1

The Search for Death Itself

The Problems of Death

Although some writers speak of "the problem of death," I think it would be better to speak instead of "the problems of death," for death is problematic in very many different ways.

There are *psychological* problems concerning death: How do we feel about our own impending death? How do we react to the death of loved ones? Are there typical psychological stages in the adjustment to the recognition of impending death? At what age do we come to have a satisfactory understanding of mortality? etc.

There are *legal* questions about death: Do dead persons have any legal rights? Should the law permit people to kill themselves?

There are *biological* problems of death: Why do organisms die? Are there any organisms that never die? Would it be possible to treat people in such a way that they would never die? Is it possible to treat people or other organisms in such a way as to make them live again after they have died?

There are also theological, literary, sociological, economic, medical, and other questions about death. All of these are interesting and deserve to be considered. But the central question of Part I is distinct from all of these questions. It is the most fundamental question that can be asked about death. It is a philosophical question. It is this: What is death?

Although the question is surely sufficiently short, it may nevertheless be a bit obscure. Its meaning will become clearer as we proceed, but it will be useful to say a few words by way of preliminary clarification. The fundamental question ("What is death?") can be formulated in a variety of different ways. We could ask for

an explanation of what we mean when we say that something has died. Another possibility would be to ask for an account of the nature of death itself. We could also state the question as a sort of challenge: can we formulate a satisfactory philosophical analysis of the concept of death itself?

Conceptual Analysis

Before we go any further, I should say a few words about conceptual analysis.

If you want to understand the engine in your motorcycle, you may study an "exploded view." This is a diagram showing all the main parts of the engine, but showing them apart from each other. Generally there are lines indicating how the parts would be connected in the functioning engine. When you come to know about the sizes and shapes of the various parts, and you come to know how they are supposed to be related to one another when the engine is properly assembled, you come to understand the engine—you begin to know how the engine works. This is a humble example of a pervasive fact about one sort of understanding. We understand complicated things by understanding their simpler parts and their relations. This is "analytical understanding."

What's true of motorcycle engines is widely thought to be true of properties. Some properties seem to be complex entities, composed of parts. Of course, the parts are not physical objects such as pushrods and valves. The parts of a property are other properties. These are related by logical relations, such as conjunction and conditionalization. For example, suppose someone is interested in the property of motherhood. We can explain this by saying that it is the conjunction of parenthood and femaleness. A person who already understands parenthood and femaleness, and who knows what 'and' means, should be able to understand motherhood. There is an ancient philosophical tradition according to which one of the best ways to come to understand a complex property is to come to understand its parts and their relations. Philosophers in this tradition focus on interesting and important complex properties and attempt to describe their parts and the way in which the

parts are organized in the whole. (Since properties are sometimes called concepts, this activity is often called 'conceptual analysis'.)

There are several different ways in which an analysis may be stated. We can consider motherhood again. We may present our analysis in any of the following forms:

M1: Motherhood is the conjunction of femaleness and parenthood.
M2: To be a mother is to be a female parent.
M3: Being a mother is being a female parent.
M4: 'Mother' means the same as 'female parent'.

For certain personal reasons, I happen to prefer to formulate analyses as definitions. Thus, if I were attempting to analyze motherhood, I might present my results in this form:

M5: x is a mother at t = df. x is a female parent at t.

One necessary condition of success for an analysis is lack of counterexamples. Thus, if M5 is correct, we will be unable to conceive of a case in which something would clearly be a mother but not a female parent (or vice versa). When searching for counterexamples, we must beware of unclear or controversial borderline cases. These do not make satisfactory counterexamples. A satisfactory counterexample must be an uncontroversial case. It must be a case with respect to which there is nearly unanimous and relatively firm agreement.

A proposed analysis of a concept must get the uncontroversial cases right. If it gets them all right, we can consider what it legislates concerning the controversial cases. We should be open-minded. Perhaps a proposed analysis will yield surprising results with respect to the controversial cases. That would be no mark against it. It would merely show that we can learn things from good philosophy.

A second necessary condition of success for an analysis is this: the proposed definition must not make use of any terms that are as obscure in meaning as the term being analyzed itself. Suppose, for example, we find the concept of death puzzling and attempt to clear things up by offering an analysis. We say that to die is "to

drift across the great divide into the other realm." Such remarks obviously shed very little light on the nature of death. Our analysis would be much more useful if it employed only clear, simple, and literal terminology. A closely related point is that the proposed definition must not be circular. If we are trying to explain what death is, the word 'dies' must not appear on the right-hand side of the definition. The problem with any such definition is obvious: we'd have to understand the meaning of 'dies' before we could use the definition to learn what 'dies' means. How could the definition be of any value?

A somewhat more subtle point concerns covertly circular definitions such as this one:

D2: x dies at t =df. x reaches thanatological termination at t.

We note a strange term in D2—'thanatological'. We look it up in the dictionary. We discover that 'thanatological' means 'having to do with death'. We thereby discover that D2 is covertly circular. We have to understand the meaning of death in order to understand the meaning of one of the terms in the proposed analysis of death. Our rule against circularity must be understood to prohibit covertly circular definitions, too.

Analysis of Death or Criterion for Death?

It is important to distinguish clearly between what I am here calling the analysis of the concept of death, and a quite different project that is sometimes called "defining death." For legal purposes, it is sometimes extremely important to determine whether someone is alive or dead. For example, before we remove the organs from someone's body for transplantation, we want to be sure that the person is dead. Similarly, before we expend huge amounts of energy and money trying to heal someone's wounds, we want to be sure that the person is not already dead. Thus, philosophers and others have set about trying to isolate some clear marks or criteria of death. Confusion arises because these criteria are sometimes called 'definitions'.

A criterion of death would be a fairly easily recognized property

that serves as an indicator of death. Someone is dead if and only if he or she displays the criterion. Since it is important to be able to specify a precise moment of death, some would insist that a really good criterion of death would be an "all-or-nothing" property, rather than a property that comes in degrees. Suppose, for example, that in absolutely every case, a certain sort of brain activity ("z-waves") could be detected in living human beings. Suppose also that this sort of brain activity abruptly stops when a person dies. Suppose it is relatively easy to determine whether a brain is emitting z-waves. We might then "define" death as the cessation of that brain activity. Thus, we could say:

D3: x dies at t iff x's brain ceases to emit z-waves at t.

No matter how good a criterion of death D3 might be, it is not an analysis of the concept of death. One obvious point is that it has no relevance to plants and other mortal things that have no brains. It applies only to human beings (and perhaps other animals whose brains emit z-waves). Another point is that if something like D3 is true, it is only contingently so. Thus, even if earthly humans always cease to emit z-waves at the moment of death, things surely could have been otherwise. If evolution had progressed in some other possible way, it might have turned out that no living human being's brain ever emitted a z-wave. We can also conceive of other possible ways in which evolution might have progressed. In some of these, it might have turned out that corpses would continue to emit z-waves long after death. Finally, and perhaps most important, D3 sheds no light on the nature of death—it doesn't help us to understand what death is. It merely purports to help us determine the moment when death arrives.

There are other important differences between an analysis of the concept of death and a criterion for human death. One of them concerns what counts as *success*. When a criterion of human death is proposed, the aim is to offer something that can be acted upon; something that might be useful; something that might be adopted. It is suggested that people in the legal, mortuary, and medical professions should adopt the criterion and use it in their activities. The proposed criterion is successful if it gains widespread acceptance. An analysis is not like this. It does not purport to be useful.

It is not intended as a solution to any practical problem. Success for an analysis is measured by the extent to which it serves to enlighten us about the nature of death itself.

A closely related and important point concerns the adoption of a criterion. People in the medical, mortuary, and legal professions are often called upon to determine whether or not someone is dead. Such determinations have profound and immediate practical implications. Thus, for example, a mortician may be asked to bury a certain body. For a variety of reasons, he doesn't want to do this if the body is not yet dead. One reason might be this: if it should later come to light that he buried a still-living body, he might be found guilty of a crime—or at least of professional incompetence. He might lose his license. Thus, the mortician wants to be sure that the body is really dead.

Furthermore, he wants to be able to demonstrate that he took proper precautions before proceeding with interment. If the z-wave criterion were universally accepted, he could proceed as follows: he could attach the z-meter to the body and check for z-waves; he could ask two reliable observers to sign the printout, certifying that there were no z-waves; he could attach a notarized copy of the printout with the death certificate. Later, if a question should arise, the mortician would be protected. He could claim that he took all the legally mandated steps. He checked for z-waves, found none, and filed the necessary papers—all *before* burying the body.

It is consistent with all this to suppose that medical technology could advance in the area of z-waves. New procedures might be introduced that make it possible to revive a person even though he has emitted no z-waves for several minutes. If an advance of this sort were to occur, the z-wave criterion would have to be abandoned. Perhaps there would be a period of transition, during which new criteria would be proposed and debated; perhaps symposia would be held during which the merits and defects of these proposals would be argued; perhaps eventually a new criterion would be adopted. Thereafter (until yet another criterion is adopted) morticians would ignore the z-waves and check instead to determine whether or not the new criterion was satisfied.

