates a change in my visual field; every movement of my hands produces a different pattern of tactile feedback. Thus the body's movements modulate the intrinsic dynamics of its nervous system via its sense organs. Perception is the regulation of neuro-dynamics by
3.4. Learning to perceive

bodily movement. Together, perception and muscle activity couple the moving body and its nervous system, generating a continuous feedback loop between these two dynamic systems.

Mechanistic theories assume that perception is passive: the nervous system waits to be affected by things in the world, and these effects are determined entirely by their causes. In the same way that the nervous system is supposed to be the linear cause of movement, things in the world are supposed to be the linear cause of the nervous system's activity (Fig. 28). Every change in the nervous system would be the effect of a discrete cause—the individual stimulus—and every stimulus would be the effect of some discrete thing in the body's surroundings. Thus the form of nervous activity would be given in advance by the form of things in the world. The total pattern of stimuli, like the pattern of bodily movement, would be no less symmetrical than its causes.

However, the nervous system is a nonlinear dynamic system, like the moving body. Thus the effects of individual "stimuli" on this system are neither constant nor additive. In other words, there is no one-to-one correlation between the activity of individual sensory neurons and the activity of the nervous system as a whole. This means that perception cannot be analyzed into discrete parts any more than movement can. A change in any single neuron is a tiny fluctuation in the nervous system's dynamics, and its effects will vary according to the global state of the system at that moment. Thus instead of trying to understand perception as a collection of individual stimuli, we must see the body's sense organs as constituting a total perceptual field: a global sensitivity to the body's surroundings which changes constantly with the nervous system's own dynamics. Just as the form of movement is not determined in advance by the nervous system's
3.4. Learning to perceive

activity, this activity is not determined in advance by the form of the body's surroundings. The nervous system, like the moving body, is a self-organizing dynamic system whose activity is more asymmetrical than its causes.

Furthermore, we have now seen that perception is not passive, but active: it is the body's own movements that produce changes in its perceptual field. Thus the body does not wait to be "stimulated" by its surroundings, but actively stimulates itself by moving. This point is easier to grasp with respect to some sense modalities than others. For example, it is not hard to see that tactile perception is active. We do not simply wait for things to touch us, but actively explore them with our hands: picking them up, palpating them, and running our fingers over them. The texture of a surface can only be felt by moving our hands over it at the right speed and in the right direction. Roughness and smoothness cannot be perceived in an instant, by a single touch, but only through movement. In the case of vision, however, we are more tempted to think that perception is passive, like a camera taking a picture—perhaps because we are less aware of the movements of our eyes than we are of the movements of our hands. Studies show, however, that movement is no less necessary to vision than it is to tactile perception. If the eyes are prevented from moving—or if an image is "stabilized" on the retina by tracking the eye's movements—visual perception fades to a homogeneous field (Martinez-Conde et al 2004). (See Fig. 30 for a simple demonstration of this.) To see is to actively explore the world with our eyes, just as we do with our hands—"palpating it with our look" (VI 131).107

---

107 See also Noë & O'Regan (2000): "Just as it would be curious to propose a mechanism whose purpose it is to compensate for hand movements in haptic object recognition (for example, a mechanism that ensures that the object is not perceived as disappearing when the hand is briefly lifted off it), there is no need for a
3.4. Learning to perceive

Fig. 30. Fix your eyes on the red dot, and you will find that the grey annulus disappears. Under voluntary fixation of the eyes, stationary objects in the periphery of vision fade. Movement of the eyes or the peripheral objects makes the objects reappear. (Adapted from Martinez-Conde et al 2004.)

Perception, then, is not simply an effect of the world on the body; rather, the body is the cause of its own perceptions. Perception and movement are thus two inseparable aspects of one sensorimotor loop: the nervous system modulates the intrinsic dynamics of the moving body through muscle contractions, and the body's movements in turn modulate the dynamics of the nervous system through perception. It follows that we

[108] "[W]e should consider the afferent sector of the nervous system as a field of forces which express concurrently the intraorganic state and the influence of external agents; these forces tend to balance themselves according to certain modes of preferred distribution and to obtain movements from the mobile
3.4. Learning to perceive

cannot understand perception simply by studying the effects of different situations on individual sensory neurons, or even on the total perceptual field of a stationary organism that has been sedated and restrained. Instead, we must study how the body's own movements produce changes in its nervous system *via* the world that it is moving through. But this requires a significant shift in the way we think about the role of the world in perception. We can no longer divide the body and its surroundings into parts, and ask how each part of the world affects each part of the body. And we can no longer think of the body as passively receiving messages from an active world. Instead, we must learn to reverse these roles—to see the body as active in perception, and the world as passive. We can then ask how the moving body uses its surroundings to generate the perceptual feedback it needs to guide its own movements, and how different situations allow the body to generate different kinds of perceptual feedback.

This is not an easy shift to make. But we can begin by noticing the following fact: in order for the body's movements to generate perceptual feedback, its surroundings must be *asymmetrical*. In a uniform environment, the movements of the body have no effect on its sense organs. For example, within a *ganzfeld*—a field of completely uniform color and illumination—one sees the same thing no matter where one looks, (provided of course that one does not look at one's own body). Thus the movements of one's eyes, head and trunk do not produce any visual feedback, any changes in one's visual field. Similarly, we can imagine a uniform tactile environment, in which one would feel the same thing no

parts of the body which are proper to this effect. These movements, as they are executed, provoke modifications in the state of the afferent system which in turn evoke new movements. This dynamic and circular process would assure the flexible regulation which is needed in order to account for effective behavior*" (SB 46/48-9).
3.4. *Learning to perceive*

matter how one moved one's hands. (This could perhaps be simulated by donning very thick, padded gloves.) In such situations, perception becomes impossible—not because the body's sense organs have suddenly become defective, but because an environment without differences offers *nothing to perceive*, no responses to the body's questing movements. Thus *perception requires an asymmetrical world*. The less symmetrical the environment is, the more rich and nuanced will be the perceptual feedback it offers to a moving body. We can see this by considering a few highly abstract, schematic "environments."

![Fig. 31](image)

First, imagine an environment that is completely uniform except for a single point (Fig. 31). This is as symmetrical as an environment can be without being completely uniform. A body exploring this space with its eyes or hands will receive very little sensory feedback. It might never find the point at all; and even if it does, moving across the point will produce only a tiny, momentary change in its perceptual field.

---

109 "A truly homogeneous area, offering *nothing to perceive*, cannot be given to *any perception*" (PP 4/26, emphasis in original).
3.4. Learning to perceive

Next, imagine a slightly more asymmetrical environment that is divided into two regions, each completely uniform, but different from each other (Fig. 32). This environment is richer than the previous one, but only slightly: it offers perceptual feedback to an exploring body only when the body crosses from one region to the other. These are the only movements that produce a change in the body's sensory field. As long as the body remains within one region, there is nothing to perceive. Only the boundary between the two regions is perceptible; it is the difference between the two regions that the body perceives, rather than any absolute property of either region taken on its own.

Gestalt psychologists long ago observed that the smallest possible percept is not a sensation or punctiform sense-impression, but a figure on a background (PP 4/26). We can now both confirm this observation and also make it more precise: the smallest possible percept is not a sensation, but a difference. A figure on a background—such as the point in Fig. 31—is the smallest possible percept because it is the contrast of figure with ground that makes them perceptible as such.
3.4. Learning to perceive

These first two "environments" were highly symmetrical. Let us now consider a less symmetrical one, which offers more perceptual feedback to a moving body (Fig. 33). Unlike the first environment, which contained just a single point against an otherwise uniform background, this one contains many points, regularly distributed. In this environment, unlike the previous two, almost any movement will produce some corresponding change in the body's sensory field. Exploring this space will generate a rhythmic pattern of perceptual variation, the rate of which will vary with the speed of the body's movements. Thus this environment, unlike the previous two, offers the body feedback on how fast it is moving—provided that the body is sensitive not only to changes in its perceptual field, but also to changes in the rate of those changes. Notice how the spatial asymmetries of the environment are converted into temporal asymmetries (patterns of periodic change) by the body's movements. To be more precise, it is temporal asymmetries in perceptual feedback that give rise to the perception of a spatially asymmetrical environment.
Finally, consider an even more asymmetrical "environment," this time one that offers a smooth gradient of differences to perception (Fig. 34). Here, as in the previous environment, almost any movement will generate some perceptual feedback, and the rate of perceptual variation will correspond to the speed of the body's movements. But unlike the previous environment, the rate of perceptual variation here will also be highly sensitive to the direction of the body's movement, varying from a maximum when the body moves along the "B" axis to zero when the body moves along the "A" axis. Thus this environment offers an exploring body feedback not only on how fast it is moving, but also on the direction of this movement.

As these schematic examples show, it is the asymmetry of our surroundings—the fact that the world is different in one place than in another—that makes perception possible. This asymmetry is the very texture of reality, which allows us to get a perceptual grip on our surroundings. However, our schematic environments have so far all been static; they are asymmetrical in space, but not in time. What about the temporal texture of reality? How do the dynamics of the changing world affect the perceptions of the moving body? We should begin by noting again that perception always involves changes in the body's perceptual field over time. Thus the difference between a spatial
3.4. Learning to perceive

asymmetry and a temporal one, from the point of view of the living body, lies in the different temporal patterns of perceptual feedback they offer to the moving body. Spatial asymmetries produce changes in the perceptual field when the body moves, but not when it is at rest. Temporal asymmetries produce changes in the perceptual field when the body is at rest, but not when it moves along with them in the right way. For example, to track a moving object visually is to move one's eyes and head in such a way that the object remains centered in one's visual field. If we treat the body's movements as transformations (in the mathematical sense) of its perceptual field, then we can say that spatial differences in the world show up as *asymmetries* in the perceptual field (i.e. transformations that *alter* the field), whereas temporal differences show up as *symmetries* (transformations that leave the perceptual field *unchanged*). A perceptual field that remains stable in response to the body's own movements is perceived as a change in its surroundings. Thus, patients whose eye muscles have been temporarily paralyzed experience the world as moving to the left when they try to move their eyes to the left (PP 48-9/74). A disruption of the body's habitual movements can create the illusion of movement in its surroundings.

Psychologists have also produced the converse effect, creating the illusion of bodily movement by disrupting the body's habitual surroundings. Psychologists studying the visual regulation of balance in human beings placed toddlers in a "room" whose walls and ceiling were suspended from above, allowing them to swing forward and backward (Lee & Aronson 1974). When the wall the toddlers were facing swung away from them, they lost their balance, stumbling or falling forwards. When the wall swung toward

---

110 Video footage of this experiment can be found online: [http://youtu.be/RJrEnK8tQxc](http://youtu.be/RJrEnK8tQxc).
3.4. Learning to perceive

the toddlers, they stumbled or fell backwards. The investigators interpreted these results as showing that toddlers use the "optic flow" patterns produced by their own movements to regulate their balance. When a toddler sways forward, asymmetries in her visual field "stream outward," moving away from the center of the field; when she sways backward, asymmetries in her visual field stream inward. The movements of the swinging room in this experiment simulated the optic flow patterns ordinarily produced by the toddler's own swaying movements. Because the subjects were accustomed to rooms with stable walls, they perceived the change in their visual fields as a movement of their own bodies, generating a compensatory movement in the opposite direction which caused them to lose their balance.

This experiment offers an excellent demonstration of the role perception plays in the regulation of bodily movement. The toddler's balancing movements are regulated by the changes they themselves produce in the toddler's visual field. When asymmetries in the visual field stream outward—as they do when the toddler sways forward—her weight shifts backward to compensate. When asymmetries in the visual field stream inward, her weight shifts forward. Thus balance arises from the feedback loop between the body's movements and the changes they produce in the nervous system via vision. This feedback loop generates upright balance as a stable state by damping out perturbations in the body's posture—namely forward and backward sway (Fig. 35). I argued earlier that perceptual feedback is necessary for the nervous system to regulate bodily movement. But we could equally say that the purpose of both nervous system and perception is to allow behavior to regulate itself. That is, the basic function of the nervous system is to
3.4. Learning to perceive

create a feedback loop through which the body's movements will be regulated by their own (perceptual) effects.\footnote{See Varela (1997): "Whenever motion is an integral part of the lifestyle of a multicellular, there is a corresponding development of a nervous system linking effector (muscles, secretion) and sensory surfaces}

![Stability landscape for balancing body as a self-regulating dynamic system](image)

If the organism's behavior is acutely sensitive to changes in its environment, it is not because the environment is the \textit{cause} of behavior, as mechanistic theories would have it, but rather because the \textit{self-regulation} of behavior depends on the asymmetries of its surroundings. These asymmetries \textit{condition} behavior, but they do not cause it. A change in the organism's environment may produce a change in its behavior, as in Lee and Aronson's experiment. But the fact that moving walls cause toddlers to lose their balance does not imply that stationary walls \textit{cause} toddlers to balance successfully under normal circumstances. It would be more accurate to say that the toddler's body takes advantage of the visual asymmetries presented by stationary walls to facilitate the self-regulation of...
3.4. Learning to perceive

its upright stance. Thus instead of asking how the body's movements are caused by things in its environment, we should ask how changes in the body's situation, or differences between one situation and another, make a difference to the self-regulation of the body's movements.

