THE ANTHROPIC COSMOLOGICAL PRINCIPLE

JOHN D. BARROW

Lecturer, Astronomy Centre, University of Sussex

and

FRANK J. TIPLER

Professor of Mathematics and Physics, Tulane University, New Orleans

With a foreword by John A. Wheeler

CLARENDON PRESS · *Oxford*
OXFORD UNIVERSITY PRESS · *New York*
1986

Oxford University Press, Walton Street, Oxford OX2 6DP

Oxford New York Toronto
Delhi Bombay Calcutta Madras Karachi
Kuala Lumpur Singapore Hong Kong Tokyo
Nairobi Dar es Salaam Cape Town
Melbourne Auckland

and associated companies in
Beirut Berlin Ibadan Nicosia

Oxford is a trade mark of Oxford University Press

Published in the United States
by Oxford University Press, New York

British Library Cataloguing in Publication Data
Barrow, John D.
The anthropic cosmological principle.
1. Man
I. Title II. Tipler, Frank J.
128 BD450
ISBN 0-19-851949-4

Library of Congress Cataloging in Publication Data
Barrow, John D., 1952—
The anthropic cosmological principle.
Bibliography: p.
Includes index.
1. Cosmology. 2. Man. 3. Teleology. 4. Intellect.
5. Life on other planets. 6. Science—Philosophy.
I. Tipler, Frank J. II. Title.
BD511.B34 1985 113 85-4824
ISBN 0-19-851949-4

Printing (last digit): 9 8 7 6 5 4 3 2 1

Printed in the United States of America

To Elizabeth and Jolanta

Ah Mr. Gibbon, another damned, fat, square book. Always scribble, scribble, scribble, eh?
THE DUKE OF GLOUCESTER

[on being presented with volume 2 of *The Decline and Fall of the Roman Empire*]

1 Introduction

The Cosmos is about the smallest hole that a man can hide his head in
G. K. Chesterton

1.1 Prologue

What is Man, that Thou art mindful of him?
Psalm 8:4

The central problem of science and epistemology is deciding which postulates to take as fundamental. The perennial solution of the great idealistic philosophers has been to regard Mind as logically prior, and even materialistic philosophers consider the innate properties of matter to be such as to allow—or even require—the existence of intelligence to contemplate it; that is, these properties are necessary or sufficient for life. Thus the existence of Mind is taken as one of the basic postulates of a philosophical system. Physicists, on the other hand, are loath to admit any consideration of Mind into their theories. Even quantum mechanics, which supposedly brought the observer into physics, makes no use of intellectual properties; a photographic plate would serve equally well as an 'observer'. But, during the past fifteen years there has grown up amongst cosmologists an interest in a collection of ideas, known as the Anthropic Cosmological Principle, which offer a means of relating Mind and observership directly to the phenomena traditionally within the encompass of physical science.

The expulsion of Man from his self-assumed position at the centre of Nature owes much to the Copernican principle that we do not occupy a privileged position in the Universe. This Copernican assumption would be regarded as axiomatic at the outset of most scientific investigations. However, like most generalizations, it must be used with care. Although we do not regard our position in the Universe to be central or special in every way, this does not mean that it cannot be special in *any* way. This possibility led Brandon Carter[1] to limit the Copernican dogma by an 'Anthropic Principle' to the effect that 'our location in the Universe is necessarily privileged to the extent of being compatible with our existence as observers'. The basic features of the Universe, including such properties as its shape, size, age and laws of change, must be *observed* to be of a type that allows the evolution of observers, for if intelligent life did not evolve in an otherwise possible universe, it is obvious that no one would be asking the reason for the observed shape, size, age and so forth of the Universe. At first sight such an observation might appear true but trivial. However, it has far-reaching implications for physics. It is a restatement, of the fact that any observed properties of the Universe that may initially appear astonishingly improbable, can only be seen in their true perspective after we have accounted for the fact that certain properties of the Universe are necessary prerequisites for the evolution and existence of any observers at all. The measured values of many cosmological and physical quantities that define our Universe are circumscribed by the necessity that we observe from a site where conditions are appropriate for the occurrence of biological evolution and at a cosmic epoch exceeding the astrophysical and biological timescales required for the development of life-supporting environments and biochemistry.