Nothing quite like this could happen with respect to the *analysis* of the concept of death. Suppose there is such a thing as the

concept of death. In other words, suppose the word 'dies' has a certain literal meaning in the English language. Experts may gather to debate about this meaning. At the end of their conference, they may reach a conclusion about the meaning of 'x dies at t'. They might hope that others would agree with them. However, their conclusion might simply be wrong. It is possible that the word just does not mean what they say. They were mistaken about the concept of death. Perhaps there is an obvious counterexample to their agreed analysis. No amount of consensus can overrule a fact of this sort. On the other hand, if their analysis is correct, it will express a metaphysical necessity. Nothing could later happen that would call for a revision of their decision. No change in medical technology could make it necessary to alter the analysis of the concept of death.

(Of course, the word 'dies' has a history; it is possible that it might come to express some new concept. In that case, a new analysis might be in order. However, it would not be a new analysis of the original concept—if the old analysis was correct, it would still be correct. What might be sought would be an analysis of the concept that the word now expresses.)

So there are many differences between the analytic project and the criterial project:

1. A criterion of death purports to help us locate the moment when death comes. An analysis of death purports to tell us what death is.
2. A criterion of death may be formulated in such a way as to apply only to human beings. An analysis of death must apply equally to anything that can die.
3. A criterion of death may be quite useful even if only contingently true. Indeed, it might be quite useful even if there are a few rare counterexamples. An analysis of death is, if true at all, then necessarily true. There cannot be even so much as a *possible* falsifying instance.
4. A criterion of death may be "in force" during a certain period of time, and then, with advances in technology, abandoned. An analysis of death, on the other hand, is eternally true if true at all.
5. A criterion of death is a success if enough people, thinking it would be useful, decide to adopt it. An analysis of death is a success if it is true—even if no one adopts it.

I am inclined to think that the analytical project has a sort of conceptual priority over the criterial project. Perhaps this may be brought out by reflecting on what might happen when experts assemble to formulate a new criterion. Suppose some experts propose a criterion according to which people are to be counted as dead when their hearts stop beating. Others object to this criterion, pointing out that hearts can easily be restarted by the careful administration of electric shocks. They claim that this shows that the proposed criterion is unacceptable—it conflicts with the "fact" that once a person dies, he cannot come back to life. Suppose advocates of the proposal grant that it has the indicated feature. But they insist that this is no defect. As they see it, people *can* come back to life after dying. Thus, the fact that their criterion permits this is a *virtue* of the proposal, not a defect.

I think this sort of disagreement would show that there is a certain amount of confusion about the nature of death itself. In the example, some of the conventioneers take death to have a certain "logic." As they understand it, someone can die and then live again. Other conventioneers think that the concept of death rules this out. Death, for them, must be *permanent.* Thus, it appears that the conventioneers do not have their eyes on the same target. Some are trying to formulate a criterion for one concept of death, and others are tying to formulate a criterion for some other concept of death. Perhaps some are trying to formulate a criterion even though they don't yet have any clear concept of death in mind.

This sort of confusion reveals one way in which the analytical project might have priority over the criterial project. It might be good for the conventioneers to agree at the outset about the nature of death. If they all agreed about what death is, they could more fruitfully go about their business of trying to agree on a clear mark of its arrival.

When I speak here of the search for death itself, I mean to indicate the effort to discover a satisfactory philosophical *analysis* of the concept of death. It is the attempt to answer the question 'What is death?', not the question 'What is the conventionally accepted indicator that death has arrived?'

The Biological Concept of Death

Some of the literature on death seems to be based on the assumption that the target concept is some special concept of death that is applicable only to *people.* Thus, for example, one writer says that to die is to cease permanently to be conscious of one's own psychological experiences.[1] He obviously does not assume that when a tree or a bacterium dies, it too ceases to be conscious in this way.

This assumption about death seems rather odd to me. I am more inclined to suppose that there is a single concept of death that has application throughout the biological realm. Perhaps I can clarify my view by appeal to some examples. Consider the following sentences:

1. JFK died in November 1963.
2. The last dodo died in April 1681.
3. My oldest Baldwin apple tree died during the winter of 1986.

I cannot think of any reason to suppose that the word 'died' has one meaning in sentence (1) and different meanings in the other two sentences. It seems to me that what we say about JFK in (1) is precisely the same as what we say about an apple tree in (3). Some slight evidence for my view can be derived from the fact that there would be nothing amusing, paradoxical, or otherwise out of the ordinary about the claim that if (1), (2), and (3) are all true, then at least three different things have died. If 'died' were used in different senses in these sentences, then the inference would be an eyebrow-raiser. It would be a play on words; it would be like the case in which a man tells us he owns two planes—one a single-engine Cessna that he uses on business trips and the other a single-bladed Stanley that he uses in his woodworking shop.

So I think there is a single concept of death that has application throughout the biological realm. This is not to say, of course, that all deaths are equally important; or that all deaths are manifested in the same way; or that one *criterion* of death must apply to every sort of entity. It is just to say, among other things, that the word 'died' has a certain intension, or meaning, and that it is possible that sentences (1), (2), and (3) may be used in such a way that the

word 'died' in each of them expresses precisely this intension. In other words, it is to say that 'died' as used in (1) might mean exactly what it means as used in (2) and (3).

Someone might agree that there is a universally applicable concept of death but insist that there is also another concept of death applicable only to people. I think this is a mistake. I do not believe that there is a special concept of death applicable only to people. I do not believe that the word 'died' has a sense for which it would be a necessary truth that if a thing died, then it must have been a person. However, quite a few philosophers seem to believe that there is such a sense, and I see no way to prove them wrong. Thus, I should be cautious in describing my project: I am here interested in the analysis of the biological concept of death—I want to know what we mean when we say that something (whether human being, apple tree, or bird) dies; if there is a specifically "personal" concept of death, then I am also interested in it. But at least at the outset, I seek to understand the biological concept of death.

Life as a Part of Death

Our question is this: What is death? What do we mean when we say, using the biological concept of death, that something dies? According to the most popular answer, death is the cessation of life. In other words, to die is to cease to live. I call this "the standard analysis":

D1: x dies at t =df. x ceases to be alive at t.

D1 may seem to satisfy some of our requirements for success in conceptual analysis. It apparently does not violate the circularity condition. It seems to satisfy the necessity condition (we will reconsider this point later in Chapter 4). However, some philosophers would say that D1 trades one mystery for another; they would point out that D1 purports to clarify the concept of death by appeal to the concept of life; and they would say that the concept of life is hopelessly obscure—too obscure to be useful in the explication of the concept of death. Perhaps this objection would be an overstatement. In any case, this much is surely true: if we don't know what

life is (what 'alive' means) then D1 is less than fully satisfactory. It is a definition per obscurius; it purports to tell us what a mysterious term means, but it makes use of yet another term at least as mysterious as the one originally in question.

We want to understand the biological concept of death. According to the standard analysis, death is the cessation of life. Thus, it may appear that we must first understand life. Let us then consider the nature of life. What does 'alive' mean? What do we mean when we say that something is alive? What is the nature of life? If we can answer this question, we can make use of the concept of life in our analysis of the concept of death. We can rebut the charge of obscurity in D1.

In Chapter 2, I consider some of the most popular analyses of the concept of life.

2

Life-Functional Theories of Life

Life itself

At least in part because it seems to play a crucial role in the analysis of death, I am interested in life—that is, I am interested in the property expressed by the phrase 'is alive'. The very same property (I think) is expressed by a variety of other phrases. For example, when we say that something 'has life' or 'is a living thing' then, if we use these terms in their central literal sense, we attribute life to it. These expressions also express the property of being alive. The challenge is to formulate a satisfactory philosophical analysis of that property.

There are other standard ways of referring to this property. Sometimes, especially when we are attending to the fact that people can be aware of properties, we call them "concepts" or "ideas." Thus, we may speak of the concept of life, or the idea of life. I take this to be the property of life, but thought of as an object of thought. In other words, it seems to me that when some property becomes the object of someone's thought, we can call that property a concept, or an idea. Properties also play a central role in the explanation of meaning. The property of being alive is the *intension* of the phrase '. . . is alive'. Since intensions are often taken to be "meanings," we may refer to the same entity when we speak of 'the meaning of "alive" '. Finally, some philosophers and biologists must surely be thinking of life itself when they speak of "the nature of living things."

It might seem that there's no difficulty in identifying our target. We all have some rough idea of what life is supposed to be. We all

understand typical sentences in which the word 'alive' appears. But even at this early stage there are serious complications.