B. Motor programs vs. sensorimotor habits

It is the self-regulating, self-organizing character of behavior that gives it the appearance of intelligence and purposiveness. The balancing body adapts to perturbations in its posture, resisting forces that push it out of balance, and finding ways of returning to an upright stance. Thus the body seems to have a certain optimal posture as its goal and to pursue this goal in creative ways, finding intelligent solutions to problems posed to it by the world. Faced with the apparent goal-directedness of behavior, it is tempting to think that the goal must be given in advance of the behavior it directs, as some kind of representation. We encountered this same logic in our discussions of growth (Ch. 1) and motor development (3.2): the robustness of developmental outcomes makes us think that they must be given in advance as plans or programs directing the course of development. Similarly, the robustness of behavioral outcomes such as balance—their resistance to perturbations and their adaptation to changing circumstances—makes us think that these outcomes must be given in advance as representations that direct the body's movements. (For example, we might posit a motor program for balance that measures the body's deviation from its desired vertical stance, and generates motor commands designed to close the gap between the body's current state and its goal-state.) However, we have

(sense organs, nerve endings). The fundamental logic of the nervous system is that of coupling movements
learned that self-organizing dynamic systems can generate seemingly goal-directed behavior without plans, designs, or representations. The robustness of the toddler's upright posture is not due to any explicit representation of its desired stance, but to the self-regulation of its movements. The body acts as if it knows the difference between a leaning and an upright stance, between a movement away from and a movement towards its goal; but this "knowledge" is not a representation or a static structure located somewhere within the body—it is not a thing that one could point to. It is a knowledge in movement, a dynamic equilibrium that emerges from the circular interactions of the body with its surroundings, and of the parts of the body with one another. Thus we find in behavior—as in growth and motor development—an intelligence without consciousness, a purposiveness without concepts or representations, and a goal-directedness whose goal is not given in advance.

Of course, there is more to intelligent behavior than robustness and stability; if a behavior is too robust, then the organism will always be doing the same thing, regardless of its circumstances. Skillful behavior requires not only that the body be able to achieve the same outcome in varying conditions, but also that it be able to change its goals as the situation changes—for example, switching from pursuing prey to fleeing from predators to attracting a mate. Thus the moving body, as a self-organizing dynamic system, must be multi-stable: it must have many stable attractors, each of which is a robustly self-regulating, skillful behavior in its own right. If any one of these behaviors is too stable, the organism will be unable to flexibly switch from one behavior to another, instead getting stuck in one "mode," one sensorimotor feedback loop. We might think of

---

with a stream of sensory modulations in a circular fashion."
3.4. Learning to perceive

obsessive behavior in human beings as an example of such an excess of stability. Conversely, an inability to stay "on task" without getting distracted might be an example of excessive instability. Thus skillful behavior depends on having the right mix of stability and instability. A body that does the same thing in every situation is not acting skillfully; but neither is one that changes its behavior in response to trivial or irrelevant changes in its environment.

Since behavior arises dynamically through the interaction of the body with its surroundings, it is intrinsically context-sensitive. Psychologists have long acknowledged the context-sensitivity of behavior while still attempting to preserve some static "essence" of behavior, arguing that the body's variable performances are imperfect reflections of more stable competencies located within the organism (Thelen & Smith 1994, pp. 25-27). On our dynamic view, this distinction between competence and performance is untenable. The body is not the sole cause of its own movements; every action is a collaboration between the body and the world, which neither is capable of producing on its own. But if the organism's own behavior is not something it can "pull off" on its own, then we can no longer locate the organism's powers and abilities exclusively within its own body. Learning a skill—such as how to walk, or how to pick things up with one's hands—is not like acquiring a new tool which one would then possess absolutely and be able to deploy in any circumstance. Rather, the body's powers belong not to the body alone, but always to the-body-in-a-situation.112 For example, it turns out that the toddler who has learned to stand on her own two feet has not acquired an unconditional skill, but

---

112 We are here discovering with respect to the living body a principle analogous to Heidegger's (1962 [1927]) definition of human subjectivity as "Being-in-the-World".
3.4. Learning to perceive

one that depends on certain environmental conditions. In a situation whose spatial and
temporal asymmetries diverge too far from those she is accustomed to—for instance, in a
room with moving walls—the toddler's habitual way of visually regulating her upright
posture may fail. Indeed, Lee and Aronson's experiment shows that in certain situations,
the same sensorimotor habits that normally enable the toddler to stand upright will cause
her to lose her balance. Thus the very habits that empower us in familiar situations may
prove disabling in unfamiliar circumstances.

In the body's habituation to certain familiar situations, we find another kind of
"knowledge in movement." The toddlers in Lee and Aronson's experiment have learned
from experience that walls do not move—but this knowledge does not reside in an
internal representation of walls as objects with certain properties. Rather, the toddler acts
as if walls were stationary; she perceives the movement of walls in her visual field as a
movement of her own body, and reacts accordingly. This is not a matter of judgment or
inference, as we have seen, but an immediate result of the body's sensorimotor loop,
which joins perception and movement together so closely that no intellectual act could
intervene between them. The optic flow pattern of the moving wall has a certain
significance for the toddler's body, which is expressed in her motor response. We can
read in the loss of the toddler's balance the meaning that the wall's movement holds for
her body. But this is not an intellectual significance, for it relies on no concepts,
judgments, or inferences. Rather, it is a motor-perceptual significance which arises from
the body's sensorimotor habits (PP 144/178).

The toddler's sensorimotor knowledge of the world can also be described as a
kind of expectation: toddlers expect that when they encounter new rooms, these rooms
3.4. Learning to perceive

will be stationary, like those they have encountered in the past. When this expectation is not fulfilled, the toddler falls flat: like someone expecting to find a step beneath her foot and encountering only air, the toddler suddenly finds that her environment is not supporting her habitual movements in the way she has come to expect. However, this expectation is not an *idea*—it is not an image or a map which might or might not resemble the "external world." The expectation cannot be located anywhere else than in the body's sensorimotor habits, its developed ways of being sensitive to the perceptual feedback generated by its own movements. The living body carries its history along with it: its habits of moving and perceiving grow out of its encounters with a particular set of concrete situations, and they bear the mark of the conditions in which they developed. It is in the context of these past experiences, with the sensorimotor habits that developed out of them, that the living body perceives and moves in every new situation it encounters. Thus, how the body will perceive and respond to a given situation can be neither understood nor predicted simply on the basis of that situation's own characteristics. Rather, the motor-perceptual significance a situation has for the organism depends not only on the situation itself, but also on the whole history of moving and perceiving that the body brings to bear on it. A toddler who had learned to stand in different circumstances—for example, one who had never before encountered a room with orthogonal walls—might react quite differently to Lee and Aronson's experimental set-up. It is precisely because of their familiarity with rectilinear spaces that the subjects of this experiment experienced the moving walls as unfamiliar and disruptive of their established motor habits.
Mechanistic theories of behavior face great difficulties in accounting for the flexibility and generality of behavior, because they appeal to static structures to explain the coordination of behavior. In order to explain a skillful motor performance in terms of a static mechanism or motor program in the nervous system, one must explain how the body is able to adapt this program to the unique demands of each particular situation. Our dynamic account of behavior does not face this difficulty, because it does not assume that the coordination of skillful movement is given in advance of its performance, prior to the body's encounter with particular situations. A sensorimotor habit is intrinsically general (SB 30/30): it is not a fixed pattern of movements or a universal algorithm, but a determinate way of being sensitive to differences in the world, which is expressed in articulate movement. Such habits are neither completely particular, nor completely universal. The toddler who has learned to stand upright in the rooms of its home can also do so in rooms that it has never seen before; its sensitivity to certain optic flow patterns, though learned in one particular set of situations, turns out to be functional in a wider range of circumstances. On the other hand, the same toddler is unable to remain upright in a room whose walls are mobile; thus the range of circumstances in which its habits generate skillful behavior is not unlimited. The context-sensitivity of behavior, which we noted above, is a double-edged sword. On the one hand, it is what gives behavior its flexibility, its capacity to adapt dynamically and creatively to different situations. On the other hand, it means that the body's powers are always conditional on circumstances beyond its control. We have all experienced the bizarre failures to which behavior is susceptible, in which skills we thought we had mastered abruptly betray us: I trip over my own feet or spill my drink on myself for no reason, despite years of experience in
walking and drinking; the right words suddenly desert me at the crucial moment of speech; I am mysteriously unable to produce my partner's name—or worse, I call him by the name of another.

**C. Perception and motor development**

Behavior is a self-organizing form which arises from the dynamic interplay of the living body with its environment. If we treat behavior as the activity of a dynamic system, then the parameters of this system at any given time include both the anatomical form of the body and the form of its situation. When psychologists study behavior, they do so by varying one or more of these parameters, and observing how the organism's behavior changes in response. In dynamic terms, we could say that they are probing the stability of the system, searching for critical points at which the quantitative variation of some control parameter produces a qualitative change—i.e. a symmetry-breaking bifurcation—in the system's dynamics. (Of course, this is not always or even usually how psychologists understand their own experiments.) We can distinguish two kinds of behavioral experiment: those that vary parameters in the body's environment, and those that vary parameters in its anatomy.

The easiest experiments to perform are those which vary some aspect of an organism's situation and study its response. These experiments seek to discover which changes in the body's environment will leave its behavior unchanged, and which will give rise to different self-organizing behaviors. The designers of these experiments are usually asking questions either about what a given organism is capable of, or what it is sensitive to. They typically study the organism over a short enough time scale that its anatomy can
be treated as fixed. Thus this kind of experiment is *synchronic*, in the sense that it studies an organism at some particular point in its development, rather than asking about how its behavior changes over time. Psychologists who are interested in *diachronic* questions about the development of behavior must proceed differently: instead of studying how a given body responds differently to different situations, they must look at how an organism's responses to a given situation change as its body develops over time. This is generally more difficult than the synchronic experiment, since it is much harder to vary the form of the living body than it is to vary its situation; anatomical changes are slow and difficult to control. One approach that psychologists have developed in response to this problem is the so-called "microgenetic study." In this kind of study, the experimenter attempts to "speed up" the organism's development by artificially enhancing its powers of movement and perception, and observing how this alters the organism's performance in a given situation, as well as the consequences for the organism's general behavior and development (as compared to similar control subjects who did not receive the experimental "enhancement").

One such microgenetic study, performed by Esther Thelen, examined the development of leg coordination in young infants. Thelen's study was based on a series of important experiments on infant learning and memory by psychologist Carolyn Rovee-Collier and colleagues (Rovee and Rovee 1969). In these experiments, infants between nine weeks and six months were placed on their backs under a colorful, noisy mobile. This mobile was attached by a ribbon to the infant's ankle, in such a way that the infant's kicks would set the mobile in motion. The sounds and movements of the mobile increased in proportion to the strength and frequency of the infant's kicks. Subjects in the
control group were placed under the same mobile, but it was not attached to their legs; instead, an experimenter continuously activated the mobile, simulating the sights and sounds experienced by the infants in the experimental group. The two groups of infants reacted quite differently to their experience with the apparatus. Those in the experimental group became alert and engaged, smiling and laughing. They quickly began to kick harder and faster than the infants in the control group, and they grew upset when their legs were detached from the mobile and it stopped responding to their movements. The infants in the control group, who were exposed to the same moving mobile, but only as passive observers, were much more subdued. This showed that the infants in the experimental group were not responding to the pleasing sights and sounds of the activated mobile per se, but to the mobile's responsiveness—the way that the sights and sounds produced by the mobile covaried continuously with their own movements.

Thelen's microgenetic study was a variation on Rovee-Collier's experiment, in which a similar apparatus was used to test whether infants could learn not only to kick more, but to kick in a different way (Thelen 1995). At three months, infants usually kick either by alternating legs, or using a single, favored leg. It is rare at this age for infants to kick with both legs simultaneously. In Thelen's experiment, three-month-old infants were attached to the mobile in such a way that the greatest activation of the mobile could only be achieved with a double kick. As predicted, infants in this experiment learned to kick their legs simultaneously—a behavior they would not normally learn until later in life.

The genius of Rovee-Collier's apparatus is that it provides the infant with a precocious link between movement, sight, and hearing—a form of perceptual feedback that would not ordinarily be available to the infant at its current stage of development.
3.4. Learning to perceive

Infants under six months have quite limited mobility: they are still developing the strength and coordination needed to support their own heads and to roll over; they typically cannot crawl yet; and they are still learning how to grasp objects and bring them to their mouths. Thus the young infant's capacities for actively exploring its world are still very limited. In particular, the movements of its legs are not yet strong enough to produce significant changes in what it sees, hears, or feels with its mouth and hands. Rovee-Collier's apparatus acts as a kind of *prosthesis*, then, which amplifies the effects of the infant's leg movements on her perceptual field, allowing her to explore her environment by moving her legs. This exploration is limited to the mobile above the infant's head. But Thelen's microgenetic study demonstrated that even this limited exploration is sufficient to generate new, precocious motor behaviors in the infant—effectively accelerating its development by several weeks. This suggests that the infant's exploration of its surroundings in the course of ordinary development plays a key role in the growth of motor coordination. In order to establish this more conclusively, one would have to perform the converse experiment, depriving infants of perceptual feedback to see if this would retard ordinary motor development. Obvious ethical concerns prevent us from performing such experiments on human infants. However, they have been performed (ethically or not) on kittens (Held & Hein 1963). As expected, it was found that kittens deprived of motor control over their own perceptions did not develop normal motor coordination.113

---

113 In this experiment, pairs of newborn kittens were placed in an apparatus which allowed one kitten to move, while the other kitten was restrained in such a way that its movements followed those of the other kitten. Thus the two kittens received the same perceptual "stimuli," but one was a passive spectator while the other was permitted to generate its own stimuli by moving. We experience something similar as passengers in a car. You may have noticed that you can be taken along the same route many times as a
3.4. Learning to perceive

Thelen's microgenetic experiment equipped infants with a prosthesis that let them precociously explore their environment with their legs. In another microgenetic study, Needham and colleagues devised a prosthetic apparatus that allowed infants to precociously pick things up with their hands (Needham, Barrett, & Peterman 2002). The experimenters had three-month-olds wear velcro-covered mittens and presented them with velcro-covered toys, allowing them to experience prehension 0.5-1.5 months before infants typically learn how to reach for and grasp small objects. Needham and colleagues found that repeated experience with this prosthesis led to precocious developments in hand-eye coordination, and a precocious increase in object exploration with the eyes, hands, and mouth. In both of these microgenetic studies, prostheses that allowed infants to engage in precocious exploration of their environment led to precocious developments in motor coordination. These and other studies have led developmental psychologists to conclude that it is the growing body's active exploration of its environment that teaches it to move in new ways (Bourgeois et al. 2005; Palmer 1989). However, they tend to interpret this phenomenon in Cognitivist terms, claiming that the body explores its surroundings in order to acquire information about the properties of objects, which it then uses to plan its own movements. As the body acquires more information about its environment over time, its action plans become more sophisticated and well-adapted to passenger without ever learning how to follow that route as a driver. One might think that this was simply a matter of "paying attention"; but our study suggests that the meaning of the visual spectacle is different for the driver, whose movements the spectacle is responding to, than for the passenger who is a passive observer. What would it mean to "pay attention," as a passenger, in such a way that one could drive the same route oneself? One would have to put oneself in the driver's place, experiencing the motor possibilities presented by each intersection as if one were choosing the route for oneself.