What we have been describing is just a grandiose example of a type of intrinsic bias that scientists term a 'selection effect'. For example, astronomers might be interested in determining the fraction of all galaxies that lie in particular ranges of brightness.[2] But if you simply observe as many galaxies as you can find and list the numbers found according to their brightness you will not get a reliable picture of the true brightness distribution of galaxies. Not all galaxies are bright enough to be seen or big enough to be distinguished from stars, and those that are brighter are more easily seen than those that are fainter, so our observations are biased towards finding a disproportionately large fraction of very bright galaxies compared to the true state of affairs. Again, at a more mundane level, if a ratcatcher tells you that all rats are more than six inches long because he has never caught any that are shorter, you should check the size of his traps before drawing any far-reaching conclusions about the length of rats. Even though you are most likely to see an elephant in a zoo that does not mean that all elephants are in zoos, or even that most elephants are in zoos. In section 1.2 we shall restate these ideas in a more precise and quantitative form, but to get the flavour of how this form of the Anthropic Principle can be used we shall consider the question of the size of the Universe to illustrate how our own existence acts as a selection effect when assessing observed properties of the Universe.

The fact that modern astronomical observations reveal the visible Universe to be close to fifteen billion light years in extent[3] has provoked many vague generalizations about its structure, significance and ultimate purpose. Many a philosopher has argued[4] against the ultimate importance of life in the Universe by pointing out how little life there appears to be compared with the enormity of space and the multitude of distant galaxies. But the Big Bang cosmological picture shows this up as too simplistic a judgement. Hubble's classic discovery[5] that the Universe is in a dynamic state of expansion reveals that its size is inextricably bound up with its age.[6] The Universe is fifteen billion light years in size because it is fifteen billion years old. Although a universe the size of a single galaxy would contain enough matter to make more than one hundred billion stars the size of our Sun, it would have been expanding for less than a single year.

We have learned that the complex phenomenon we call 'life' is built upon chemical elements more complex than hydrogen and helium gases. Most biochemists believe that carbon, on which our own organic chemistry is founded, is the only possible basis for the *spontaneous* generation of life. In order to create the building blocks of life—carbon, nitrogen, oxygen and phosphorus—the simple elements of hydrogen and helium which were synthesized in the primordial inferno of the Big Bang must be cooked at a more moderate temperature and for a much longer time than is available in the early universe.[7] The furnaces that are available are the interiors of stars. There, hydrogen and helium are burnt into the heavier life-supporting elements by exothermic nuclear reactions. When stars die, the resulting explosions which we see as supernovae, can disperse these elements through space and they become incorporated into planets and, ultimately, into ourselves. This stellar alchemy takes over ten billion years to complete. Hence, for there to be enough time to construct the constituents of living beings, the Universe must be at least ten billion years old and therefore, as a consequence of its expansion, at least ten billion light years in extent. We should not be surprised to observe that the Universe is so large. No astronomer could exist in one that was significantly smaller. The Universe needs to be as big as it is in order to evolve just a single carbon-based life-form.