One complication concerns the fact that 'life' is also used in English to indicate something like the "history" of a living thing. For example, consider the statements "Her life lasted only sixteen and one half years," "A cat has nine lives," "His life was filled with joy," "His life was more interesting than mine." It would be implausible to interpret 'life' (or 'lives') in these sentences as referring to a certain property. In each case, it seems to refer to something with duration—something that lasts through a period of time. Life itself seems not to be like that. Furthermore, in these examples 'life' is used in such a way that it makes sense to say that different people have different lives. Yet one of the most striking facts about the property of life is that (apparently) every living thing exemplifies that very same property. It is a "universal"; it is common to all living things. Thus, these "lives" are distinct from life itself. For present purposes, we must put aside these "lives."[1] We are interested in a property that each living thing has throughout its life—the property of being alive, or "life."

A closely related complication concerns the fact that 'life' is sometimes used as a mass term indicating the aggregate of "living stuff." For example, when a man says that a certain region is "filled with life," or is "teeming" or "bursting" with life, he probably is not using 'life' as the name of a property. He's probably using it as a mass term for a kind of stuff—living stuff. For present purposes, we must also put aside this stuff. We are interested in the property, life itself, that characterizes this stuff.

Another complication is that 'life' and 'alive' are used in a variety of semantically distinct ways. Some of these are uncontroversially literal, and others are uncontroversially metaphorical. For example, suppose there has been a train wreck, and casualties are being removed from the wreckage. We might want to know if a certain passenger survived the accident. We ask, 'Is she still alive?' 'Alive' here is almost certainly used literally. On the other hand, when someone sings 'The hills are alive with the sound of music,' she almost certainly uses 'alive' metaphorically. Similarly, when we say that an automobile 'springs to life' when you turn the key, we use 'life' metaphorically (except perhaps in the Stephen King novel *Christine*). The complication arises because

there are many intermediate cases—cases in which it is not clear whether 'alive' is being used literally or metaphorically. These can lead to confusion.

An interesting example concerns ponds. Environmentalists sometimes tell us that ponds are living things. They describe the geological conditions that "give birth to" a pond; they describe the way in which the pond "grows" and "flourishes"; they may describe the way in which the pond "becomes sick" when it fills up with silt or rubbish; finally, they may describe the way it "dies" when it ultimately turns into a swampy bog or a meadow. There are several possibilities here. (1) The environmentalists may think that ponds are literally alive—that they exemplify true life just as their fish and frogs do. (2) They may think that ponds are not literally alive, but that the history of a pond is like the history of a genuinely living thing, such as a fish or a frog. This recognition of similarity may provoke the use of 'living' in a strictly metaphorical way. (3) Finally, the environmentalists may be engaging in loose talk. If pressed, they might admit that they do not know whether they meant their expression literally or figuratively.

Another troublesome case is the so-called Gaia Hypothesis, introduced by J. E. Lovelock in his book *Gaia: A New Look at Life on Earth*. The biosphere is the part of the Earth that contains living stuff—roughly the part starting at the bottom of the topsoil and rising to the lower part of the atmosphere. The biosphere also contains the seas, lakes, rivers, and other bodies of water. Gaians claim to believe that the biosphere is alive. Perhaps they are using 'alive' metaphorically. Perhaps all they mean is that the biosphere is in interesting ways *like* a living thing. In this case, their claim seems uncontroversial. It's obvious that the earth is like a living thing in some ways. Perhaps Gaians find these similarities interesting.

On the other hand, it is possible that Gaians are using 'alive' in what they take to be its literal sense. Consider an ordinary fish happily swimming in an ordinary pond. We can use the word 'alive' in its literal sense to characterize the fish. It is possible that some Gaians use the word 'alive' in just this sense to characterize the biosphere. If so, their view seems to me to be silly. However, in the absence of a widely accepted understanding of the concept of life, it may be hard to adjudicate such disagreements.

I intend to proceed on the assumption that 'alive' has a central,

literal sense. We generally use the word with that sense when we use it to describe an ordinary fish swimming in an ordinary pond. I will also assume that when used in this literal sense, 'alive' expresses a certain property. This is "life itself." I take it that the attempt to "define life" is the attempt to formulate an adequate definition of 'alive' when used in this way. This project is equivalent to the project of formulating an analysis of the concept of life, or an analysis of the property of being alive.

Some Preliminary Objections

In *The Growth of Biological Thought,* Ernst Mayr makes the following statement:

> Attempts have been made again and again to define "life." These endeavors are rather futile since it is now quite clear that there is no special substance, object, or force that can be identified with life.[2]

A few pages later, Mayr says:

> Life . . . is simply the reification of the processes of living. Criteria for living can be stated and adopted, but there is no such thing as an independent "life" in a living organism. . . . The avoidance of nouns that are nothing but reifications of processes greatly facilitates the analysis of the phenomena that are characteristic for biology.[3]

It is not entirely clear that Mayr means to say that 'alive' cannot be defined, but his words strongly suggest this. He says that there is something that has been tried "again and again" and that is "rather futile." He characterizes this as the attempt to "define life." If he is not talking about our topic, it is hard to understand what he is talking about. In any case, let us consider the objection.

I think that Mayr is probably correct in saying that there is no special substance, [physical] object, or force that can be identified with life. But it surely does not follow from this that life cannot be defined. There is no substance, physical object, or force that can be identified with motherhood, but motherhood can be defined. Equally, there is no substance, physical object, or force that can be

identified with primeness (as in "three is a prime number"), but that fact does not rule out analyses of the concept of primeness.

Mayr's words suggest that he thinks that there is no such thing as life. One who thinks that life is a property, as I do, does not think that life is a physical object or substance, and he or she very well may reject the notion that it is a force. No property is a physical object, and quite a few fail to be forces. Thus, Mayr's remarks give us no reason for agreeing that there is no such thing as life. So long as we are careful to keep life in its proper ontological category, I see no danger in supposing it to exist.

Others have given other reasons for thinking that life cannot be defined. One reason is that living things are a very varied and heterogeneous group: Some have suggested that living things are so varied that they have nothing in common—hence there is no such property as life. This objection to our project strikes me as being premature. In the first place, it is obvious that living things in fact do have a lot in common—they share all necessary properties, and they are all alive. In the second place, huge variations in other respects are compatible with important similarities in some respects. For example, consider the class of mothers. It displays enormous heterogeneity—some are people and others are bugs; some weigh only a fraction of an ounce, and others weigh over a ton. Nevertheless, its members have something in common. They are all female parents.

So it seems to me that we may continue. No good reason has been given for supposing that our quest is wrongheaded from the start. Let us then turn to some of the most important traditional attempts to define life itself.

Aristotle's Life-Functional Analysis of Life

No approach to the analysis of life itself has a more distinguished history than the one we may call the "life-functional" approach. Aristotle himself is the founder and chief advocate of this approach. Since his time, hundreds of philosophers and biologists have defended variants of his view. Let's begin by considering Aristotle's version of the theory.

Although different texts suggest different lists, Aristotle seems to have recognized the following main life functions:

Nutrition: the capacity to get food, absorb it into oneself, and thereby to grow. Surprisingly, Aristotle maintains that nutrition is inseparable from reproduction; that these are both functions of the same "soul"—the nutritive soul. "Nutrition and reproduction are due to one and the same psychic power."[4] Reproduction is the capacity to produce offspring. Aristotle maintains that the capacity to engage in nutrition is the most basic and widely distributed life function. His view seems to be that every living thing, plant or animal, has this capacity; no nonliving thing has it. Thus, apparently according to Aristotle, a thing is alive if and only if it can engage in nutrition and reproduction.

Sensation: An organism has sense if it is capable of "receiving into itself the sensible forms of things without. . . ."[5] This capacity operates in two different modes. Certain senses (touch and taste) operate on objects that are in contact with the organism. Aristotle seems to have believed that all animals have the capacity to engage in this "immediate sensation."[6] Plants, however, do not have it.

The second mode of sensation involves the perception of objects that are not in direct contact with the organism. Sight, smell, and hearing are the instances of this mediate form of sensation. Aristotle claims[7] that these senses are not found in every sort of animal, but only in animals that can move. He apparently felt that there would be no point in giving sight, for example, to a fixed animal such as a barnacle. After all, even if the barnacle could see a tasty bit of food a few inches away, it would not be able to do anything about it. Having sight would not improve the well-being of a barnacle. Hence, nature did not give barnacles eyes.

Motion: An organism has "the locomotive soul" if it is capable of moving itself from place to place. Aristotle apparently held that the locomotive soul is always found in conjunction with the far sensitive soul. Stripped of its soulful terminology, the claim amounts to this: a creature can move itself about if and only if it can see, hear, or smell.

Thought: some organisms are able to think. This includes people and perhaps some other rational beings. Aristotle says[8] that this sort of "rational soul" is "capable of existence in isolation from all other psychic powers." Perhaps he is thinking of gods. There seems to be a slight tension between this remark and his later

remark that "the nutritive soul then must be possessed by everything that is alive. . . .'[9] Perhaps he means to say this: among mortal beings, the nutritive soul is universal. If we include immortal beings, we find instances of things with rational souls but without nutritive souls.