114 Video footage of an infant using this prosthesis can be seen here: http://www.youtube.com/watch?v=hJGRM4LFJjU.
3.4. Learning to perceive

its surroundings. This account recognizes that the body learns about its environment by moving around in it, and that this learning manifests itself in new motor behaviors. Unfortunately, it succumbs to the temptation to reify this learning as "information" captured from the environment and stored somewhere in the body. In other words, this account recognizes that the body learns to move by moving—that the body's own movements are what drives its motor development. But it insists on interpolating, somewhere between one movement and the next, a static representation of the body's changing surroundings, and a static plan for its subsequent movements. I have argued, to the contrary, that the goal of perception is not to construct an internal model of the world, but to regulate movement in real time; and that the body's knowledge of its environment does not reside in any representation, but rather in its developed ways of using its surroundings to perceptually regulate its own movements. On this view, what the body gains from exploring its environment is not new information about the world, but new sensorimotor habits. The infant exploring its surroundings with mouth, eyes, and hands is not cataloguing abstract facts about objects and their properties, but learning to move and perceive in new ways.

In rejecting the traditional distinction between perception and movement, we have also ruled out any absolute distinction between exploration and action. When the psychologist is interested in the real-time coordination of the infant's movements, she calls them "actions"; when she is interested in how these movements develop over time, she calls them "exploration." But they are the very same movements in each case. The

---

115 "The information that infants register about objects guides their manual actions, but at the same time, these actions help infants to register important information about objects' properties" (Bourgeois et al. 2005).
3.4. *Learning to perceive*

body's actions are always perceptually guided, which means that every action is a perceptual exploration of the world. When I walk, I am exploring the ground with my feet and the world moving past me with my eyes; when I hit a tennis ball, I am exploring the ball with my racket. Conversely, perception takes place only through movement, which means that all perception is a kind of action. When I look at the world around me I am actively exploring it by moving my eyes; when I feel the shape, weight, and texture of an object, I am actively exploring it by running my fingers over its surfaces and hefting it in my hand. As the body moves, it is also learning how to move; as it perceives, it is also learning how to perceive. Thus the body is always both active and passive, acting and exploring, doing and learning.

Although we must reject any absolute distinction between exploration and action, we may still recognize that some movements seem more exploratory than others. When I am running at full speed, there is a sense in which I am exploring the ground with my feet; but there is another sense in which I am simply taking the ground for granted, trusting that it will be there to support my next stride. When I suddenly trip, or step on something sharp, or the ground goes soft beneath my feet, my behavior changes abruptly. Where before my movements had been smooth and sure, now they become halting, tentative, exploratory—I am suddenly unsure of how to move in and with this new situation. Something has changed, but I do not yet know what. My situation is suddenly ambiguous, indeterminate; I do not know what I am seeing or touching, and I must feel my way around until I get my bearings again. Suddenly something "catches," and the situation resolves itself: I have stepped on a nail, or landed in mud, or tripped for no
3.4. Learning to perceive

reason at all. Things are familiar again, and I know what I am seeing, how to move, and what to do next.

Skillful actions are smooth and decisive, attuned to and harmonized with their situation; the body seems to know in advance what it will encounter, and to already be preparing for it. Body and world seem to be cooperating in a common project: the body finds just what it needs to support its actions, and minor fluctuations in the situation or the body are easily compensated for. When the body's skillful movement encounters an unexpected or unfamiliar situation that interrupts its comfortable operation, it becomes exploratory: the body shifts from taking the situation for granted to feeling its way around in it, unsure of how to move. In dynamic terms, we could say that the moving body finds itself hovering unstably on the edge of several basins of attraction, dipping into one and then another without falling completely into any one stable behavior. Sometimes the body ends up finding its way into an existing basin, an established sensorimotor habit that turns out to be stable in the current situation. In such cases, the result of the sudden instability and subsequent exploration is simply that the body shifts from one familiar way of moving and perceiving to another. However, if the new situation is sufficiently unfamiliar, the stability landscape of the body-world system may undergo a substantial change, such that all of the body's established motor habits are destabilized. This can be a moment of profound crisis, in that the body's movements may initially be incoherent, confused, and dangerously maladaptive. However, this can also be a moment of great opportunity, in which the body discovers new stable ways of moving and perceiving, new sensorimotor habits made possible for the first time by its new situation. At the same time that it is exploring its environment, then, the moving body is also exploring the "space"
of possible self-organizing behaviors. The organism is constantly bringing its established
sensorimotor habits to bear on new situations, which have the power to destabilize those
habits and so to give rise to new ones.

In the microgenetic studies of Thelen and Needham et al., we saw how growing bodies can use prostheses to explore their environment, and thus to develop new
sensorimotor habits. This same phenomenon is pervasive in ordinary human
development: we are constantly learning to use new prostheses that allow us to explore
the world in new ways, and thus teach us new ways of moving and perceiving. When we
first encounter a new tool—be it a pencil, a baseball bat, a car, or a telescope—we begin
by exploring it with our eyes, hands, and mouth. Depending on the possibilities it seems
to offer us, we may lift it, tap it, squeeze it, turn it over in our hands, or look at it from
different angles; we may sit on it, push it around, press its buttons, or pull its levers. In
other words, we try out various habitual ways of moving and perceiving, to see whether
and how they cohere with this new situation. As we explore the unfamiliar object, we
discover new possibilities in it: banging it makes a pleasing noise that we have never
heard before; rubbing it on a surface produces a lasting mark; swinging it amplifies our
power to hit other objects; riding in it amplifies our powers of locomotion. The tool is
initially an alien object, an unfamiliar situation that presents a challenge to our existing
habits. As we explore the object and become familiar with it, however, it ceases to be an
object of exploration, and becomes instead a means for exploring other objects and
situations. Thus the tool goes from alien object to prosthesis: it becomes an extension of
my body, an integral part of my sensorimotor habits, which I rely on for the
accomplishment of my everyday tasks.
Any part of my environment can become a prosthesis in this sense, not just the objects we commonly identify as tools. The ground is initially an unfamiliar part of the infant's environment, which she explores with her eyes, hands, and feet. But over time, as she learns to crawl, walk, and run, the ground becomes an integral part of the child's sensorimotor habits—a familiar part of her situation whose support she takes for granted, and which allows her to move around and explore other aspects of her world. This prosthetic structure is the structure of learning and behavior in general: familiar situations and established habits "scaffold" the exploration of new situations and the development of new habits. Learning to walk frees the child's hands and elevates its gaze, allowing it to explore the world in new ways, to encounter new challenges, and to develop new powers and skills. This hierarchical structure of habits built on other habits, prostheses built on other prostheses, is the source of human behavior's enormous power, flexibility, and adaptability. However, it also renders behavior susceptible to unexpected "regressions." When the ground beneath my feet becomes too steep, uneven, or unsteady to support walking, I revert to crawling on all fours—something I may not have done for years. In this archaic posture I lose other powers that I have come to take for granted, such as the ability to carry things in my hands or to see what lies far ahead of me. Similarly, when I return to my childhood home, I may find myself reverting to old ways of acting and relating to others; the familiar surroundings call forth from my body the behaviors they once enabled, even though these may no longer be appropriate.
D. Learning as differentiation

Alongside the Cognitivist view of learning as the accumulation of information about the world, the traditional behaviorist view of learning as *conditioning* continues to flourish in psychology. On this view, organisms learn to *associate* certain behaviors with certain perceptual stimuli through positive or negative reinforcement: pleasurable results strengthen the association between stimulus and response, while painful consequences weaken it. Thus the organism learns by trial and error to respond to its environment in adaptive ways: with time and experience, responses that produce positive results increase in frequency, while those that fail or actively cause harm become less frequent. For example, Rovee-Collier and colleagues interpret the results of their infant learning study as a case of "operant conditioning" (Rovee and Rovee 1969). The pleasurable stimulus of the moving mobile acts as a "conjugate reinforcement" of the kicking response—a reinforcement whose strength varies with that of the subject's response. Thus the infant learns to *associate* kicking with the pleasurable sights and sounds of the activated mobile. Because more intense kicking produces a stronger pleasurable stimulus, the infant learns to kick harder and more frequently.

The fundamental assumption of this theory of learning as conditioning is that learning consists in the *association of a discrete stimulus with a discrete response*. I want to point out three related problems with this assumption. The first problem is that it assumes what needs to be explained—namely, how a given "response" is learned in the first place. For example, Rovee-Collier and colleagues take for granted that the infant is

---

116 "Operant conditioning" is the reinforcement of a voluntary response, as opposed to "reflex conditioning" which reinforces an involuntary response such as the salivation of Pavlov's famous dogs.
already capable of kicking, and of varying the speed and frequency of these kicks. In general, however, the question facing the student of child development is not simply how the child learns to perform old behaviors in new contexts, but how the child learns to do new things which it could not do before. This is precisely the point of Thelen's microgenetic study, which demonstrates that Rovee-Collier's ingenious apparatus can be used not only to encourage existing behaviors in the infant, but also to generate qualitatively new behaviors. This genesis of new behaviors cannot be interpreted as an association of existing elements, unless we take every behavior to be composed of units that are given in advance.

This is the second problem with the association theory of learning: it takes behavior to be composed of discrete responses to discrete stimuli. But we have now seen that neither movement nor perception can be decomposed into discrete elements. Both the moving body and its nervous system are nonlinear dynamic systems. Thus there is no one-to-one correlation between individual muscle contractions and bodily movement, or between the activity of a single sensory neuron and that of the whole nervous system. The infant's kicks in Rovee-Collier's experiments are not discrete or instantaneous "responses," but self-organizing forms spread out in space and time; and they do not produce discrete or instantaneous "stimuli," but rather an ongoing pattern of changes in the infant's perceptual field. The concept of a "stimulus" in psychology is notoriously ambivalent, referring sometimes to the thing in the world that is doing the stimulating, and sometimes to the effect of that thing on the sensory surfaces of the living body.

---

117 The focus of their experiments was on infant memory, rather than learning per se. Rovee-Collier and colleagues used their apparatus to test whether infants would exhibit a learned behavior again if returned to the same context some time after the learning took place.
Unfortunately, this ambivalence tends to produce an illusion of rigor: in theory, the stimulus has a precise anatomical definition as the "firing" of a single sensory neuron; but in practice, psychologists often identify the cause of these stimuli as "the stimulus," thus avoiding the difficult task of analyzing the effects of the organism's environment on its individual nerve-endings. This slippage of meaning allows the experimenter to describe as a single, discrete stimulus what must in fact be a complex pattern of changing stimuli. It also conceals from view the crucial question of how things in the organism's environment become "stimuli" for it, i.e. how environmental differences come to make a difference to the organism and its behavior.

This brings us to the third problem with the association theory of learning: it treats the stimulus and the response as two mutually external events occurring in succession, the first causing the second. But we have learned that the relation between perception and movement is a circular one: the moving body is the cause of its own perceptions, "stimulating itself" by moving around in its asymmetrical environment. In Rovee-Collier's and Thelen's experiments, it is the ongoing feedback between movement and perception that teaches the infant to move in new ways. Thus it is impossible not only to divide movement and perception into discrete parts, but also to separate them from each other. Perception and movement are two aspects of a single ongoing process, the

---

118 This is a form of what Merleau-Ponty calls the "experience error" (PP 5/27).
119 "On analysis, the equivocal notion of stimulus separates into two: it includes and confuses the physical event as it is in itself, on the one hand, and the situation as it is 'for the organism,' on the other, with only the latter being decisive in the reactions of the animal" (SB 129/139).
120 "Situation and reaction are linked internally by their common participation in a structure in which the mode of activity proper to the organism is expressed. Hence they cannot be placed one after the other as cause and effect: they are two moments of a circular process" (SB 130/140).
121 "The mutual exteriority of the organism and the milieu is surmounted along with the mutual exteriority of the stimuli" (SB 161/174).
3.4. Learning to perceive

sensorimotor feedback loop. This process is not built up out of any minimal units: it
cannot be understood as a series of discrete events, or a linear chain of causes and effects.
Thus the very concepts of the discrete stimulus and the discrete response must be
abandoned, and with them the view of learning as association.

Mechanistic theories of behavior are obliged to conceive of learning as a synthetic
process of association because they assume that behavior is a collection of mutually
external events. All behaviors are built out of the same basic elements, and so learning
can only be a matter of putting these constant units together in more and more elaborate
ways. For these theories, the moving body is all too differentiated, with its myriad
muscles and sensory nerve endings; the problem is to understand how all of these parts
get assembled into one coordinated, responsive movement. On the view I have been
arguing for, on the contrary, the moving body begins as a relatively undifferentiated
whole, which is relatively insensitive to differences in its environment. Thus learning is
not a matter of association, but rather of differentiation (Gibson & Gibson 1955): the
development of increasingly articulated movements which respond to increasingly subtle
differences in the body's surroundings.