We should emphasize that this selection of a particular size for the universe actually does *not* depend on accepting most biochemists' belief that only carbon can form the basis of spontaneously generated life. Even if their belief is false, the fact remains that *we are a carbon-based intelligent life-form which spontaneously evolved on an earthlike planet around a star of G2 spectral type, and any observation we make is necessarily self-selected by this absolutely fundamental fact.* In particular, a life-form which evolved spontaneously in such an environment must necessarily see the Universe to be at least several billion years old and hence see it to be at least several billion light years across. This remains true even if non-carbon life-forms abound in the cosmos. Non-carbon life-forms are not necessarily restricted to seeing a minimum size to the universe, but *we* are. Human bodies are measuring instruments whose self-selection properties *must* be taken into account, just as astronomers *must* take into account the self-selection properties of optical telescopes. Such telescopes tell us about radiation in the visible band of the electromagnetic spectrum, but it would be completely illegitimate to conclude from purely optical observations that all of the electromagnetic energy in the Universe is in the visible band. Only when one is aware of the self-selection of optical telescopes is it possible to consider the possibility that non-visible radiation exists. Similarly, it is essential to be aware of the self-selection which results from our being *Homo sapiens* when trying to draw conclusions about the nature of the Universe. This self-selection principle is the most basic version of the Anthropic Principle and it is usually called the *Weak Anthropic Principle*. In a sense, the Weak Anthropic Principle may be regarded as the culmination of the Copernican Principle, because the former shows how to separate those features of the Universe whose appearance depends on anthropocentric selection, from those features which are genuinely determined by the action of physical laws.

In fact, the Copernican Revolution was initiated by the application of the Weak Anthropic Principle. The outstanding problem of ancient astronomy was explaining the motion of the planets, particularly their retrograde motion. Ptolemy and his followers explained the retrograde motion by invoking an epicycle, the ancient astronomical version of a new physical law. Copernicus showed that the epicycle was unnecessary; the retrograde motion was due to an anthropocentric selection effect: we were observing the planetary motions from the vantage point of the moving Earth.

At this level the Anthropic Principle deepens our scientific understanding of the link between the inorganic and organic worlds and reveals an intimate connection between the large and small-scale structure of the Universe. It enables us to elucidate the interconnections that exist between the laws and structures of Nature to gain new insight into the chain of universal properties required to permit life. The realization that the possibility of biological evolution is strongly dependent upon the global structure of the Universe is truly surprising and perhaps provokes us to consider that the existence of life may be no more, but no less, remarkable than the existence of the Universe itself.

The Anthropic Principle, in all of its manifestations but particularly in its Weak form, is closely analogous to the self-reference arguments of mathematics and computer science.[54] These self-reference arguments lead us to understand the limitations of logical knowledge: Gödel's Incompleteness Theorem demonstrates that any mathematical system sufficiently complex to contain arithmetic must contain true statements which cannot be proven true, while Turing's Halting Theorem shows that a computer cannot fully understand itself. Similarly, the Anthropic Principle shows that the observed structure of the Universe is restricted by the fact that we are observing this structure; by the fact that, so to speak, the Universe is observing itself.

The size of the observable Universe is a property that is changing with time because of the overall expansion of the system of galaxies and clusters. A selection effect enters because we are constrained by the timescales of biological evolution to observe the Universe only after billions of years of expansion have already elapsed. However, we can take this consideration a little further. One of the most important results of twentieth-century physics has been the gradual realization that there exist invariant properties of the natural world and its elementary components which render the gross size and structure of virtually all its constituents quite inevitable.[8] The sizes of stars and planets, and even people, are neither random nor the result of any Darwinian selection process from a myriad of possibilities. These, and other gross features of the Universe are the consequences of necessity; they are manifestations of the possible equilibrium states between competing forces of attraction and repulsion. The intrinsic strengths of these controlling forces of Nature are determined by a mysterious collection of pure numbers that we call the *constants of Nature*.[9]

The Holy Grail of modern physics is to explain why these numerical constants—quantities like the ratio of the proton and electron masses for example—have the particular numerical values they do. Although there has been significant progress towards this goal during the last few years[10] we still have far to go in this quest. Nevertheless, there is one interesting approach that we can take which employs an Anthropic Principle in a more adventurous and speculative manner than the examples of self-selection we have already given.