Aristotle maintains that "Of the psychic powers above enumerated, some kinds of living things, as we said, possess all, some less than all, others one only."[10] If we simplify slightly, and emphasize certain texts rather than others, we can present the outlines of Aristotle's view in a chart.[11]

	Plants	*Fixed Animals*	*Beasts*	*People*	*[Gods]*
Nutrition	Yes	Yes	Yes	Yes	No
Reproduction	Yes*	Yes*	Yes*	Yes*	No
Near Sensation	No	Yes	Yes	Yes	No
Far Sensation	No	No	Yes	Yes	No
Motion	No	No	Yes	Yes	No
Thought	No	No	No	Yes	Yes

A few comments may be in order. First, it is important to recognize that Aristotle views nutrition and reproduction as functions of the same "soul"—the nutritive. Furthermore, he explicitly acknowledges[12] that there are many cases in which individual plants and animals are incapable of engaging in reproduction. Thus, I have asterisked the "yes" in each occurrence concerning reproduction. Second, the column marked "[Gods]" is just a guess on my part. Aristotle merely indicates that in his view, people are the only *mortal* beings with rational souls. This leaves open the question of whether there might also be some *immortal* beings with souls.

While the chart is surely suggestive, it does not yet constitute an answer to our fundamental question, What is Life? One natural answer, based on these Aristotelian ideas, would be this:

> LF1: x is alive at t =df. x is able to perform at least one of the life functions at t.

It seems to me, however, that there are very serious problems for LF1. One of these problems concerns motion. If we understand

motion in a straightforward manner, we will have to say that any mechanical device that is capable of setting itself into motion displays this sort of life function. As a result, LF1 seems to imply that alarm clocks, robots of various sorts, automatic lawn sprinkling devices, and the like, are all alive.

Aristotle himself seems to have been particularly impressed by the apparent universality of the nutritive soul. In a widely quoted passage, he says:

> This [the power of self-nutrition] is the originative power the possession of which leads us to speak of things as *living* at all. . . .[13]

And in another just a few paragraphs later we find:

> First of all we must treat of nutrition and reproduction, for the nutritive soul is found along with all the others and is the most primitive and widely distributed power of soul, being indeed that one in virtue of which all are said to have life. The acts in which it manifests itself are reproduction and the use of food.[14]

On one interpretation, Aristotle might be taken to be maintaining this view about life itself:

> LF2: x is alive at t =df. x is able to engage in nutrition and reproduction at t.

Obviously, however, this won't do. As Aristotle himself pointed out, very many living things cannot engage in reproduction.[15] He cites three sorts of cases in which living things lack the capacity to reproduce. Some cannot reproduce because they are too young. (I suppose we could add that some cannot reproduce because they are too old.) He goes on to mention organisms that are "mutilated." Finally, he mentions organisms whose mode of reproduction is spontaneous. Aristotle seems to have thought that some creatures were produced by spontaneous generation. Be this as it may, the first two points are surely conclusive. Many things cannot engage in reproduction even though they are alive. Thus, LF2 is clearly wrong.

We could modify LF2 merely by deleting the second conjunct. This would yield:

LF3: x is alive at t =df. x is able to engage in nutrition at t.

Aristotle devotes several pages to his discussion of nutrition. He presents a fascinating discussion of the analysis of the concept of food, and he talks about the nature of growth. In the end, however, stripped of complexities, his view seems to be this: a creature can engage in nutrition at a time if and only if it is able at that time to acquire some food, absorb that food and make it part of itself, and as a result, grow and have the energy needed to do what needs to be done.

By way of criticism of LF3, I here quote at length a moving passage from *The Nature of Living Things* by C. Brooke Worth and Robert K. Enders. Worth and Enders have been describing an experiment with a cecropia moth. Shortly after the moth emerged from its pupal shell, Worth and Enders tied a string around its waist and placed it outdoors where a male could find it. According to their account, a male did find the moth:

> Copulation lasts through the day. In the evening we untether the female and put her in a shoe box for an hour or so. By this time she has laid fifty eggs, so we let her go. For the next few nights she will dot the rest of her eggs, some two or three hundred, on various trees. What then? Already her gorgeous wings are a bit tattered. Her abdomen has shrunk and she is beginning to tremble. But naturally! She has been so busy that food has been forgotten. Now she is faced with tragedy greater than one would suspect, for her entire race has forgotten about food. The caterpillar's digestive tract, taken to bed in the pupa's interior, was completely demolished during moth-formation, but no substitute was provided. So here flits the *cecropia*, completely absolved of her responsibilities to posterity, but unable to taste the rewards of accomplishment. No mouth, no stomach—only a small additional reserve of stored energy. The moth flies about bright lights for a few evenings more but then falls ragged and quivering to the ground, where ants slowly extinguish the rest of its waning life.[16]

The case of the cecropia moth demonstrates, I think, that a creature can be alive at a time even though it is not then able to engage in nutrition. Other examples come to mind. While undergoing abdominal surgery, a person's digestive system might be temporarily detached and shut down. Furthermore, the patient might be unconscious and paralyzed. Such a patient is clearly alive, yet unable to take in any food (because of being paralyzed and unconscious) and unable to absorb any food (because the intestines are detached). As it stands, LF3 is unsatisfactory.

It should be obvious that we cannot modify LF3 in anything like this way:

LF3′: x is alive at t =df. either (i) x is able to engage in nutrition at t, or else (ii) x was able to engage in nutrition at some time earlier than t.

The problem with LF3′ is that nearly every dead organism satisfies the second disjunct of the definiens. Corpses are nonliving things that *formerly* were able to engage in nutrition.

It seems clear, then, that Aristotle's version of the life-functional approach suffers from some serious problems. Perhaps two thousand years of biological research has provided the basis for a more plausible formulation. Let us therefore consider some typical modern examples of the life-functional approach to the analysis of the concept of life.

Some Modern Life-Functional Analyses of Life

In *Philosophy: an introduction to the art of wondering,* James Christian says:

> At present, it appears that "life" can be defined with two qualities: self-replication and mutability. Any organism possessing these two qualities can be considered alive. In these two characteristics is contained the essential processes of evolution: continuity and adaptation. . . . But mutability—the ability to effect changes from one generation to another and adapt to a fluid environment—is essential. Without the ability to change and adapt no species could long

> survive. Environmental conditions are forever changing; species must be able to change along with their environments. So far as we know, only living organisms have these two qualities, and an organism must possess both qualities to be considered alive.[17]

It is interesting to compare Christian's view with a view presented in Richard Goldsby's *Biology*. Goldsby reports that exobiologists at NASA have been interested in the nature of life. One of their missions, apparently, is to send spaceships to other planets on a search for living things. Since conditions on other planets are undoubtedly quite different from conditions here on earth, the NASA exobiologists recognized that it would be a big mistake to design their equipment to recognize "life as we know it." The equipment would have to recognize life even if it appeared in forms quite different from the forms we know. According to Goldsby:

> These scientists have tried to reduce the functional definition of life to the most simple, general, and abstract criteria. Their conclusion is that only two characteristics distinguish living entities from inanimate nature: the ability to *reproduce* themselves, and the means of producing and perpetuating *genetic variations* among the offspring.[18]

Goldsby goes on to claim that this very abstract definition of life has certain corollaries. In order to reproduce, an organism has to stay alive at least for a little while. This requires *metabolism* (the ability to "absorb, transform, and use material from the environment") and *adaptation* (the ability to make useful, genuinely "homeostatic" responses to changes in the environment). Christian also mentions these other life functions[19] along with a few others, but it isn't clear that he views them as "corollaries" of the core definition.

Thus it appears that NASA (at least according to Goldsby) and Christian would agree that life itself can be defined as follows:

> LF4: x is alive at t =df. x is able to reproduce at t, and x is able to produce and perpetuate genetic variation among offspring at t.

The influence of Charles Darwin is obvious in this account of life. As both Christian and Goldsby point out, this analysis of the concept of life very naturally leads to the conclusion that living things will be able to evolve as their environment changes (so long as the environment changes at a suitable rate, and the mutations occur at a suitable rate). Because of this emphasis, this sort of analysis is sometimes called "the genetic analysis of life." It is nowadays quite popular with philosophers and biologists.

Nevertheless it is clear that, as it stands, LF4 won't work. As we have already noted, lots of living things are unable to reproduce. In some cases, infertility is only temporary, but in other cases it is permanent and lifelong. Among ants and bees, for example, many living individuals are permanently sterile. The same holds true for certain hybrids, such as mules. Obviously, if a thing can't produce offspring at all, then it surely cannot produce offspring manifesting genetic variations from itself. Thus, each conjunct of the the proposed analysis of life is clearly too narrow, and the analysis itself fails.

One natural modification of LF4 suggests itself. We must distinguish between the concrete, individual organism (this particular mosquito—the one that just bit my ear) and the species (in this case, I suspect, *Culex pipiens*). Living individuals may be unable to reproduce. But as a number of authors going back to Aristotle have remarked, a viable species must have some standard method (or methods) of reproduction. Typical adult, unmutilated instances of the species generally reproduce in the method standard for the species.