Psychologists have long noted how people differ in their capacities for perceptual
discrimination. William James, in his Principles of Psychology (1890), wrote:

That 'practice makes perfect' is notorious in the field of motor accomplishments. But motor accomplishments depend in part on sensory discrimination. Billiard-playing, rifle-shooting, tight-rope dancing demand the most delicate appreciation of minute disparities of sensation, as well as the power to make accurately graduated muscular response thereto. In the purely sensorial field we have the well-known virtuosity displayed by the professional buyers and testers of various kinds of goods. One man will distinguish by taste between the upper and lower half of a bottle of old Madeira. Another will recognize, by feeling the flour in a barrel, whether the wheat was grown in Iowa or Tennessee. (p. 509, my emphasis)
3.4. Learning to perceive

However, psychology's understanding of these phenomena has been held back by its mechanistic assumption that sensation is *passive*. On this view, the novice and the expert placed in the same situation must receive the very same sensations. In order to explain why they nevertheless *perceive different things*, psychologists and philosophers have long distinguished between sensation—the passive reception of stimuli—and *perception*: a second stage of intellectual operations that would synthesize the body's disconnected and ever-changing sensations into the perception of a world of stable objects (PP 34-5/56-8; Gibson & Gibson 1955). But this whole second stage becomes superfluous once we realize that sensation is not the passive reception of disconnected stimuli, but the active exploration of the body's environment. Thus in overturning the distinction between perception and movement, we have also overcome the distinction between sensation and perception. The expert and the novice move in different ways because they perceive the world differently; but they perceive the world differently because they move in different ways. In learning how to move, we are also learning how to perceive.¹²²

Consider an infant learning to explore its surroundings, or an adult in a new and unfamiliar situation, learning how to play a new sport or a new musical instrument. As novices, we start out making relatively gross and uncoordinated movements: we swing the racket wildly, or bash the guitar pick against all the strings at once. The perceptual feedback we receive from these movements is as blunt and unsubtle as the movements themselves. Large, uncoordinated movements reveal only the largest and most prominent

---

¹²² “The gaze obtains more or less from things according to the manner in which it interrogates them, in which it glances over them or rests upon them. Learning to see colors is the acquisition of a certain style of vision, a new use of one's own body” (PP 154-5/190).
3.4. Learning to perceive

differences in our situation: for example, the difference between hitting the ball or striking the strings and missing them completely. Thus we at first experience our new situation as relatively formless and undifferentiated; we seem to be moving in a void, struggling to find reference points on which to anchor our movements. But as we explore our new surroundings, our motions become more sensitive and refined, revealing smaller, more subtle differences. Our movements become more attuned to the form of our situation, and we begin to perceive it in greater detail. What started out as a formless void takes on new contours, becoming rich and inviting, and offering many avenues for new explorations. Thus learning takes place through a positive feedback loop between movement and perception: as our movements become more articulated and differentiated, we encounter ever more subtle differences in our situation, and these encounters in turn give rise to even more asymmetrical movements. The world we perceive becomes more articulated at the same time as our movements do. Thus learning is the simultaneous growth of determinacy and differentiation in the body and in its situation.

Mechanistic theories of perception suffer from a prejudice of determinacy (PP 31-33/53-55): they assume that the body's perceptions must be as determinate as the world it perceives. Thus the organism's perceptions must be composed of discrete, fully determinate parts. Learning to perceive, on this view, can only be a matter of learning to synthesize these determinate elements in the right way. The difficulty, for the beginner, is not that her perceptions are ambiguous or indeterminate, but rather that they are all too distinct from one another: she is overwhelmed by their sheer variety and unable to organize them into coherent categories. Take the wine novice, for example. On the view
that learning takes place through association, the wine novice's predicament is that every wine tastes completely different from every other one: each one offers a unique combination of many determinate flavors.\textsuperscript{124} To become expert, the novice must learn to identify the flavors that certain wines have in common, which mark them as coming from the same region or being made from the same grape varietal. In other words, she must learn which of the many flavors to pay attention to, and then to associate particular flavors with particular names. But this account of learning is at odds with the actual experience of being a beginner. The wine novice's predicament is not that every wine tastes \textit{different}, but that \textit{they all taste more or less the same}. This is why she cannot distinguish one varietal from another, or an excellent wine from a mediocre one. What confuses the novice is not how \textit{many} different flavors she can taste, but how \textit{few}. Before she can ask what various wines have in \textit{common}, she must first be able to \textit{tell them apart}; the question of classification or similarity is secondary to the recognition of \textit{multiplicity} or \textit{difference}. When beginners exclaim in despair that "There is so much to learn!", it is not the variety of their own perceptions that is daunting, but the variety of distinctions that the connoisseur expects them to absorb—distinctions that \textit{do not yet correspond to any differences in their own experience}. The expert may inform the beginner that there are so many different kinds of grape, or so many different genres of music; but these distinctions will be meaningless to the novice until she begins, through practice, to taste or hear these differences for herself. Indeed, the beginner is often skeptical of the expert's

\textsuperscript{123} More precisely, mechanistic theories assume that the body's perceptions must be as determinate as the world that the \textit{scientist} perceives; we will return to this point in the next chapter.
\textsuperscript{124} One might go even further, and say that even a single wine yields a different flavor profile with every sip.
claims, suspecting that her many fine distinctions have no basis in reality. There is a scene in the film "The Devil Wears Prada" in which the editor of a fashion magazine is trying to decide which of two green belts to use in a photo shoot. "I don't know," she says, "they're just so different." The protagonist, who is new to the world of fashion, betrays her ignorance by laughing. To her, the belts look identical, and the editor seems pompous and ridiculous. To the expert, of course, it is the novice who seems ridiculous, for she cannot see what is right in front of her. In cases such as this, there is no way for the beginner to verify the expert's claims except by entering the expert's world and gaining some expertise for herself.

I argued above that bodily movement is self-organizing, and that the purpose of the animal nervous system is to allow movement to regulate itself via perception. However, we could say with equal justification that it is the perceptual field that regulates itself via the body's movements. From this perspective, behavior and motor development appear as the self-organization of the perceptual field. Psychologists have long noted that the living body can perform the same task in a variety of ways: playing the same melody on a new instrument, or signing one's name with a different part of one's body. From the body's point of view, to carry out these tasks is to generate a certain pattern of perceptual feedback: I can play the same melody on a different instrument because I am listening to myself play; I can sign my name with my toes because I am watching myself write. The robust outcome of these various performances is not any particular sequence of movements, but the perceptual form that the movements produce. The body perceives

---

125 "[H]abits acquired by one group of muscles can be transferred immediately to another: my handwriting on the blackboard resembles my handwriting on paper although the muscles concerned in each case are not the same" (SB 30/30).
in order to move; but it also moves in order to perceive. Gilchrist and colleagues report the remarkable case of a young woman, "AI", whose eye muscles have been paralyzed since birth (Gilchrist et al. 1997). We learned above that vision requires constant "saccadic" eye movement: we see by exploring the visual texture of the world with our eyes. Yet AI exhibited little or no visual impairment: she was able to read, write, and perform visually guided motor tasks. What Gilchrist and colleagues found was that AI saw by making saccadic movements with her head instead of with her eyes. Evidently, saccadic movement is a robust outcome of human motor development—not because saccadic eye movements are "hard-wired" in the brain, but because the body learns to move by seeing, which is to say by being attuned to visual differences in its environment.

What is self-organizing in the development of vision is not the movement of the eyes *per se*, but rather the determinacy of the visual field: the body robustly discovers whatever patterns of movement will yield the most detailed and differentiated perceptions.\(^{126}\)

In a world constructed out of discrete parts, there can be no ambiguity or indeterminacy. But the world that the living body perceives is not built out of already-determinate sensations. It is an ambiguous, indeterminate whole that grows more determinate over time as the organism learn to perceive differences within it. Thus the situation that the living body perceives and that its behavior is responding to will be more or less differentiated, depending on how familiar the body is with its situation and how skillful its sensorimotor habits are. The perceived world of the organism grows along with its body and behavior, and in the same way: through a cascade of symmetry-breaking bifurcations.

\(^{126}\) "[T]he eye always places itself in such a way that it receives the richest possible stimulations from the
5. Conclusion

I have argued in this chapter that the living body's behavior, like its anatomy, is self-organizing—a form that arises dynamically through symmetry-breaking rather than being constructed out of discrete parts according to a plan. I have tried to show, contra mechanistic theories, that the living body's behavior is not determined in advance either by its nervous system or by its environment. In each case, I argued that the purported cause of behavior is also its effect. The nervous system is regulated by the body's movements, both in its real-time activity and in its growth over time. Similarly, the environment that the organism perceives, and to which its behavior responds, depends on the organism's own exploratory movements and grows over time along with them. It is the world that teaches the organism how to move—but a world that is only revealed to the organism by its own movements. Thus the organism and its world grow together dialectically, each driving the other to become more determinate through its own increasing determinacy. The self-organization of the living body is also the self-organization of its perceived world.

That which we call "the living body" is not a static thing made out of discrete parts, like a machine, but a self-organizing movement. Being, acting, and becoming are all one for the living body: what it is, what it does, and how it comes to be are one and the same. "Growth," "learning," and "behavior" are all names for this same movement; they designate not three different phenomena, but one self-organizing activity observed at three different time-scales. Thus they are each a valid answer to the question, "What is object looked at" (SB 36/36).
3.5. Conclusion

Because it is self-organizing, the movement of life is constantly generating new forms which are nowhere fully given in advance; life is a genuinely creative process whose outcomes are never completely pre-determined. This is not to say that they are arbitrary or unintelligible, but rather that life is the source of its own intelligibility: in searching for the causes of life, we have found that nothing external or indifferent to the organism can explain it. The living body is self-explanatory: not in the way of a tautology, but because it is the cause of itself, and can only be explained in terms of its own activity. This is not to say that the organism is self-contained; on the contrary, it can only be understood in relation to its situation, and to the whole history of situations it has lived through. However, these situations can themselves be understood only in relation to the organism that inhabits them. Thus our very definition of the organism must undergo a peculiar expansion, coming to encompass not only the body itself but the whole world as it is perceived by the body. What appeared to us initially as an object in our own world—a thing among things—turns out to be a whole world of its own.

The world of the other appears within my own world without being reducible to it: it is neither completely alien nor completely familiar, neither perfectly transparent nor utterly opaque. The other appears as a point of view on the same world that I inhabit, which is nevertheless not perfectly convertible with my own. Thus my own perspective on the world is revealed to me as a limited point of view, which does not exhaust the reality onto which it opens.

Up until now, we have taken up the perspective of the scientist who studies the living body from the outside, seeking a causal explanation for its anatomy and behavior.
3.5. Conclusion

However, this very study has now revealed the limitations of its own perspective, and indeed has carried us beyond them. Mechanistic theories seek the causes of behavior in the world that we, as scientists, perceive, assuming that this is "the real world." But we have discovered that the causes of the organism's behavior do not lie in the world that we perceive, but in the world that the organism perceives—a world that can only be revealed to us by its behavior. We have thus established the causal reality of meaning in nature: the world has a motor-perceptual significance for the living body, and its behavior can only be understood in terms of this significance. Moreover, this significance is not the same as that which things have for us as scientists. Thus our study of the living body has brought us to a point where we can no longer take our own perspective for granted. We are now forced to acknowledge that we ourselves are living bodies; that we are not outside of nature looking in; and that the scientist who studies behavior is one organism perceiving another. The natural world that we as scientists observe is revealed to us by our own sensorimotor habits. Thus the scientific project is a motor-perceptual project of a living body, and scientific progress is nothing other than the education of our powers of perception by nature. I will now conclude the present study by taking up this new perspective in earnest, in order to stake out the philosophical implications of the discoveries we have made in the last three chapters.
CONCLUSION: The meaning of life

The problem that motivated the present study was the problem of meaning. Ever since Descartes, the question of the place of consciousness in nature has been understood as a question of whether and how it is possible for a machine to think. Contemporary theories of consciousness claim to have overcome Cartesian dualism by showing that machines can think, provided that thinking is understood as computation: the mechanical manipulation of symbols according to fixed rules. I argued in the Introduction that these theories take the meaningfulness of their symbols for granted. They confuse the meaning things have for us as scientific observers with the meaning they have for the systems (living or artificial) that we are observing. As a result, they fail to explain how symbols become meaningful in the first place. In order to explain the origin of meaning, I proposed that we return to the phenomenon of the living body. I claimed that a careful study of the living body would demonstrate that the problem of consciousness and nature has been poorly posed. The living body is not a machine; and as long as we continue to think of nature in terms drawn from human techne, we will never be able to understand how meaning arises in nature.

In order to show that the living body is not a machine, I had to face two orthodoxies: atomism and Neo-Darwinism. Atomism is both a philosophical doctrine and a scientific research program, which holds that all apparently complex wholes can be reduced to (i.e. completely explained in terms of) their simple, indivisible parts. Neo-Darwinism is the theory in biology that evolution by natural selection produces genetic blueprints or programs for building adult organisms. In Chapter One, I argued that both
of these dogmas collapse in the face of the phenomenon of development. According to atomism, nature is made of atoms in the same way that a machine is made of parts. The central problem for this view is the problem of form: how to explain the arrangement of these atoms into the complex wholes we find in nature? Who or what plays the role for nature that the human manufacturer plays for the artifact? Nowhere is this problem more acute than in the case of the living body, because of the astonishingly complex and harmonious arrangement of its many parts. Neo-Darwinism claims to resolve this problem by showing that the living body is a machine assembled according to a genetic blueprint or program, which is itself produced by the blind forces of natural selection. However, I argued that this theory is completely incapable of accounting for the facts of development, and has persisted only by ignoring them. No pre-established plan or program could explain the growth of the organism, for the organism has no builder whose synoptic view could guide its development. The living body, unlike the machine or artifact, is not synthesized out of pre-existing parts according to a pre-established plan; it grows by division, differentiating itself into parts that organize themselves into a harmonious whole without needing to be shaped from the outside by an intelligent designer. Thus the phenomenon of development calls into question atomism's definition of form as a shape imposed on passive matter, and calls for a new concept of form that is not drawn from human techne, but rather from nature itself.