It is possible to express some of the necessary or sufficient conditions for the evolution of observers as conditions on the relative sizes of different collections of constants of Nature. Then we can determine to what extent our observation of the peculiar values these constants are found to take is necessary for the existence of observers. For example, if the relative strengths of the nuclear and electromagnetic forces were to be slightly different then carbon atoms could not exist in Nature[11] and human physicists would not have evolved. Likewise, many of the global properties of the Universe, for instance the ratio of the number of photons to protons,[12] must be found to lie within a very narrow range if cosmic conditions are to allow carbon-based life to arise.

The early investigations of the constraints imposed upon the constants of Nature by the requirement that our form of life exist produced some surprising results. It was found that there exist a number of unlikely coincidences between numbers of enormous magnitude that are, superficially, completely independent; moreover, these coincidences appear essential to the existence of carbon-based observers in the Universe.[13] So numerous and unlikely did these coincidences seem that Carter proposed[1] a stronger version of the Anthropic Principle than the Weak form of self-selection principle introduced earlier: that the Universe *must* be such 'as to admit the creation of observers within it at some stage.' This is clearly a more metaphysical and less defensible notion, for it implies that the Universe could not have been structured differently—that perhaps the constants of Nature could not have had numerical values other than what we observe. Now, we create a considerable problem. For we are tempted to make statements of comparative reference regarding the properties of our observable Universe with respect to the alternative universes we can imagine possessing different values of their fundamental constants. But there is only one Universe; where do we find the other possible universes against which to compare our own in order to decide how fortunate it is that all these remarkable coincidences that are necessary for our own evolution actually exist?

There has long been an interest in the idea that our Universe is but one of many possible worlds. Traditionally, this interest has been coupled with the naive human tendency to regard our Universe as optimal, in some sense, because it appears superfically to be tailor-made for the presence of living creatures like ourselves. We recall Leibniz' claim that ours is the 'best of all possible worlds'; a view that led him to be mercilessly caricatured by Voltaire as Pangloss, a professor of 'metaphysico-theologo-cosmolo-nigology'. Yet, Leibniz' claims also led Maupertuis to formulate the first Action Principles of physics[14] which created new formulations of Newtonian mechanics and provided a basis for the modern approach to formulating and determining new laws of Nature. Maupertuis claimed that the dynamical paths through space possessing non-minimal values of a mathematical quantity he called the Action would be observed if we had less perfect laws of motion than exist in our World. They were identified with the other 'possible worlds'. The fact that Newton's laws of motion were equivalent to bodies taking the path through space that minimizes the Action was cited by Maupertuis as proof that our World, with all its laws, was 'best' in a precise and rigorous mathematical sense.

Maupertuis' ensemble of worlds is not the only one that physicists are familiar with. There have been many suggestions as to how an ensemble of different hypothetical, or actual' universes can arise.[15,16] Far from being examples of idle scholastic speculation many of these schemes are part and parcel of new developments in theoretical physics and cosmology. In general, there are three types of ensemble that one can appeal to in connection with various forms of the Anthropic Principle and they have rather different degrees of certitude.

First, we can consider collections of different possible universes which are parametrized by different values of quantities that do not have the status of invariant constants of Nature. That is, quantities that can, in

principle, vary even in our observed Universe. For example, we might consider various cosmological models possessing different initial conditions but with the same laws and constants of Nature that we actually observe. Typical quantities of this sort that we might allow to change are the expansion rate or the levels of isotropy and spatial uniformity in the material content of the Universe. Mathematically, this amounts to choosing different sets of initial boundary conditions for Einstein's gravitational field equations of general relativity (solutions of these equations generate cosmological models). In general, arbitrarily chosen initial conditions at the Big Bang do not necessarily evolve to produce a universe looking like the one we observe after more than fifteen billion years of expansion.[17] We would like to know if the subset of initial conditions that does produce universes like our own has a significant intersection with the subset that allows the eventual evolution of life.