In the case of mutation, or variation, the focus on the species rather than the individual is even more obvious. It makes virtually no sense to say that an individual mosquito undergoes genetic variation from generation to generation. The individual has the same genetic makeup throughout its existence and is a member of exactly one generation, no matter how long it lives. However, it does make sense to say that a *species* undergoes genetic variation from generation to generation. Roughly, what this means is that individuals of one generation are genetically different from individuals of other generations.

In order to simplify our discussion, let us introduce some simpli-

fying terminology. We can say that a species is *reproductive* when there is a method of reproduction such that typical members of that species reproduce by that method. Thus, the amoeba is reproductive because typical members undergo fission; the tomato is reproductive because typical members produce viable seeds; the lion is reproductive because males inseminate females who then carry their cubs to term; and so on. Going beyond this, we can say that a species is *variably reproductive* when it is reproductive and when individuals of one generation are capable of producing offspring that manifest small genetic differences from their parents. We need not attempt to define what is meant by "small" genetic differences.

Making use of these abbreviations, and taking note of the distinction between species and individuals, we could replace LF4 with:

LF5: x is alive at t =df. x is a member of some variably reproductive species at t.

The advantage of LF5 over LF4 is clear. Immature, "mutilated," and postreproductive individuals are not counterexamples to LF5. Such individuals are counted as living not because they can reproduce, but because they are members of reproductive species. Furthermore, sterile ants and bees also count as alive, since they are members of variably reproductive species. Their own sterility is here irrelevant. Unfortunately, a moment's reflection will reveal that LF5 casts the net of life much too widely. It correctly counts the senile as alive; it incorrectly counts the deceased as alive. A dead chicken is still a chicken; it's still a member of a variably reproductive species. LF5 therefore tells us that each such chicken is still alive.

Someone might insist that a dead chicken is really not a chicken. Such a person might claim that the corpse of a member of a species is not itself a member of that species. This seems to me to be wrong. If we reflect for a moment on the activities of taxonomists, its wrongheadedness will become even clearer. Entomological taxonomists, for example, do virtually all of their work with dead specimens. They sort individuals into species—but the individuals are rarely living. They point to their cases of dead butterflies and say, "This is the Monarch; that is the Viceroy. Notice the difference

in the pattern." If the current proposal were correct, the taxonomists would be wrong. Strictly speaking, there would be no Monarchs or Viceroys in their case. Only a dedicated philosopher could say such a thing with a straight face.

This sort of approach gives rise to further profound difficulties. Perhaps the most intractable of these is this: LF5 makes use of the notoriously obscure concept of "species." In order to make LF5 fully satisfactory as a philosophical analysis of the concept of life, we would have to give some account of the concept of species, and that would be a most difficult task. However, if we insist on altering our conception of species in such a way that, as a matter of conceptual necessity, each species contains only *living* members, then the task becomes vastly more difficult. Furthermore, as should be obvious, the task presupposes a solution to our present problem. In order to define 'species', we would have first to define 'alive'.

The Matthews Approach

A modern, sophisticated version of Aristotle's approach has been developed by my colleague, Gary Matthews.[20] Matthew's idea involves several refinements of Aristotle's approach. One of these is the idea that capacities that are life functions for the members of one species might not be life functions for the members of another species—it varies from species to species. Thus, a fundamental concept is expressed by 'F is a life function for the members of species S' (which Matthews expresses by 'x is a psychic power for species s').

It would be natural to suppose that a *life-functional* property is one that an individual needs in order to be alive, but it is clear that the properties on Aristotle's list don't have this feature. Individuals can live without them. Obviously, an individual can continue to live even if it cannot engage in reproduction or locomotion. Matthew's idea is that life functions are capacities without which a *species* cannot be preserved.

More exactly, the idea is that a certain capacity is a life function for a species if and only if *most* members of that species must have that capacity in order for the *species* to be preserved.

Consider the capacity to reproduce. An individual mosquito might be able to survive quite well even if it were unable to reproduce. However, if most mosquitoes were unable to reproduce, the species would soon begin to die out. Eventually, there would be no more mosquitoes. Thus, the capacity to reproduce in the mosquitoish way is a life function for mosquitoes. It is a capacity that members of that species must have if the species is to be preserved. In a comment, Matthews mentions the following as plausible candidates for vital functions: reason, sense perception, locomotion, appetite, metabolism, and reproduction.[21] Presumably, he means that certain forms of these are vital functions for certain species.

Making use of his novel concept of vital function, Matthews proceeds to offer an Aristotelian analysis of life itself. He says that for a thing to be alive is for it to be able to exercise at least one vital function for its species.[22] In other words:

LF6: x is alive at t =df. at t, x is able to exercise at least one capacity that is a vital function for x's species.

Matthew's proposal is subtle and insightful. However, there are a few difficulties.

I think that most of the properties Matthews mentions are not vital functions according to his definition. Consider reproduction, for example, among ants or bees. In certain species of social bees, one female out of ten thousand engages in reproduction.[23] Most of the members of the hive are sterile workers or drones. Now consider the property of *reproducing in the manner peculiar to bees.* It is clear that it is *not* the case that individual organisms belonging to *Apis mellifera* must exercise that property in order for the species to be preserved. In fact, most individuals do not exercise that property, and the species has been preserved for thousands of years. It is not on the endangered species list.

Consider any species that is profligate, that is, far more offspring are produced than are needed to keep the species going. Just a few individuals produce enough offspring to populate the whole next generation. Making allowances for the need for genetic variation, we can suppose that it is important that 10 percent reproduce. So it is not necessarily for the survival of the species that most of them reproduce, or that members of the species *in general* reproduce.

Indeed, human beings could get by indefinitely if only 5 percent of us engaged in reproduction (provided that each reproductive female worked at it full time).

Similarly for distant perception. Suppose 51 percent of humans became deaf, blind, and unable to smell. This would be very bad for those individuals, but surely they could reproduce. Even if they were somehow unable to reproduce, their difficulties might not interfere with the reproductive efforts of the other 49 percent, who could have somewhat larger families if need be. The species might continue to exist for thousands of years. Thus, far sensation is not a psychic power on Matthew's definition.

It goes almost without saying that our species could continue to exist even if all of us lost the distinctively human power to reason. Suppose we all behaved as irrationally as ponies. We would probably be somewhat better off, and we surely would be able to have lots of children.

Other properties may be vital functions according to Matthews's analysis. For example, some evergreens have this feature: the seeds are in cones. The cones burst open only if exposed to very considerable heat. The heat is produced by forest fires. The forest fires occur only if the old trees burn. Consider this property: being able to burn. If most evergreens of the relevant sort lacked this property, the species would die out. So it is a psychic property of the species. So any still-combustible tree is alive. But some dead trees are still combustible.

A deeper problem with Matthews's proposal is that it may be circular. As we saw above, Matthews defines a vital function as one without which a species will not long be *preserved.* What do we mean when we say that a species is "preserved"? Surely not this:

D9: S is preserved up to t =df. at t, S still exists.

The problem is that if species are such things as the property of being a tiger, then they exist necessarily (I think). Instances need do nothing in order to keep the species in existence. Surely Matthews didn't mean that. What then? Presumably this:

D10: S is preserved up to t =df. at t, there are still some living members of S.

But if this is what 'preserved' means, then Matthews's approach as a whole is clearly circular. We propose to analyze the concept of life by appeal to the concept of vital property; we analyze the concept of vital property by appeal to the concept of preservation; we analyze the concept of preservation by appeal to the concept of life. On the other hand, he might try:

D11: S is preserved up to t =df. at t, there are still some existing members of S.

But then you don't have to be alive to preserve your species. And then a species could be preserved even if none of its members were alive—all we would need would be suitable formaldehyde baths for at least some members of the species.

Conclusion

In spite of its magnificent pedigree and its popularity, the life-functional approach to the analysis of life is unsuccessful. I see no satisfactory way to define life by appeal to some set of life functions. Let us then turn to some other approaches to the analysis of life itself.

3

Vitalist Theories of Life

I want to understand the nature of death itself. In other words, I want to discover a suitable philosophical analysis of the concept of death. According to the most popular view, death is the cessation of life. It appears then that in order to understand death properly, we must first understand life. So I have been considering some proposed analyses of the concept of life. In Chapter 2, I presented and discussed several versions of the Aristotelian idea that life can be explained by appeal to the life functions. None of these seemed successful. I know of no more plausible variant of the life-functional approach. In this chapter, I turn to a consideration of some other approaches to the analysis of life. These are vitalism, DNA-ism, and the genetic informational theory of life.

Vitalism

Aristotle is the first and perhaps greatest advocate of the life-functional approach to the analysis of life. Strangely enough, he is also the first and perhaps greatest advocate of a wholly distinct approach to the analysis of life. This second approach is generally known as "vitalism."