In search of this new concept of form, I turned in Chapter Two to the emerging science of complexity. There I discovered a novel account of form as asymmetry, which mathematicians define as variation under a transformation. We are accustomed to associating form and order with symmetry; however, it turns out that the most
CONCLUSION: The meaning of life

symmetrical state is one of total formlessness or uniformity in space and time. Thus what we call form is not symmetry, but asymmetry—differences in space and changes in time. The human experience of order and form is that they take work to produce and maintain. Thus we are accustomed to thinking of form as a positive quantity that must be added to things, a shape that must be imposed on indifferent or even recalcitrant materials. The materials we build things out of are forever trying to go their own way: rusting, rotting, sprouting, eroding, and just generally falling apart. But what seems, from the point of view of human artifice, to be a breakdown of order is in fact nature's way of generating new forms. Form arises in nature through the loss or breakdown of spatial and temporal symmetries.\(^{127}\)

Atomism assumes that nature is linear, and that symmetry is conserved. To say that nature is linear is to say that causes and effects are additive: a change in one part of a system produces a proportional change in the system's global state, and the system's global behavior is a simple sum of the changes in each of its parts. A linear system's apparently complex behaviors can be decomposed into simple parts, each with its own discrete cause. Thus symmetry is conserved in linear systems: a symmetrical cause will have an equally symmetrical effect, and an asymmetrical effect must have an equally asymmetrical cause. It follows that the form of a system's behavior is always given in advance in its causes and conditions, just as the form of a human artifact is fully determined in advance by its design. Form, on this view, is like matter and energy: it can neither be created nor destroyed. Thus nature is entirely predictable, for it is incapable of generating anything truly new.

\(^{127}\) See Bergson 2004 [1896], Ch. 2; see also the discussion of constraints in Morris 2004, Ch. 2.
CONCLUSION: The meaning of life

For a long time, scientists assumed that we lived in a linear world. However, we now know that this is not the case. The second law of thermodynamics states that a closed system will always evolve toward thermal equilibrium—a state of total uniformity or maximum symmetry. However, in an open system that receives constant energy inputs from the outside—such as oil in a frying pan, the Earth bathed in the light of the Sun, or a living body that absorbs nutrients from its environment—the energy input can destabilize the system's initial, more symmetrical state. At this point, the slightest external disturbance or internal fluctuation will abruptly push it into a more stable, but less symmetrical state. Dynamic systems that exhibit these sorts of abrupt, qualitative changes are nonlinear: they do not respond proportionally to changes in their parts or parameters. Nonlinear systems can be extremely robust, maintaining the same stable behavior in the face of large external fluctuations; but they can also be extremely sensitive when destabilized, such that a tiny disturbance can shift the system from one behavior into another.

These new concepts of form and morphogenesis allow us to pose the problem of embryogenesis in a new and fruitful way, and thus to resolve the paradox of development. Rather than asking how the complex form of the adult organism is assembled out of pre-existing parts, we can ask how the relatively symmetrical zygote becomes progressively less symmetrical over the course of its development. Nature does not synthesize: the living body grows not by addition, but by division. Cell division is a kind of paradigm of morphogenesis in general: the body grows by dividing itself into parts which become increasingly differentiated from one another in and through their
CONCLUSION: The meaning of life

reciprocal regulation of one another's internal dynamics. Embryogenesis takes place through a cascade of symmetry-breaking events at multiple spatial and temporal scales. Development reliably generates viable organisms not because it is guided by a prospective view of this outcome, but because each stage of development—each symmetry-breaking event—is a stable, robust behavior of the stage that precedes it.

The view that living bodies are machines has long shaped our understanding of animal and human behavior. The rise of Neo-Darwinism has brought with it a widespread genetic determinism: if our bodies are machines, then the functioning of these machines must be determined in advance by their genetic designs. My arguments against Neo-Darwinism and its view of the living body as a machine also tell, a fortiori, against all mechanistic theories of behavior. Nevertheless, I chose in Chapter Three to examine these theories of behavior in detail, in order to refute them on their own terms. I argued that Behaviorism, Cognitivism, and Connectionism share a common view of the organism as a mechanical input-output device. These theories all attempt to explain behavior by breaking it down into a collection of discrete inputs and outputs, or causes and effects. Motor development, on this view, is a synthetic process in which discrete muscle contractions are assembled into progressively more complex wholes. The increasing adaptedness of behavior to the body's surroundings can only be understood as a process of association, in which experience conditions the body to associate certain sensory inputs with certain motor outputs. By studying the empirical details of motor development, I argued that the living body's behavior cannot be decomposed into independent parts, any more than its anatomy can. The effects of individual muscle contractions on gross bodily movement are not additive; the body's movements are more
asymmetrical than their causes. Thus the moving body, like the growing body, is a nonlinear dynamic system. Indeed, the movements of anatomical development cannot be distinguished absolutely from the gross bodily movements of behavior: growth, motor development, and behavior are one continuous movement seen at three different temporal scales. What the organism is (its body) and what it does (its behavior) cannot be separated. New behaviors grow like new organs: through a cascade of symmetry-breaking events.

Since behavior is nonlinear, it cannot be decomposed into a collection of discrete causes and effects: if movement is not a collection of discrete muscle contractions, then perception is not a collection of discrete stimuli. Indeed, the very distinction between perception and movement, input and output, breaks down: unlike a machine, which passively waits to be set into motion by outside forces, the living body is always already in movement; and it is this very movement that exposes the body's sense organs to its environment. Thus perception is equally cause and effect of movement: perception regulates movement and movement regulates perception—which is to say that behavior regulates itself through the sensory feedback it receives from its surroundings. This perceptual self-regulation of movement is only possible, however, if the body's environment is asymmetrical. A uniform environment produces no changes in the body's sense organs when the body moves. Thus perception takes place through the moving body's encounter with differences in its surroundings. Precisely which differences the body encounters, however, depends on the form of its own movements: the more subtle and differentiated the organism's movements are, the more subtle the differences it discovers will be. Learning to perceive is not a matter of synthesizing discrete sensations
into meaningful wholes, or associating them with discrete motor responses. Rather, it is a matter of differentiation: learning to perceive increasingly subtle differences in one's surroundings by moving in increasingly differentiated and responsive ways. It is the world that teaches the living body how to move—but a world that the body only encounters in and through its own movements. Thus development is a dialectical process in which the organism's behavior grows together with the situations it is responding to, each driving the other to become more complex, differentiated, and determinate.

With the revelation that perception is the discovery of differences in the body's environment, we have come full circle. I argued in Chapter Two that nature is not composed of atoms, but of differences; and that the real project of science is not to explain how nature is put together of out pre-existing parts, but to explain why there is difference rather than indifference: why things are different in one place than another, and why things are different now than they were before. We have now discovered that perception is precisely the discovery of differences in nature, or asymmetries in our surroundings. The definition of asymmetry that I offered in Chapter Two—variation under a transformation—was a mathematician's definition, and therefore an abstract one: it said nothing about how to discover which of the infinite possible transformations would reveal a particular system's asymmetries. But we have now discovered the roots of this abstract definition in embodied perception: bodily movements are the original "transformations," which discover asymmetries in the body's surroundings by producing "variations" in its own sense organs. In speaking of asymmetries or differences, then, we have always already installed ourselves at the level of perception. We cannot conceive of
CONCLUSION: The meaning of life

a difference in nature except by reference (implicit or explicit) to a bodily movement that would reveal this difference.\textsuperscript{128} This is not to say that the living body creates the differences it discovers in nature; on the contrary, as we have seen, it is nature's self-differentiation that creates the living body. However, it is to say that we cannot give an account of nature that is not an embodied account, that does not take up the point of view of a moving body situated within the nature it describes.

This may seem like a disappointing or even a skeptical conclusion. In saying that we can only know nature from our own finite perspective, are we not admitting that we cannot know nature at all? We have become accustomed to atomism's promise of a "view from nowhere"—a non-perspectival account of the natural world. But in fact, atomism takes up a very particular perspective on reality: that of a creator contemplating her creation. Though it long ago ceased to appeal to God as an explicit hypothesis, atomistic science continues to appeal implicitly to a God's-eye-view of the cosmos. It thus remains "metaphysical" in the Heideggerian sense: the truth of this world lies elsewhere; we can understand nature only by transcending it. Atomism claims to strip nature of all anthropological predicates in order to arrive at an account of reality as it exists "in itself." But in fact, nothing could be more anthropological than this way of describing nature as if human beings had manufactured it. The meaning of a manufactured thing lies not in the thing itself, but in the mind of its creator. The parts of an artifact are strictly indifferent to one another; it is only in the human mind that these parts are constituted as a unified whole. If nature is an artifact—a blind mechanism, atoms in a void—then it has no

\textsuperscript{128} In this sense, Merleau-Ponty is right to say that physical form is "conceivable only as an object of perception" (SB 144/156).
meaning of its own. The meaningful wholes that we perceive in nature are our own projections, products of the mind's synthetic activity. Thus materialism turns into its supposed opposite: idealism is simply a self-conscious atomism that has made its own standpoint explicit. The idea of a constituting consciousness for whom all of nature would be a single, transparent idea is a direct corollary of the view that nature is a machine. If there is no room for consciousness within Descartes' view of nature, that is only because Descartes' nature is already an idea for a pure consciousness. The Cartesian mind is no less mechanical than the Cartesian body; they arise from the same ontology, and neither can exist without the other. Thus materialism is not a true monism, but only a half-hearted dualism; it can only be maintained so long as the scientist's own perspective remains implicit, taken for granted. As soon as this perspective becomes explicit, we find ourselves inexorably committed to a transcendental idealism. But this idealism is no less mechanistic than the materialism it replaces. The transcendental turn cannot consist in this move from materialism to idealism, which really changes nothing. It can only consist in a rejection of the ontology that underlies this false choice.

A nature that can only be known from the outside cannot truly be known at all, but only mastered and controlled. It has no meaning of its own, and so it can only have a meaning imposed upon it. To reject this ontology is to affirm that nature has its own endogenous meaning which is not constituted by consciousness. It is precisely this nascent meaning that we have discovered in the phenomenon of form. Form is neither thing nor idea, neither atom nor artifact. It cannot be known by a disembodied mind, but only perceived by a living body. The perception of form is not a matter of representation; it is not a question of duplicating nature's forms within a consciousness. Perception arises
CONCLUSION: The meaning of life

within nature, through nature's own self-differentiation into an inside and an outside, a living body and its environment, perceiver and perceived. Nature is a creative power, a continuous process of morphogenesis; it is a whole that complicates itself of its own accord into different parts or regions, creating form out of uniformity, difference out of indifference. The appearance of life in nature was a fundamental ontological mutation, a singular event in the history of form. In the behavior of the living body, we witness form folding back upon itself to generate a new kind of form, which I have called motor-perceptual significance. Behavior is a natural form that develops in and through its encounters with other natural forms—a whole whose asymmetries reflect the history of these encounters, and thus is oriented toward a future. The world has a significance for the organism which is not an idea, but a developed way of responding differentially to differences in the world. Thus behavior is a form that points beyond itself to a meaningful world. The organism's behavior reveals to us the differences that have come to make a difference to it—differences that we may never have perceived before.

We deliberately began our investigation from the point of view of the scientist, for whom the living body is just another natural phenomenon—albeit an especially complex one—that calls for scientific explanation. However, this phenomenon has now turned our investigation inside out. Our study of the living body has compelled us to make a transcendental turn: to recognize that we cannot simply study this natural phenomenon from the outside. We began by assuming, naturally enough, that the world we perceive is "the real world," and that the causes of behavior must be found in this world. What we discovered, however, is that behavior can only be understood as a response to differences that it itself reveals—differences that may be entirely new to us. Thus the phenomenon of
CONCLUSION: The meaning of life

the living body reveals the limited and perspectival character of our own scientific standpoint. We ourselves are living bodies; the scientist who studies perception and behavior is one organism perceiving another. This means that the results of our investigations rebound upon us as investigators. We can no longer take our own perspective for granted; we must explain not only the behavior and perception of the other, but also our own perception of this behavior. That is, we must explain how the meaning of one organism's behavior can be perceived by another.