In order to explain the distinguishing characteristic of vitalism, we have to appeal to the distinction between "substances" and "attributes." Unfortunately, the substance/attribute distinction is notoriously difficult to draw. I will not be able to draw it very precisely here. Perhaps I can hint at the distinction in this way: such things as my body, the blood in my body, the air in my lungs, the earth, and the like, are *substances*. The shape of my body, the

color of my blood, the temperature of the air, the weight of the earth, and the like, are *attributes*, or properties.

Substances are also known as objects, individuals, things, particulars. Attributes are also known as properties, characteristics, features, or universals. A given substance, such as my body, for example, is said to *exemplify* various attributes, such as being six feet tall, weighing over one hundred and fifty pounds, being at a temperature of approximately 98.6°F. These attributes are said to *characterize* this substance.

It is natural to think that living things differ from nonliving things fundamentally because living things have some special organization or capacity—that being alive is primarily a matter of having "vital properties." Thus, in this view, life itself is an attribute that living things have because they also have certain other attributes. Vitalists, however, disagree. They think living things differ from nonliving things in virtue of the fact that living things contain a special substance, which can be called their "life." Thus, for the vitalists, life itself is a substance rather than an attribute. They think that living things are alive because they contain some of this substance. Aristotle used the term 'psyche' (or 'soul') apparently to refer to this substance. In a famous passage in *De Anima* ("On the Soul"), Aristotle says: "[W]hat has soul in it differs from what has not in that the former displays life."[1]

If this is meant to be a suggestion of a definition of 'alive', and 'soul' is meant to refer to a substance rather than an attribute, then this passage contains an ancient formulation of the vitalist conception of life. According to this view, to say that a thing is alive is to say that it contains a certain object, a "soul."

Later vitalists characterized life in other ways. One of the most popular forms of vitalism is based on the idea that life is a fluid—"vital fluid." Hans Driesch, one of the best known of modern vitalists, insisted that life is some sort of substance, but he wrote in such an impenetrable style that it is nearly impossible to figure out what sort of substance he took it to be.[2] Perhaps in part because it was generally presented in such obscure terminology, vitalism has passed from the scene.

Vitalism is dead. Since there is no point in beating a dead horse, it may seem that there is no point in trying to refute vitalism. However, my main purpose here is not to refute vitalism but to engage in

a sort of postmortem examination. I want to consider why vitalism died. Having done this, I will turn to some living alternatives, to consider the state of health of vitalism's descendants.

During its life, vitalism took many forms. It will be necessary at the outset, therefore, to sketch one relatively clear version of the view.

At the heart of our form of vitalism is the commitment to "vital fluid." Vital fluid is an "imponderable substance." When we say, in this context, that vital fluid is a "substance," we imply that it is a kind of "stuff" rather than, for example, a force, a property, or a set of dispositions. Perhaps it would also be correct to say that each unit of vital fluid is supposed to be a concrete, individual thing rather than a property or attribute. Thus, a drop of vital fluid belongs in the same ontological category as a drop of blood, a drop of cerebrospinal fluid, or a drop of transmission fluid.

But vital fluid is different from transmission fluid in several important respects. One of the most important of these is that vital fluid is supposed to be "imponderable." To say that it is imponderable is to say that it has no weight. In this respect, vital fluid is like ether and phlogiston. Furthermore, vital fluid is colorless, tasteless, odorless, and in general unobservable. We can recognize that it is present (if at all) only by the consequences of its presence.

Let us also agree that from the perspective of chemistry and physics, vital fluid is not an ordinary liquid or gas. It does not have any chemical structure. It is not made of atoms and molecules. In light of this, we have to recognize that it would not be possible to synthesize vital fluid in a laboratory. At best, it might be possible to *collect* some. No one will ever be able to *make* any—at least, no one will be able to make any *by ordinary chemical processes*.

According to our version of vitalism, every living thing contains at least some vital fluid. In the case of human beings, vital fluid is contained in the sperm. When the sperm and egg unite, the vital fluid "vitalizes" the resulting embryo, and it is therefore alive. The living embryo has the capacity to enlarge its original portion of vital fluid. As the embryo grows, a portion of vital fluid is incorporated into every new cell. Eventually, the mature, living individual is permeated with vital fluid—every living cell in the body contains a tiny drop. As death approaches, the vital fluid begins to leave the body. In some cases, it may gradually ooze away as the various

formerly living parts of the body become moribund and start to die. At some point, the body contains so little vital fluid that it can no longer function. We say that it is dead. After a few days, perhaps no vital fluid is left. The body is thoroughly dead—no part of it remains alive.

The version of vitalism that I mean to discuss here is not an empirical hypothesis about living things. That is, it is not just the contingent claim that, as a matter of fact, every living thing happens to contain some vital fluid. Rather, it is a theory about "life itself." That is, it is a theory about the nature of life, or the property of being alive. In its simplest form, the theory can be stated as a definition:

V1: x is alive at t =df. there is some vital fluid in x at t.

If V1 is correct, then it is impossible for something to be alive without vital fluid, and it is impossible for something to contain vital fluid without being alive. To say that something is alive is, in this view, simply to say that it contains vital fluid. This is not merely a contingent, empirical hypothesis.

The Empirical Problem

One of the most serious problems for V1 arises from the fact that there probably is not any vital fluid. If there is not any vital fluid, then, according to V1, nothing is alive.

Of course, in light of the fact that vital fluid is supposed to be unobservable, it is not easy to determine whether any exists. Driesch apparently performed a number of experiments on sea urchin eggs in an effort to establish the existence of vital fluid.[3] In one experiment, he allowed a fertilized sea urchin egg to undergo one cell division. Then he separated the two cells. Each cell continued to grow and eventually developed into a full-fledged (but pint-sized) sea urchin. Driesch declared that the experiment had succeeded—it showed, he said, that the original fertilized sea urchin egg contained vital fluid. Driesch's reasoning here eludes me. I can not see any connection between the facts about the sea urchins and the existence of vital fluid. The experiment seems to me to be utterly pointless.

The Jonah Problem

There is an another obvious and unanswerable objection to V1. This objection conclusively refutes V1 even if every living thing contains the appropriate amount of vital fluid. A trivial variant of the story of Jonah and the whale will make the objection clear. Suppose Jonah is swallowed by the whale, just as the Bible says. Suppose, however, contrary to what the Bible says, that as a result of eating Jonah, the whale suffers a huge bout of indigestion and dies. Suppose several days pass. Jonah remains very lonely and frightened but fully alive inside the whale. The whale becomes utterly dead. Suppose, in order to give the story a not-too-grisly ending, that the whale washes up on a beach, some fishermen hear Jonah's cries and cut open the whale, thereby freeing Jonah at last. Perhaps Jonah's ordeal teaches him that he cannot escape God's commands, and so he goes off to Nineveh to preach to the Ninevites as God originally told him to do.

For present purposes, the crucial period of time is the period during which the whale is dead and Jonah is alive inside the whale's carcass. According to V1, every living thing contains vital fluid. If V1 were true, it would follow that Jonah contained vital fluid during his captivity. Since Jonah was entirely inside the whale during this period, it follows that there was some vital fluid (Jonah's) inside the whale during the period of Jonah's captivity. If V1 were true, therefore, it would follow that the whale was alive during that period. But it wasn't. So V1 is false.

Perhaps someone will say that Jonah's vital fluid was not really "inside" the whale during his captivity. It might be insisted that for something to be truly "inside" the whale, it has to be inside the cells in the whale's body—merely being in the whale's belly would be insufficient. This suggests:

V2: x is alive at t =df. there is some vital fluid inside the cells of x at t.

Since Jonah's vital fluid was not inside the cells of the whale's body, V2 generates more plausible results in the Jonah case. However, V2 is hardly an improvement over V1. Suppose the whale died because of an infection. Suppose the still-living infectious agents are located inside some of the whale's cells. Then the vital fluid in the infectious agents is really "inside" the whale—it is inside his cells. But the whale is still dead.

These examples show quite conclusively that V1 and V2 are false. To state a more plausible version of vitalism, we might try to distinguish between cases (such as the one involving the whale and Jonah's vital fluid), in which some vital fluid is in an object in the "wrong way," and other more typical cases (such as the one involving Jonah and Jonah's vital fluid) in which the vital fluid is in the object in the "right way"—the way that serves to make the thing alive.

We can say that vital fluid "animates" an organism when it is in the organism in the "right way." Thus, during his captivity, Jonah's vital fluid animated Jonah, but it did not animate the whale. Now, instead of V1, we can consider:

V3: x is alive at t =df. some vital fluid animates x at t.

Since we have not even tried to explain what is meant by 'animates', V3 suffers from a certain amount of obscurity, but the cases I have described do not refute it.

The real problem with V3, as I see it, is that the appeal to vital fluid now begins to seem irrelevant. Perhaps the difficulty can be brought to the surface by comparing V3 with a nonvitalistic competitor:

V4: x is alive at t =df. something animates x at t.

Suppose some sea urchin were animated by ether, or by phlogiston, or by some other imponderable fluid, instead of by vital fluid. So long as it were truly *animated,* it would behave in every way like a living thing. We could never tell that it was animated by the wrong substance. But what is most important in the present circumstances is that there seems to be no plausibility to the claim that such a sea urchin really would not be alive. We would all say that it

is alive, and I think we would be right. In light of this, it is hard to see what difference the vital fluid makes. Animation (whatever that may be) seems to be the crucial factor. So the nonvitalistic V4 is just as plausible as the vitalistic V3. Each, of course, is hopelessly obscure.

The Failure of Analyticity

Suppose a flying saucer lands on the Mall in Washington. A door opens, and a little green man walks out. Suppose he breathes air, eats hamburgers and french fries, and talks to us about his hopes and fears. Suppose also he shows us photos of his wife and children back home on Mars. We would all say that he is alive, and we would be literally right.

However, it is consistent with all this to suppose that life on Mars developed in a genuinely "unearthly" way. Suppose the Martian allows some scientists to take a tiny sliver of his skin for laboratory study. Suppose the scientists discover that there is no vital fluid in the Martian skin sample. Perhaps the Martian's cells contain "zoetic fluid" instead of vital fluid. (How this could be determined is another matter.) The scientists say:

1. The Martian is alive, but he contains no vital fluid.

It is important to notice that the scientists' statement, (1), is not self-contradictory. Such a statement *could* be true. That is, there is no conceptual problem in the supposition that something is alive but has no vital fluid. This shows that 'alive' does not entail 'has vital fluid'. Hence, "life itself" does not involve vital fluid. Even if every living thing on earth contains vital fluid, the property we ascribe to such things when we say that they are alive is not properly defined by V1, V2, or any other vitalistic definition. The case of the zoetic Martian shows that the property of being alive cannot be identified with the property of containing, or being animated by, vital fluid. We can easily imagine a case in which something is alive even though it contains no vital fluid.

So there are several reasons for rejecting vitalism. In the first place, there is no reason to suppose that there is any vital fluid.

This is the empirical problem. In the second place, even if there is vital fluid, it is hard to state in any useful way how the vital fluid is supposed to bear on the life of living organisms. Consideration of the story of Jonah demonstrated the difficulty. Finally, there is the failure of analyticity. The case of the zoetic Martian shows that there is no conceptual link between being alive and having vital fluid. With these reflections in mind, let us now turn to a consideration of a much more respectable view. According to this view, which we can call "DNA-ism," the essence of life is the containment of DNA or RNA.

DNA-ism

In *The Growth of Biological Thought,* Ernst Mayr seems to maintain a somewhat confusing combination of positions on the question concerning the definability of life itself. On the one hand, in a passage I quoted earlier, he asserts that attempts to define 'life' "are rather futile, since it is now quite clear that there is no special substance, object, or force that can be identified with life."[4] On the other hand, he says that "The process of living, however, can be defined. There is no doubt that living organisms possess certain attributes that are not or not in the same manner found in inanimate objects."[5] Mayr lists eight attributes that are allegedly characteristic of living things. One of these attributes is the possession of a genetic program. Mayr says:

> All organisms possess a historically evolved genetic program, coded in the DNA of the nucleus of the zygote (or in RNA in some viruses). Nothing comparable to it exists in the inanimate world, except for manmade computers. . . . Except for the twilight zone of the origin of life, the possession of a genetic program provides for an absolute difference between organisms and inanimate matter.[6]

Mayr does not say that the "process of living" can be defined simply by appeal to DNA. Seven other characteristics of living things are also mentioned. Yet his remarks clearly entail some sort of strong conceptual link between life and the possession of DNA or RNA. Let us consider a non-Mayrian form of DNA-ism—the

view that life can be defined strictly by appeal to the containment of DNA.

This view may be stated as a new definition of life itself:

DNA1: x is alive at t =df. x contains some DNA or RNA at t.

One of the serious objections to vitalism is that there is no evidence to support the claim that there is vital fluid. Earlier, I called this "the empirical problem." DNA1 faces no comparable difficulty. There is overwhelming empirical evidence that DNA and RNA actually exist. My older daughter once worked in a biology laboratory. Together with others, one of her jobs was to extract DNA from ground-up mouse spleens. One day she brought home a little test tube containing a few drops of purified DNA. So DNA exists. I've seen it with my own eyes. Obviously, then, DNA-ism has certain advantages over vitalism.

However, this little story about my daughter and the test tube provides the basis for a thoroughly devastating objection to DNA1. My daughter's test tube contained DNA. According to DNA1, it would follow that the test tube was alive. But it was not. Therefore, DNA1 is false. The reader is invited to consider the analogy between this objection to DNA1 and the Jonah Problem discussed above in connection with V1. I think the similarity is quite striking.

The natural way to modify the theory is to say that living things differ from nonliving things, not simply by virtue of the fact that they *contain* some DNA or RNA, but that the DNA or RNA is contained *within their cells.* Thus we can modify DNA1 as follows:

DNA2: x is alive at t =df. DNA or RNA is contained in the cells of x at t.

Since my daughter's test tube was not made of cells, it would not be correct to say that its cells contained DNA. Thus, DNA2 yields more satisfactory results in that case. Nevertheless, DNA2 is obviously wrong. It is vastly too generous with life. Consider any reasonably fresh corpse. Its cells contain just as much DNA as they contained when the organism was alive, yet it is now dead. According to DNA2, the corpse is alive. But it isn't. Thus, DNA2 is false.

Cases such as this last one concerning the fresh corpse show that

vitalism has certain advantages over DNA-ism. It is open to the vitalist to insist that all the vital fluid has gone out the fresh corpse. Since vital fluid is unobservable, no one can prove the vitalist wrong. But the defender of DNA-ism has no such "out." DNA is observable. Unfortunately for DNA-ism, it can readily be observed in dead organisms.

There are a number of ways in which we might try to modify DNA2. One follows the pattern of V3. Instead of saying (as we do in DNA2) that DNA or RNA is merely "contained in" the cells of x, we can require that the DNA or RNA be more productive—we can require that x be *animated* by the DNA or RNA in its cells. This would yield:

DNA3: x is alive at t =df. x is animated at t by some DNA or RNA in x's cells.

One could reasonably maintain that DNA3 does not run afoul of the corpses in the morgue. Although their cells contain plenty of DNA, we can always insist that their DNA fails to animate them. Thus, DNA3 yields the correct results in such cases.

Nevertheless, DNA3 faces a very serious difficulty. The difficulty is quite like the difficulty that faced V3. 'Animate' is a pretty fancy word. What does it mean? Let us consider three main possibilities.

1. My dictionary (*The Oxford Paperback Dictionary*, 1983) says that 'animate' means 'to give life to'.[7] In other words:

A1: x animates y at t =df. x gives life to y at t.

For certain purposes, this might be a fine dictionary definition of 'animate'. However, we cannot make use of it in the present context. We are trying to give an account of the concept of life. The proposal currently under consideration is DNA3. It makes use of 'animate'. If, in this context, we define 'animate' as we do in A1, the project as a whole becomes hopelessly circular. We define life by appeal to animation, and animation by appeal to life. Obviously, the combination of DNA3 and A1 sheds no light on the nature of life.

2. A second proposal concerning animation would involve an appeal to the "life functions." These are such activities as nutrition, reproduction, motion, and thought. The precise details of the list of life functions should not detain us here, since they were discussed in Chapter 2. The important point is that we give some independent account of the life functions, and then we say:

A2: x animates y at t =df. x causes y to be able to perform the life functions at t.

It should be obvious that the combination of DNA3 and A2 is a failure. In the first place, as I attempted to establish in Chapter 2, no one has been able to give a satisfactory account of the life functions. However, if someone were to give such an account, it would make DNA3 and A2 pointless. We could define life directly by appeal to the life functions, and skip the business about DNA as well as the business about animation. What I have in mind, of course, is just this:

L1: x is alive at t =df. x is able to perform the life functions at t.

3. A third proposal concerning animation would involve taking this concept as a conceptual primitive—that is, we propose that animation is both indefinable and familiar. In virtue of its indefinability, we are absolved of the responsibility of providing a definition. In virtue of its familiarity, we are permitted to make use of it in the effort to define other, less familiar concepts.

Clearly, however, we cannot take animation as a primitive here. That would be tantamount to taking life itself as a primitive. In other words, it would be tantamount to admitting that we cannot discover an analysis of the concept we set out to analyze.

Earlier, in connection with V3, I presented what I called "the analyticity problem." I pointed out that the sentence "The Martian is alive, but contains no vital fluid" is not self-contradictory. Even if all and only actual living things contain vital fluid, this shows that V3 is unacceptable. Let us now consider whether a similar objection can be raised against DNA3.