We could have avoided a great deal of trouble by simply pointing this out at the outset, and thus beginning with the question of how life appears to life. But we would not then have encountered the problem of form, or the paradox of development. We would not have discovered the extent to which our understanding of life and nature is impregnated with artifice. And we would not have been compelled to search for a new account of form which was not drawn from human *techne*. Moreover, had we simply begun from the point of view at which we have now arrived with such difficulty, we would have had to place ourselves in opposition to the scientific point of view, arguing on *a priori* grounds that the living body challenges or escapes the scientist's "naturalistic attitude", and that the scientist is "naïve" because she takes her own point of view for granted. Such arguments have little purchase in scientific discourse. They criticize science from the outside, juxtaposing the philosophical attitude against that of the

---

129 "The object of the preceding chapters was not only to establish that behavior is irreducible to its alleged parts. If we had had nothing other in view, instead of this long inductive research—which can never even be finished, since behaviorism can always invent other mechanical models with regard to which the discussion will have to recommenced—a moment of reflection would have provided us with a certitude in principle…But by following this short route we would have missed the essential feature of the phenomenon, the paradox which is constitutive of it: behavior is not a thing, but neither is it an idea. It is not the envelope of a pure consciousness and, as the witness of behavior, I am not a pure consciousness. It is precisely this which we wanted to say in stating that behavior is a form." (SB 127/138)
CONCLUSION: The meaning of life

scientist; thus they cannot hope to persuade the scientist, and do not offer her any resources for pursuing her own projects in a different way. By beginning from within the universe of science, we have been able to criticize it on its own terms. We have discovered that the transcendental turn—the move from consciousness to self-consciousness—does not need to be imposed on science from without, but arises within the scientific project as an internal necessity. And we have seen, moreover, that what this turn requires is not simply a change in attitude, but a change in ontology. Without this ontological mutation, our purportedly transcendental philosophy will remain infected with artificial concepts. Thus we began from the scientific perspective not only to criticize it, but also to learn from it what nature is, what life is, and what we are. Self-consciousness cannot be achieved all at once, or once and for all. The only legitimately transcendental philosophy is the phenomenological one: the empirical, dialectical investigation that allows itself to be called into question by its own results, even to the point of starting all over again.

If nature is a creative power, then it can never be known once and for all. The scientific project can never be completed, because nature is always changing, always becoming something new. We would have to give up the dream of mastering and controlling nature; but what we gain in return is the possibility of actually knowing nature, of learning new things and being transformed by this knowledge. The scientist who seeks a causal explanation for the complex forms she observes in nature is engaged in a perceptual project. Her goal is to discover which differences make a difference to the phenomenon she is studying; but this will often mean learning to perceive differences that had been invisible to her, and to ignore differences that had seemed to be of great
CONCLUSION: The meaning of life

significance. The goal of science is not to discover the real world behind the world that we perceive, but to allow nature to educate our powers of perception, so that we may perceive the natural world in more determinate and insightful ways. This is precisely what atomism fails to do: it has decided in advance how nature is to be divided, in terms that are drawn from human techne rather than from the observation of nature itself. It is thus a perceptual stance that refuses to be educated by the world that it perceives.

If I am myself a living body, then what does it mean for me to study the behavior of another organism? How is it possible for me perceive another perceiving, to see the world through another's eyes? How can the world of the other appear within my own? If perception were a mental representation, then the perceptions of the other would be forever hidden from my gaze, which sees only the movements of the other's body. But we now know that perception is nothing but the body's way of moving and exploring its environment. Thus perception does not take place within a private mental realm, but publically in the world. This is not to say that the meaning of the other's movements is immediately apparent to me; it is neither completely transparent nor completely opaque. To understand the behavior of another takes work. I can learn to perceive the world through another's eyes—but only by allowing the behavior of the other to teach me how to move and perceive in a new way.

In fact, this is precisely how we learn to move and perceive in the first place: the first objects of perception for the human infant are not inanimate objects, but other human beings. The infant's powers of mobility—and thus its capacity to explore its environment—are at first extremely limited. The most responsive parts of the infant's surroundings are its caregivers, who respond to the infant's cries and facial movements
with sounds and facial movements of their own. Thus newborns are primarily interested in the faces of others, and only become interested in inanimate things around 4-6 months of age (Poulin-Dubois 1999). The adult is a kind of first prosthesis for the infant, providing her with the kind of rich and responsive sensory feedback that the rest of her environment cannot yet offer. As the child grows, her caregivers continue to act as prostheses, responding to her cries with food and physical contact, helping her to explore her surroundings, interpreting her desires and attempting to fulfill them. Thus human beings learn to move and perceive through our relations to others, who "scaffold" our behaviors and interpret the meaning of our gestures. Through the child's relation to the other, the meaning of her behavior can shape the development of that behavior, long before it has a meaning for the child herself. The child's relation to her adult caregivers is thus a kind of pre-reflective reflection, prior to the child's learning to reflect on the meaning of her own actions.

The child does not initially encounter the other as a problem, as a challenge to her own point of view; indeed, she does not encounter the other as an other at all, but as an object of perception, a prosthesis, an extension of her own body. It takes around 9 months for the typical child to become aware of the perspectives of others on inanimate things. Until the age of four, most children are incapable of recognizing that other people may not know everything that they themselves know (Poulin-Dubois 1999). To construe the child as a self in relation to others and a mind in relation to its own body would be to confuse once again the meaning things have for us as observers with the meaning things have for the child, and thus to foreclose any possibility of understanding how things come to have the meanings for us that they do. The problem of consciousness is precisely to
understand how the living body becomes a self-conscious self; how the anonymous, pre-reflective relation between the living body and its surroundings can differentiate itself into self and other, mind and body.

The problems of consciousness and symbolic meaning with which we began have not been answered; we have only learned to pose them in a better way. Rather than asking how a machine can think, or how a disembodied mind can synthesize an objective world out of disconnected sensations, we can now ask how nature becomes conscious of itself, how it differentiates itself into knower and known. By studying the development of the living body, we have borne witness to nature's division into perceiver and perceived. In perception, form is folded back upon itself, giving rise to a new kind of motor-perceptual significance. Symbolic meaning must arise, in a parallel way, from the folding back of behavior upon itself in reflection: behavior that thematizes its own motor-perceptual significance. A self-conscious self is not a disembodied mind, but a body for whom the meaning of its own behavior has become an issue. Consciousness, on this view, is itself a form of behavior: not something we have, but something we do. Thus we ought to be able to study it in the same way we that we studied anatomy and behavior: by observing its development. If we follow the gradual development of a single cell into an adult self, we may be able to observe the cascade of symmetry-breaking bifurcations that give rise to subjectivity, and thus catch life in the act of becoming conscious.
APPENDIX A.
Dynamic systems theory:
A different kind of scientific model

The purpose of this appendix is to offer an introduction to the basic concepts and approach of dynamic systems theory (DST), for the reader who wants more mathematical detail than Chapter Two provides. We will begin by looking at the origins of DST in classical physics. Then we will study some of the basic concepts of DST. Finally, I will develop the implications of these concepts for the present project.

1. Origins of Dynamic Systems Theory

Modern science is characterized by its method of constructing ideal models of natural phenomena on the basis of experiment. These models are typically expressed in the language of mathematics. The natural phenomenon is reduced to a set of variables which together describe the state of the phenomenon at a given time; the model consists of a mathematical description of how these different variables depend on one another, i.e. how they vary with respect to one another. The purpose of experiment is to determine these quantitative relationships by varying a single aspect of the phenomenon (corresponding to a single variable in the model) in a controlled way, and measuring the effect of this variation on other aspects of the phenomenon (represented by other variables). Of special interest to Modern science are models that describe the dynamics of a phenomenon, i.e. how it changes over time.
The first great mathematical model of a natural phenomenon was the pre-Modern Ptolemaic model of the solar system, which could successfully predict the positions of the celestial bodies with great accuracy. Astronomy, however, does not permit experiments, but only observations. The first truly Modern scientific model—which could justifiably be said to have inaugurated Modern science—was Galileo's mathematical model of free fall (Galileo 1638/2000). In order to familiarize ourselves with the practice of modeling, let us construct a mathematical model of a small body falling freely near the surface of the Earth.\textsuperscript{130} To keep this model simple we will neglect air friction and assume that the body is only moving in one dimension (straight down).\textsuperscript{131}

The first step is to choose variables that describe the relevant features of the natural phenomenon. In this case, we are interested in how the falling body moves over time. We can describe this in terms of three variables: the distance \((x)\) travelled by the falling body; the speed \((s)\) at which it is falling; and its acceleration \((a)\), the rate at which its speed is increasing or decreasing. Now, we are interested in how these variables change over time, or the dynamics of the system, so we will need a fourth variable \(t\) which represents the time elapsed since the start of the body's fall. What Galileo discovered, by careful experiment, is that the vertical distance travelled by a falling body is proportional to the square of the time elapsed. In algebraic terms, this relationship can be expressed as

\textsuperscript{130}Galileo's mathematical methods are of considerable interest: the tools of algebra and calculus were not yet available; instead, Galileo proceeds by geometrical proofs, relying heavily on the ancient theory of proportion. For simplicity of exposition, however, I will present a model of free fall in modern algebraic terms.

\textsuperscript{131}In what follows, I have tried not to assume any mathematical knowledge on the part of the reader beyond basic algebra. We will need to employ one concept from calculus, that of the derivative; but all the reader needs to know is that a derivative represents the rate of change of one variable with respect to another. Most of the derivatives we will encounter will represent rates of change with respect to time.
\[ x = kt^2 \]

where \( k \) is a constant whose value has been determined experimentally to be approximately 5 m/s\(^2\). This equation states that the distance (in meters) travelled by a falling body near the Earth's surface is equal to the square of the time elapsed (in seconds) multiplied by five. We can plot this equation on a graph, showing how the position of the falling body varies over time. (Since \( x \) measures displacement from some initial position, I have chosen an arbitrary initial height of 500 m and plotted the height \( (h) \) of the body versus time.)

\[ h (\text{m}) = 500 - x \]

It can be shown mathematically, on the basis of this relationship between distance and time, that the speed of the falling body is proportional to the time elapsed, or

\[ s = gt \]

where \( g \) is a constant whose value is approximately 10 m/s\(^2\).\(^{132}\)

\[^{132}\text{ Using basic calculus, it is easy to show that the speed } s, \text{ which is the rate of change of } x \text{ with respect to } t, \text{ is } \frac{dx}{dt} = 2kt.\]
And it can also be shown that the acceleration of the falling body is constant. In fact, it is equal to the constant $g$:

\[ a = g \approx 10 \text{ m/s}^2 \]
APPENDIX A

(One of the surprising results of Galileo's investigations was that this constant rate of acceleration is the same for all falling bodies near the Earth's surface, regardless of their mass.\textsuperscript{133})

Equations (1)-(3) describe how our three variables (distance, speed, and acceleration) vary as functions of time. (If we wanted to know how speed varies as a function of distance, or any other such combination, this could be found with a bit of simple algebra.) These equations can be expressed as a system of differential equations: a set of equations describing a function and its derivatives (rates of change) of various orders. In this case, distance \((x)\) is a function of time; speed \((s)\) is the derivative or rate of change of \(x\) with respect to time; and acceleration \((a)\) is the rate of change of \(s\) with respect to time, or the second-order derivative (the rate of change of the rate of change) of \(x\). In modern calculus notation, we can express these relationships as follows:

\[
\begin{align*}
(4) \quad s &= \frac{dx}{dt} = x' \\
(5) \quad a &= \frac{ds}{dt} = \frac{d^2x}{dt^2} = x''
\end{align*}
\]

(Reading the final symbol in equation (4) as "x-prime" and the final symbol in (5) as "x-double-prime.") We have just constructed a simple model of a falling body as a dynamic system.\textsuperscript{134}

Isaac Newton's great achievement was to see the parallel between the small body falling freely near the Earth's surface and the orbit of the Earth around the Sun, and to

\textsuperscript{133} To be precise and to avoid anachronism, we should note that Galileo did not have any concept of "mass"—he would have used the term "weight."

\textsuperscript{134} Note the ambiguity in the term "system" here: it can refer either to the physical phenomenon being modeled, or (as in this case) to the model itself, (the "system" of differential equations). We must be careful not to confuse the two.
unite these two apparently disparate phenomena under the single concept of gravity.

Newton showed that if we lift the small body up above the Earth's atmosphere (to avoid the drag of air friction), and give it a sufficient horizontal velocity, its free fall will take it in a stable orbit around the Earth. Thus vertical free fall near the Earth's surface and the orbit of one body around another are two particular cases of a more general phenomenon, the gravitational attraction between two bodies. Newton's model of this attraction is captured in the following equation:

\[ F = \frac{G m_1 m_2}{r^2} \]

where \( m_1 \) represents the mass of the first body, \( m_2 \) the mass of the second, \( r \) the distance between their respective centers of mass, \( G \) the gravitational constant (determined by experiment to be approximately \( 6.7 \times 10^{-11} \text{ N m}^2/\text{kg}^2 \)), and \( F \) the force of attraction acting on each body.\(^{135}\) This equation states that every particle in the universe attracts every other particle with a force that is proportional to both their masses and inversely proportional to the square of the distance between them. In order to model the movement of two bodies that are attracted to each other gravitationally, we need to combine equation (7) with Newton's Second Law of Motion,

\[ a = \frac{F}{m} \]

which states that the acceleration produced by a force acting on a body is proportional to the magnitude of the force and inversely proportional to the body's mass. From these two equations together we can deduce Galileo's model of free fall, as well as Kepler's (empirically verified) model of planetary orbits around the Sun. The model of planetary

\(^{135}\) In fact there are two equal and opposite forces at work, \( F_1 \) and \( F_2 \), each acting on one of the two bodies in the direction of the other, and both of magnitude \( F \).
motion is more complex than our model of free fall near the Earth's surface, since it involves the movement of two bodies in two dimensions, rather than one body in one dimension. However, it has the same basic form, in that the positions, velocities and accelerations of each body can be represented as functions of time.