Consider this scenario: A flying saucer lands on the Mall in Washington, D.C. A door opens and out steps a little green man. He breathes air; he eats hamburgers and french fries; he speaks to us about his hopes and fears. He shows us photos of his wife and

children back home on Mars. As an expression of interplanetary scientific generosity, he allows some scientists to remove a sliver of skin for analysis in the laboratory. To their surprise, they find that the Martian's skin cells contain no DNA or RNA. Perhaps they find that his cells contain a substance never before seen on earth—"ZNA". They say:

> 2. The Martian is alive, but he contains no DNA or RNA.

It seems to me that (2) is not self-contradictory. It *could* be true. Therefore, it is metaphysically possible for a thing to be alive even though it does not contain DNA or RNA. This shows that the concept of life does not involve the concept of containment of DNA or RNA. Any definition along the lines of DNA1 or DNA2 has got to be wrong.

It may be interesting to note that Francis Crick, one of the co-discoverers of the structure of DNA, commented on this very point. In his book *Life Itself,* in the chapter "The General Nature of Life," he discusses the idea that there may be forms of life utterly different from everything we know here on earth. He describes what he takes to be some of the necessary conditions for life (these involve storage and replication of genetic information). He points out that the earthly system, based on chemical features of the carbon atom, is well suited to these tasks. But he goes on to say:

> Of course, elsewhere in the universe life may exist based on other materials. At lower temperatures liquid ammonia might serve as the solvent, though it is not as versatile a solvent as water, which is an exceptionally good one. Instead of carbon, silicon has been suggested. . . . Thus, a form of life based on other materials is not impossible.[8]

It appears, then, that Crick recognized that there is no conceptual connection between life and DNA containment. He saw that statements similar to (2) might be true. Thus, he would not endorse any definition of life that entails that life essentially involves DNA.

As we saw above, Mayr mentioned eight attributes in his attempt to "define the process of living." It appears that Mayr intended his "definition" to be conjunctive in form. Thus, he might endorse some definition according to which 'x is alive at t' means the same as some eightfold conjunction whose final conjunct mentions the containment of DNA or RNA.

Any such definition yields the result that there is a necessary connection between life and DNA- or RNA-containment. If such a definition were correct, it would be *necessary* that every living thing (or its zygote) contain DNA or RNA. On conceptual grounds alone, DNA or RNA containment would be a necessary condition of life. Crick's reflections concerning life on other planets show that this is wrong. While life as we know it seems to involve DNA or RNA, life itself does not. There is no conceptual connection between being alive and having DNA or RNA.

We saw above that vitalism faces three main objections. In the first place, it runs into difficulty because it is doubtful that there is any vital fluid. We called this "the empirical problem." DNA-ism has no similar difficulty. The second problem for vitalism was the Jonah problem. If there were vital fluid, merely *containing* it would not be sufficient for life. We found that DNA-ism suffers from a variant of the Jonah problem. The mere *containment* of DNA is also not sufficient for life—a nonliving test tube might *contain* DNA. A fresh corpse contains plenty of it. It is very hard to describe in any clear or helpful way how life is related to DNA. The third objection to vitalism was the analyticity problem. There seems to be no conceptual link between life itself and the containment of vital fluid. This objection applies equally to DNA-ism. There is no conceptual link between life itself and the containment of DNA. We can easily understand what would be meant by saying that something is alive but contains no DNA.

Genetic Informationism

In another passage in *Life Itself,* Crick presents a more positive suggestion concerning the nature of life. He says that life itself may be identified with the conjunction of five features involving "genetic information." The five features are these: (1) each living thing contains some genetic information that it is able to repro-

duce, together with "machinery needed for execution"; (2) when it produces replications, it will do so with a high degree of accuracy, but not perfectly; (3) the genetic information and the objects (on earth, it is cells) thereby created are held in close contact; (4) the system is open, and able to receive some raw material for its chemical operations; and (5) it is also able to receive energy from the environment.

It seems to me that Crick's view is a combination of a pure genetic information theory (like the one suggested by Mayr in the passage cited above) and some version of the metabolism theory (similar to Aristotle's). If we emphasize conditions (1) and (2), we get something quite reminiscent of what Mayr said.

All this is quite interesting and surely contains more than just a grain of truth, but what precisely is the view? One serious problem concerns the notions of "instructions," "information," and the like. What do we mean when we say that the genetic material in a living cell contains "information" about the cell? I think it is clear that one thing we do not mean is "semantic information." A book might contain semantic information about a cell. It would do so if it contained sentences that expressed facts about that cell. The genetic material in a cell does not in this way express facts about the cell.

Some information seems to be "non-semantic." Consider a footprint in the mud. It might contain lots of information about the person who left it. From its depth, we might infer the weight of the person; from the shape of the print, we might infer the sort of shoe the person was wearing and that it was a man's shoe; from the location of the print, we might infer that he was here; from the amount of stuff that has fallen into the footprint, we may infer how long ago he was here; and so on. In general, then, one thing contains information about another when properties of the first are related to properties of the second in such a way that observations of properties of the first enable sufficiently well-informed persons to infer corresponding properties of the second.

What about "instructions"? Strictly speaking, instructions are semantic entities. Instructions for creating a cell would be statements such as: "first take some carbon atoms; then add some oxygen; then mix that with a bit of nitrogen; and so on." The genetic material in a cell does not contain any literal instructions.

Once again, however, we can give sense to the concept of "nonsemantic instructions." Suppose we want the lights to go on when it gets dark, and to go off again when it gets light. We might hire a lightkeeper, and give him or her suitable semantic instructions. On the other hand, we might rig up a photosensitive switch. When sufficient sunlight shines on the switch, a current is induced, a relay is tripped, the circuit is broken, and the lights go off. When the sun goes down, insufficient current goes across the circuit, a spring trips the relay, and the lights go back on. This is a purely causal process—there is nothing semantic about it—but we might want to say that the photosensitive switch acts in accord with the "instructions" to turn the lights on when there is insufficient sunlight; to turn them back off when there is sufficient sunlight.

A more complicated sort of switch might be made in such a way as to act in accord with more complex "instructions" such as these: when there is insufficient sunlight, then turn the lights on, except if they have been off for less than a minute. Maybe it would be possible to characterize a program as a sufficiently complex sequence of nonsemantic instructions such as these.

What about the genetic material in living cells? It contains nonsemantic information about many of the important features of the cell itself and cells that could be produced by replication from that cell. It also contains nonsemantic information about the organism as a whole from which it has been taken. It also contains nonsemantic instructions about the productions of replicated cells, and whole reproduced organisms. All of this information and instruction is contained, among earthly organisms, in the DNA or RNA.

Obviously, however, it would be a mistake to try to define life by appeal to DNA-containment. As we saw in above, most organisms contain just as much DNA when they are dead as when they were alive. Furthermore, as Crick points out, "a form of life based on other materials is not impossible."[9] Thus, for all we know, living Martians may contain ZNA rather than DNA or RNA. It is more plausible to suppose that the idea is this: living things contain tiny parts that contain nonsemantic information and instructions concerning the construction of near replicas of themselves; furthermore, living things are able to make use of this information in the production of such near replicas of themselves.

Let us say that an object contains a "genetic representation" of itself just in case it contains some tiny parts that encode some nonsemantic information and instructions concerning the construction of near replicas. Thus, since we contain DNA, and DNA encodes such information and instructions, we contain genetic representations of ourselves. If there are Martians, and they contain ZNA, then they also contain genetic representations of themselves. In this way, we focus on the *function* of DNA rather than on the particular chemical features of the stuff that serves that function in us.

Now we can formulate a theory about life itself based on these ideas:

GI: x is alive at t =df. x contains a genetic representation of x at t.

I think that GI generates correct results in a wide variety of cases, and that it is close in spirit to a central component of the proposals suggested by Mayr and Crick in the passages cited.

Problems for Genetic Informationism

Any living thing that is capable of reproducing is alive according to D3. However, I believe that many dead things are counted as alive by this definition. Consider a dead female frog. Suppose this frog contains a lot of still-viable eggs. If the eggs are removed from the frog, they can be fertilized, and they can then grow into "near replicas" of the original frog. In this case, though the frog is clearly dead, it still contains genetic representations of itself. Thus, GI implies, incorrectly, that it is alive.

Similarly, consider what happens in the gardens of careless gardeners. They may neglect to remove their tomato plants at the end of the season. Frost will kill the plants. During the winter, the plants stand dead and blackened against their stakes. Miserable-looking, shriveled tomatoes may hang from the branches. In those tomatoes are many still-viable seeds. Though the plants are clearly dead, they therefore still contain genetic representations of themselves. In the spring, seeds from the dead tomatoes may sprout

into near replicas of the now-dead parents. Thus, GI is clearly false.

While there are still other proposed analyses of the concept of life itself, I think I have discussed most of the interesting variants. Each of them has proved unsatisfactory. Other theories, it seems to me, are no more successful. Thus, my view is that life is a mystery. Though professional biologists may have huge amounts of very detailed information about typical features of living things here on earth, it appears that no one has succeeded in formulating a satisfactory philosophical analysis of the concept of life itself.