It is easy to see how Newton's accomplishment led to Laplace's view of a mechanistic universe (1814/1951):

Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes. (p. 4, translation modified)

The Newtonian model made it possible to find the positions, velocities, and accelerations of two point-masses at any time in the future or past, given their present state. Laplace was simply generalizing from Newton's model of a two-body system to nature as a whole, conceived of as a system of many mass-points. Laplace takes a purely additive view of nature: he is imagining the cosmos as a simple sum of some vast number of two-body systems, each of which can be understood in isolation. Force in Newton's model is additive: to find the total gravitational forces acting on any given body in a system of \( n \) bodies, one can simply add up the gravitational force contributed by every other body in the system (according to equation (7)). Having found the total force acting on the body, one can find its acceleration using equation (8); and from this, it would seem, one ought to be able to predict the future position of the body.

\[136\] The motion of the two bodies with respect to each other always takes place in a plane, and so can be treated as two dimensional.
In one of the great surprises of Modern science, this turned out not to be the case. Laplace's assumption that Newton's solution of the two-body problem could be extended to systems of \( n \) bodies proved to be false—and not only for high values of \( n \). On the contrary, we now know that systems of just three bodies exhibit complexities of an entirely different order than two-body systems. As a result, it is impossible to "solve" the three-body problem, i.e. to find equations for all the variables of the system as functions of time, which would make the system entirely predictable. This surprising result was proved by the great French mathematician Henri Poincaré in 1890. This discovery spurred him to invent the new mathematical methods which form the core of what is now called dynamic systems theory. We now know that the difficulties presented by the three-body problem are not unique. On the contrary, the vast majority of dynamic systems (i.e. systems of differential equations with time as the independent variable) cannot be solved as functions of time.\(^{137}\) It may be hard to imagine how this can be the case on the basis of our simple falling-body example, so let's look at another model which exhibits more complex behavior. One might think that the way to make a system more complex would be to add more variables. However, we will see that it is not the number of variables that determines whether a dynamic system can be solved for time—rather, it is the way these variables are related to one another.

Imagine that the falling body we discussed above is hanging from a spring. Now in addition to the force of gravity, the falling body is also subject to a force in the opposite direction, whose magnitude depends on how far the spring has been stretched.

\(^{137}\) More generally, we can say that only a small subset of differential equations have a closed-form or
So instead of a system in which acceleration is constant, as in the free-fall case, we now have a system in which the body's acceleration depends on its position. The number of variables we need to describe this system remains the same, but our description is nevertheless more complex, because of the way these variables are interrelated: the position $x$ is changing at rate $v$ (for *velocity*), which is itself changing at rate $a$ (for *acceleration*), which is in turn changing according to $x$. We thus have a *circular* or *auto-regulatory* system, a *feedback loop* which can be described by the following differential equation:

\[ ma + kx = 0 \]  

where $m$ represents the mass of the body, $a$ the acceleration, $k$ the spring constant (which measures the spring's resistance to being stretched or compressed), and $x$ the displacement of the body from the system's equilibrium point (the point where the force of gravity and the resistance of the spring are in balance). This equation can also be rewritten (through simple algebraic manipulation) as

\[ a = v' = x'' = - \left( \frac{k}{m} \right) x \]

Notice that time does not appear explicitly in this equation; it is present only implicitly, in the fact that $a = v' = \frac{dv}{dt}$, i.e. that $a$ represents the rate of change of $v$ *with respect to time*. We would like to be able to express $a$, $v$, and $x$ as functions of the independent variable $t$, so as to be able to find the state of the system at any given time. But it is not immediately obvious how to do that; all we know about the system is how $a$ varies as a function of $x$. 

---

*analytical solution.*
As it happens, this dynamic system can be solved analytically.\textsuperscript{138} It belongs to a class of systems known as \textit{linear differential equations}, which all have analytical solutions. However, most systems of differential equations are \textit{nonlinear} ones, which can almost never be solved analytically; and most of the phenomena we encounter in nature can only be modeled by nonlinear equations. Strogatz (1994, pp. 8-9) explains the difference this way:

Why are nonlinear systems so much harder to analyze than linear ones? The essential difference is that linear systems can be broken down into parts. Then each part can be solved separately and finally recombined to get the answer. This idea allows a fantastic simplification of complex problems, and underlies such methods as normal modes, Laplace transforms, superposition arguments, and Fourier analysis. In this sense, a linear system is precisely equal to the sum of its parts. But many things in nature don't act this way. Whenever parts of a system interfere, or cooperate, or compete, there are nonlinear interactions going on.

Although most nonlinear systems cannot be solved analytically, they can be \textit{simulated} using methods of numerical approximation. The simulation begins with some set of initial conditions, and uses the differential equations of the model (which give the instantaneous rates of change for the system's variables) to calculate the approximate state of the system a short time later. (These values will be only approximate because the simulation cannot take into account how the system's rates of change have themselves changed over this short span of time.) The new system state can then be plugged back into the differential equations to make another short step forward in time. This process can be repeated indefinitely. However, the further forward in time the simulation goes, the less accurate its predictions will be, because the error in each approximation is compounded in the next one.

\footnote{The general solution for $x$ in terms of $t$ is $x = x_0 \cos(\omega t) + \omega v_0 \sin(\omega t)$, where $x_0$ and $v_0$ are the initial}
Predicting the future of nonlinear dynamic systems is made even more difficult by their sensitivity to initial conditions, (a phenomenon made famous under the name of "Chaos Theory"). In a linear system, a small change in the system's initial conditions produces a proportionally small change in the state of the system at some future time. In nonlinear systems, on the other hand, a small change in the system's initial conditions can set the system onto a very different path, resulting in a disproportionately large change in the future state of the system. (This is the phenomenon captured by the famous image of a butterfly flapping its wings and causing a hurricane on the other side of the planet.) This sensitivity to initial conditions has important implications for our practice of modeling natural phenomena mathematically. Our mathematical models never describe the world precisely; they are always approximations. We generally assume that the approximate nature of our measurements is not a problem, since a small error in our approximations will produce only a small error in our results or predictions. If this error turns out to be too large, we can always go back and try to make our initial measurements more precise. However, this assumption only holds for linear systems. In a nonlinear system, a small error in our initial measurements may produce a disproportionately large error in our predictions. What's more, this problem may not be ameliorated by going back and making the initial measurements more precise; an increase in the precision of our initial measurements is not guaranteed to produce a proportional increase in the accuracy of the model's predictions.

The mechanical understanding of nature gained widespread acceptance in no small part because of the success of Newton and his successors in modeling natural position and velocity of the body, and $\omega = (k / m)^2$. 
phenomena as linear dynamic systems. It was not until Poincaré's pioneering work near the end of the 19th C. that scientists began to realize the limitations of this approach, and to develop new concepts and techniques for understanding nonlinear systems. We now turn to a study of these concepts.

2. Basic concepts of Dynamic Systems Theory

Faced with the impossibility of finding an analytical or quantitative solution to the three-body problem in astrophysics, Poincaré pioneered a new, qualitative approach to the study of nonlinear differential equations (Hirsch 1984, 18). To see how this approach works, let's return to our model above of a mass hanging from a spring, its motion described by the equation

\[ a = - \left( \frac{k}{m} \right)x \]  

We saw that this equation can be solved analytically in terms of \( t \); but that does not mean that we cannot apply Poincaré's qualitative methods to it. Indeed, the reader may find that these methods offer more insight into the system than does the analytical solution given above. We would like to be able to visualize the motion of the body through space; but we have no way of doing this, since we do not know how its position varies with respect to time. All we know about this system is the relation between the acceleration of the body and its position. Is there some way to use this information to visualize the dynamics of the system?

Doing so requires a difficult conceptual shift, one which is absolutely central to dynamic systems theory. We must move from modeling the motion of bodies in space to modeling changes in the state of the system as a kind of "motion" through "state space"
APPENDIX A

(also known as "phase space"). Instead of the familiar spatial dimensions of height, breadth, and depth, we can treat each of the system's variables as a dimension. These dimensions together constitute a space of all possible system states, and each point in this space represents one such state (i.e. one determinate value for each of the variables that describe the system). This conceptual move will allow us to apply to the dynamics of the system all of the geometrical tools developed for modeling the motions of bodies in space.

In this case, we can think of our system as having two "dimensions": position \((x)\) and velocity \((v)\). The state space of this system can then be visualized as a plane with two axes, and the dynamics of the system can be plotted as a movement in this plane, much as we plotted the movement of a falling body above in the \(x\) vs. \(t\) plane.

\[
\begin{align*}
\vec{n} \\
\vec{x}
\end{align*}
\]

But how can we plot the relation between \(x\) and \(v\)? We do not have an equation that expresses \(v\) in terms of \(x\). But we \textit{do} know something about the \textit{rates of change} of \(v\) and \(x\). The velocity \(v\) is itself the rate of change of the position \(x\), by definition. And the rate of change of \(v\) is the acceleration \(a\), whose relation to \(x\) is given by equation (11) above.

In algebraic terms,

\[
\begin{align*}
(12) & \quad x' = v \\
(13) & \quad v' = -\left(\frac{k}{m}\right)x
\end{align*}
\]
These equations allow us to plot a vector field on the $xv$ plane, which indicates, for any point on the plane, what direction (in state space) the system is moving in, and how quickly. (Recall that movement in the state space represents a change in the state of the system.)

At the origin, where $x$ and $v$ are both equal to 0, $x' = 0$ (by equation (12)) and $v' = 0$ (by equation (13)). Along the $x$-axis, where $v = 0, x' = v = 0$, while $v'$ is negative for positive $x$ and positive for negative $x$ (because of the minus sign in equation (13)). Thus the vectors along the $x$-axis point straight up when $x$ is negative and straight down when $x$ is positive, and they get longer as $x$ gets larger in magnitude. Along the $v$-axis, where $x = 0, v' = 0$ and $x' = v$. Thus the vectors point horizontally to the right where $v$ is positive, and left where $v$ is negative, with the size of the vectors increasing as $v$ increases in magnitude. To find the vector at any other point in the field, simply find the vector on the $v$-axis corresponding to the point's $v$-value, do the same thing on the $x$-axis for the point's $x$-value, and add these two vectors together. Doing this will yield the plot shown above.

This vector field gives us a way of visualizing the dynamics of our system. For any point on the plane, representing an initial position and velocity, this field will produce a trajectory in state space, a set of states that the system will pass through over time. You can think of the vector field as a fluid, and its trajectories as steadily flowing
APPENDIX A

currents; an imaginary particle placed anywhere in this field will be carried along by these currents. The trajectories of our vector field look like this:

This is called the *phase portrait* of the system: it gives an overall picture of the various trajectories possible in this system. The phase portrait does not allow us to predict the state of the system at a given time, the way a *quantitative* solution would; but it does reveal certain *qualitative* features of the system's dynamic. The origin (the point \( \{0,0\} \)) is special: a particle placed here will remain motionless. This is called a *fixed point*. A particle placed anywhere else will circulate around the origin, eventually returning to where it began. Such cyclical or *periodic* trajectories are called *closed orbits*.

What does this phase portrait tell us about the physical system of a mass hanging from a spring? The fixed point at the origin corresponds to static equilibrium of the system: if the mass begins at rest \( (v = 0) \) in its equilibrium position \( (x = 0) \), then it will not move from this position. The closed orbits correspond to *periodic motions* or *oscillations* of the hanging body. (Indeed, this system is what physicists call a *simple harmonic oscillator*.\) We can see from the phase portrait that the velocity of the oscillating body is 0 where the body reaches its maximum displacement in either direction; this is where the body is about to change direction. And we can also see that it reaches its maximum
velocity when \( x = 0 \), i.e. when the body is passing through the spring's equilibrium point, which is also the mid-point of the oscillation.\(^\text{139}\)

Notice that the phase portrait does not merely show us a single trajectory of the system, but \textit{all possible trajectories}. Each of these trajectories corresponds to a certain set of \textit{initial conditions} of the system. For any initial \( x \) and \( v \), the phase space gives a particular trajectory describing the path the system will follow, starting from that initial state. Thus it allows us to study the dynamics of the system, not as a function of time, but \textit{as a function of its initial conditions}.\(^\text{140}\) This allows us to answer qualitative questions

---

\(^{139}\) If we were to plot the movement of this oscillating mass in terms of position over time (using the analytic solution to the differential equation) we would see that it forms a sinusoidal wave.

\(^{140}\) “Earlier analysts, with few exceptions, had studied individual solutions in isolation. It was Poincaré who systematically studied the mutual relations between all the solutions. In particular, … he made the
about the *stability* of the system: Are all of its trajectories *periodic*, (i.e. do they trace out the same path over and over)? How will a trajectory be affected by small changes to its initial conditions?\textsuperscript{141} We can see from this system's phase portrait that all of its trajectories are periodic, outside of the fixed point at the origin. However, none of these trajectories—including the fixed point—are *stable*: any random *fluctuation* or small change in the state of the system will shift the system out of one trajectory and into another. In physical terms, this means that if some external force acts on the oscillating mass—increasing or decreasing its velocity, say—it will be "knocked" into a new oscillation with a different amplitude and maximum velocity.

The question of stability under small fluctuations is crucial in modeling natural phenomena, for two reasons. One is that the systems we are modeling are never completely isolated—there are always external forces acting on the system and introducing small perturbations. Thus the question of how the system's different trajectories respond to fluctuations is of great importance in making realistic *qualitative* predictions about the system's behavior. Second, as we noted above, our measurements of a system's initial conditions are always approximate. Thus, when it comes to making *quantitative* predictions (using simulations), it is important to know how sensitive a given trajectory is to small differences in its initial conditions. This lets us know when a small

\textsuperscript{141} Questions about the stability of the solar system motivated Poincaré's original development of these methods. "If there is only one planet and the sun, then Newton's equations of motion are easy to solve explicitly; the orbits are the Keplerian ellipses, and all questions are easily answered. When there are two or more planets the situation is quite obscure. Very little is known about such questions as these: will two of the planets ever collide? Do the mutual distances stay bounded away from zero and infinity? Is it probable that a planet can escape to infinity, or be captured from infinity? If we slightly alter the mass of a planet, how does that change the trajectories?" (Hirsch 1984, 20).
error in our initial measurements may result in a large discrepancy between our predictions and the system's future state.

The concept of stability will be more clear if we consider a further example. We have been pretending, for simplicity's sake, that our mass hanging from a spring is not subject to friction. But now let us modify our model to take friction into account, turning it from a simple harmonic oscillator into a damped oscillator.\(^{142}\) This new model has the following phase portrait:

\[ a = -\left(\frac{k}{m}\right)x - hv, \]
\[ x' = v, \quad v' = -\left(\frac{k}{m}\right)x - hv. \]

Whereas the frictionless system is described by a differential equation of the general form \(ax'' + bx = 0\), the frictional system's equation introduces a first-order derivative term, and has the form \(ax'' + bx' + x = 0\).

\(^{142}\) We can do this by assuming that the force of friction acting on the hanging body is proportional to its velocity. Friction acts as a drag force, reducing the body's acceleration. Thus equation (11) above becomes \(a = -\left(\frac{k}{m}\right)x - hv\), where \(h\) is some constant coefficient. Equations (12) and (13) then become \(x' = v, v' = -\left(\frac{k}{m}\right)x - hv\). Whereas the frictionless system is described by a differential equation of the general form \(ax'' + bx = 0\), the frictional system's equation introduces a first-order derivative term, and has the form \(ax'' + bx' + x = 0\).
mass eventually settles down to rest at the static equilibrium point \( x = 0 \). In the previous system, the fixed point at the origin was \textit{unstable}: any little fluctuation would set the hanging mass into motion, and this motion would continue forever unless some later fluctuation cancelled it out. In the \textit{damped} oscillator, on the other hand, the fixed point at the origin is \textit{stable}: if some fluctuation or external force sets the system into motion, it will not continue to oscillate forever, but rather return to rest at the fixed point. Such a stable fixed point is also known as an \textit{attracting fixed point}, since all nearby points and trajectories get "pulled into it." Once again, this phase portrait does not answer the quantitative question of \textit{how long} it will take for the system to settled down at rest; but it does give us the \textit{qualitative} prediction that the system will always end up at static equilibrium in the long term.

So far, for the sake of simplicity, we have been considering linear systems. However, the real beauty of Poincaré's qualitative approach can be seen when we turn to more interesting, nonlinear systems. Consider the system with the following phase portrait:\textsuperscript{144}

\textsuperscript{144} This phase portrait can be generated by the following pair of (nonlinear) differential equations: 
\[ x' = y + x - xy^2 - x^3, \quad y' = y - x - x^2y - y^3 \]
Here we see a new pattern, which combines the stability of an attracting fixed point with the periodic motion of a closed orbit. This is called a limit cycle. It represents a stable periodic orbit of the system, a repetitive cycle which the system generates of its own accord, regardless of its starting conditions. Notice how trajectories beginning inside the limit cycle spiral out to join it, while those beginning outside of the limit cycle spiral in to join it. Thus the limit cycle is a stable attractor: all nearby trajectories are pulled into it, which means that the system will always settle into this cycle in the long term, and that fluctuations pushing the system away from this cycle will be damped out.

Many other behaviors are possible. There are unstable limit cycles: periodic orbits in which the slightest fluctuation will send the trajectory spiraling away; and half-stable limit cycles, which attract trajectories on one side and repel them on the other, so that they are stable under fluctuations in one direction and unstable under fluctuations in the other direction.
Similarly, several different types of fixed point are possible:

\[
\begin{align*}
\text{(a) } a &< -1 \\
\text{(b) } a & = -1 \\
\text{(c) } -1 &< a < 0
\end{align*}
\]

In order to understand these different patterns of stability and instability, it can be very helpful to visualize the phase portrait in a different way. Rather than plotting the system's dynamic as a vector map, we can plot it as a potential landscape. Let's go back to our earlier example of a damped oscillator: a mass hanging from a spring, subject to friction. Instead of plotting the body's position \(x\) vs. its velocity \(x'\), we can plot \(x\) vs. the potential function \(V\), which is defined by the equation

\[
(14) \quad \frac{dV}{dx} = -\frac{dx}{dt}
\]

This yields the following graph:
The potential function derives from the physical concept of potential energy in a gravitational field. To read this graph, think of it as a landscape, and think of the state of the system as a heavy ball rolling through this landscape.

Places where the landscape is flat correspond to the fixed points of the system's phase portrait; they are places where the system can come to rest. The potential landscape gives us an intuitive way to see whether such fixed points are stable or not. In the landscape above, the fixed point at \( x = 0 \) is stable: if a fluctuation moves the system "up the slope" in either direction, it will roll back down again to rest at the fixed point. This corresponds to the phase diagram above in which every trajectory spiraled down to this point. Stable fixed points in the potential landscape are denoted by a black ball:
However, consider the following potential landscape:

Here we see a system with an *unstable* fixed point at $x = 0$, denoted by a white ball. The system can come to rest there, but the slightest fluctuation will send it "rolling down-hill" in one direction or the other. Finally, consider a third potential landscape:

This represents a system with a *half-stable* fixed point at $x = 0$, denoted by a black and white ball. If the system comes to rest at this point, it will be stable in one direction, since fluctuations that send the ball rolling up-hill will be damped out as the ball rolls.
back down to rest at the fixed point again. On the other hand, the slightest fluctuation in the other direction will send the ball rolling down-hill away from the fixed point.

These potential landscape diagrams become especially useful when we consider systems with more than one fixed point, which are known as multistable systems.

Consider the system with the following potential landscape:\textsuperscript{146}

![Potential Landscape Diagram]

This system has five fixed points: three stable, and two unstable. The fixed point at $x = 0$ is the global minimum, meaning that it is the system's most stable fixed point. However, a large enough fluctuation in $x$ could push the system over into the basin of attraction of one of the system's other two fixed points. Thus a multistable system is capable of multiple stable behaviors; and a sufficient "push" in the right direction can shift it from one stable state into another. Multistable systems can also be visualized using vector maps:

\textsuperscript{146} This landscape was generated by the potential function $V = -\cos(x) - \cos(2x)$. 
Notice that a single system multistable system can exhibit multiple fixed points and limit cycles, each of which may be stable, unstable, or half-stable. Such a system is capable of multiple behaviors, both static and cyclical, and it can be shifted from one stable behavior to another by fluctuations or external forces.

We have one final complication to consider. Up until now, we have been treating the dynamic of a system—as represented by phase portraits and potential landscapes—as itself fixed. However, we sometimes need to move to an even higher level of abstraction, to think about how the system's different ways of changing over time can themselves be changed. For example, when we considered the oscillating body hanging from a spring, we treated the mass of the body and the resistance of the spring as constant parameters—and rightly so, since for any given mass and spring, these do not vary. But what if we wanted to consider all the possible mass-and-spring systems as a class? We would then need to treat the mass and spring-resistance as variable parameters, and ask how variations in these parameters alter the dynamics and stability of the system. Let's consider a more interesting, multistable system, whose potential landscape we looked at above. The potential function of this system is

\[
V = -\cos(x) - b\cos(2x)
\]
Notice that this function has one parameter, $b$. When $b = 1$, the landscape looks like the one we saw above:

As we noted above, this system has three stable points (at $x = 0$, $-\pi$ and $\pi$) and two unstable states. But let's see what happens to the system when we vary the parameter $b$ from 1 to 0:
As $b$ decreases, the stable points at $\pi$ and $-\pi$ become less and less stable. Finally, at $b = 0.25$, these points become unstable, so that the system goes from having three stable points to only one.
Let's consider a more interesting nonlinear system, described by the equations

\[
\begin{align*}
    x' &= y + (a - x^2 - y^2)x \\
    y' &= -x + (a - x^2 - y^2)y
\end{align*}
\]

where \(a\) is a parameter. For \(a \leq 0\), the phase portrait of this system looks like this:

Like the damped oscillator we considered above, this system has a single attracting fixed point at (0,0), and all trajectories spiral down into this point. However, when \(a > 0\), the stable fixed point at the origin becomes unstable, and a new attractor appears: a limit cycle whose radius is the square root of \(a\). Here is a phase portrait of the system when \(a = 0.5\):
(The limit cycle is highlighted in blue.) In this class of systems, we see that varying the parameter $a$ can produce an important qualitative change in the dynamics of the system. At $a = 0$, the stable behavior of the system changes from static equilibrium (the fixed point at the origin) to periodic oscillation (the limit cycle of radius $\sqrt{a}$). We can visualize this change by graphing the behavior of the system as a function of $a$:

The single solid line represents the single attracting fixed point that exists when $a < 0$. At $a = 0$ this fixed point becomes unstable, which is indicated by the beginning of a
dotted line, and a stable limit cycle appears (indicated by the solid lines splitting off and circling around the formerly stable fixed point). The point $a = 0$ is called a critical point, and the change in the system's stability at this point is known as a bifurcation, (because of the way that the line representing the stable point splits in two). Accordingly, diagrams like this one are called bifurcation diagrams. This particular sort of bifurcation, in which a stable fixed point becomes unstable and spawns a limit cycle, is called a Hopf bifurcation.

The parameters of a system do not vary as part of the system's intrinsic dynamics. However, sometimes these parameters can be varied by outside forces. Critical points thus take on an enormous significance in our modeling of physical systems: they represent points at which a quantitative variation in some feature of the system can produce a sudden qualitative change in the system's behavior—a behavior that was stable can suddenly become unstable, and the system may jump quite quickly to some new stable state.

We have just done a lot of difficult conceptual work in a very condensed space. It took the work of several great geniuses and many human lifetimes to generate the conceptual shifts that we have quickly run through in just a few pages. Let us briefly review the mathematical practices we looked at, each founded on the one before:

1. Modeling a dynamic system (e.g. a moving body) as variables related by algebraic equations.

2. Modeling a dynamic system with differential equations, (algebraic equations in which some variables represent the rates of change of other variables).
3. Visualizing these differential equations as a vector field in state space (a.k.a. phase space), which is called a phase portrait.

4. Visualizing this state space as a potential landscape.

5. Looking at how this potential landscape changes with variations in parameters.

6. Visualizing these changes with a bifurcation diagram.

3. Implications of Dynamic Systems Theory

We saw above how DST emerged from attempts to generalize the Newtonian model to more complex dynamic systems, and the realization that this was not possible. Nevertheless, our thinking about scientific models—and about nature itself—has tended to remain stuck in a Newtonian mode. Newton's model describes the motion of bodies in terms of universal forces acting on individual point-masses according to universal laws. It allows the future and past of the systems it describes to be predicted and retrodicted with great accuracy. The models of DST, on the other hand, are more like the model organisms of biology than they are like the models of classical physics. Biologists study simple organisms such as Drosophila (the fruit fly) and E. Coli, which are easy to observe and manipulate, in order to draw inferences about more complex organisms such as ourselves. Similarly, DST allows us to construct and study model systems which exhibit behaviors similar to those of the natural systems we are studying. Like the model organisms of biology, the dynamic systems model is understood to be a simplification of the more complex phenomenon we are studying. This model is easier to observe and

---

147 Or at least, with an accuracy proportional to the accuracy of our measurements of the system's present state.
APPENDIX A

manipulate than the phenomenon itself, and allows us to gain insights into the latter by analogy.

The Newtonian model is not just a simpler version of the phenomenon it models; it belongs to a different order entirely. It renders the phenomenon transparent, flattening it out so that past and future can be encompassed simultaneously by our gaze. In DST, the model is itself a phenomenon which requires careful study. It can be run forward and backward in a computer simulation, but this simulation itself takes time; the model must be allowed to exhibit its behavior in its own time and on its own scale, and the behaviors it exhibits may be quite unexpected. Thus the model in DST does not replace the phenomenon with an idea, transparent and intelligible. Instead, it offers us a new phenomenon, one which is easier to investigate but still retains its opacity and independence—even though it is we who have constructed it.

Unlike the Newtonian model, models in DST do not yield quantitative predictions, but qualitative insights into the forms of a system's behavior. Phase portraits, potential landscapes and bifurcation maps make visible to us forms which would otherwise be imperceptible, because they reside in the variations between the various possible trajectories of a system. They allow us to perceive the stable patterns in a dynamic system's flow, and the sensitivity of these patterns to initial conditions, perturbations, and parameter variation. The diagrams of DST are not merely visual aids which the mathematical mind could dispense with; rather, they embody the geometrical (or more accurately, topographical) approach which is the very essence of Poincaré's method.
Whereas the models of classical physics produces universal laws, the models of DST generate insights into *patterns* which possess a certain generality. Because these patterns concern the *relations* between the parts of a system, rather than the nature of these parts, the same dynamic pattern can appear in many different materials, at many different scales. For example, in the two-dimensional dynamic system we considered above which generated a stable limit cycle, the two variables $x$ and $y$ could represent the concentrations of two chemical reagents, two animal populations in an ecosystem, or the position of a particle moving in a plane. Rather than attempting to explain nature in terms of laws that operate only at the micro-scale, DST allows us to recognize the same patterns arising in self-organizing systems at multiple space-time scales, and to say something about how these different scales interact.

The models of classical physics suggest (or at least support) a mechanistic understanding of nature: they divide nature into point-masses whose movements can be computed individually as linear effects of additive forces operating according to universal laws. DST, like classical physics, begins by dividing natural phenomena into different measurable features which can be represented by different variables. However, the complex interdependence of these variables requires that we consider the system as a circular whole, and study its behavior qualitatively. Moreover, the behavior of the system itself suggests new, higher-order variables—such as the *period* of a limit cycle, (the amount of time it takes for the system to return to a given state)—which arise out of the interaction of lower-level variables, and describe the behavior of the system in its own terms. Thus the models of DST suggest a non-additive view of nature, in which matter
organizes itself into complex systems whose parts constrain one another in nonlinear ways, generating new and unexpected forms which cannot be explained by analysis.
WORKS CITED


Works Cited


Works Cited