Another way of generating variations in quantities that are not constants of Nature is possible if the Universe is infinite, as current astronomical data suggest. If cosmological initial conditions are exhaustively random and infinite then anything that can occur with non-vanishing probability will occur somewhere; in fact, it will occur infinitely often.[18] Since our Universe has been expanding for a finite time of only about fifteen billion years, only regions that are no farther away than fifteen billion light years can currently be seen by us. Any region farther away than this cannot causally influence us because there has been insufficient time for light to reach us from regions beyond fifteen billion light years. This extent defines what we call the '*observable*, (or *visible*), *Universe*'. But if the Universe is randomly infinite it will contain an infinite number of causally disjoint regions. Conditions within these regions may be different from those within our observable part of the Universe; in some places they will be conducive to the evolution of observers but in others they may not. According to this type of picture, if we could show that conditions very close to those we observe today are absolutely necessary for life, then appeal could be made to an extended form of natural selection to claim that life will only evolve in regions possessing benign properties; hence our observation of such a set of properties in the finite portion of the entire infinite Universe that is observable by ourselves is not surprising. Furthermore, if one could show that the type of Universe we observe out to fifteen billion light years is necessary for observers to evolve then, because in any randomly infinite set[19] of cosmological initial conditions there must exist an infinite number of subsets that will evolve into regions resembling the type of observable Universe we see, it could be argued that the properties of our visible portion of the infinite Universe neither have nor require any further explanation. This is an idea that it is possible to falsify by detecting a density of cosmic material sufficient to render the Universe finite. Interestingly, some of the currently popular 'inflationary' theories of how the cosmic medium behaves very close to the Big Bang not only predict that if our Universe is infinite then it should be extremely non-uniform beyond our visible horizon, but these theories also exploit probabilistic properties of infinite initial data sets.

A third class of universe ensembles that has been contemplated involves the speculative idea of introducing a change in the values of the constants of Nature, or other features of the Universe that strongly constrain the outcome of the laws of Nature—for example, the charge on the electron or the dimensionality of space.[23] Besides simply imagining what would happen if our Universe were to possess constants with different numerical values, one can explore the consequences of allowing fundamental constants of Nature, like Newton's gravitation 'constant', to vary in space or time. Accurate experimental measurements are also available to constrain the allowed magnitude of any such variations.[24] It has also been suggested[48] that if the Universe is cyclic and oscillatory then it might be that the values of the fundamental constants are changed on each occasion the Universe collapses into the 'Big Crunch' before emerging into a new expanding phase.

A probability distribution can also be associated with the observed values of the constants of Nature arising in our own Universe in some new particle physics theories that aim to show that a sufficiently old and cool universe must inevitably display apparent symmetries and particular laws of Nature even if none really existed in the initial high temperature environment near the Big Bang. These 'chaotic gauge theories', as they are called,[25] allow, in principle, a calculation of the probability that after about fifteen billion years we see a particular symmetry or law of Nature in the elementary particle world.

Finally, there is the fourth and last class of world ensemble. A much-discussed and considerably more subtle ensemble of possible worlds is one which has been introduced to provide a satisfactory resolution of paradoxes arising in the interpretation of quantum mechanics.[26] Such an ensemble may be the only way to make sense of a quantum cosmological theory. This 'Many Worlds' interpretation of the quantum theory introduced by Everett and Wheeler requires the simultaneous existence of an infinite number of equally real worlds, all of which are more-or-less causally disjoint, in order to interpret consistently the relationship between observed phenomena and observers.

As the Anthropic Principle has impressed many with its apparent novelty and has been the subject of many popular books and articles,[27] it is important to present it in its true historical perspective in relation to the plethora of Design Arguments beloved of philosophers, scientists and theologians in past centuries[28] and which still permeate the popular mind

today. When identified in this way, the idea of the Anthropic Principle in many of its forms can be traced from the pre-Socratics to the founding of modern evolutionary biology. In Chapter 2 we provide a detailed historical survey of this development. As is well known, Aristotle used the notion of 'final causes' in Nature in opposition to the more materialistic alternatives promoted by his contemporaries. His ideas became extremely influential centuries later following their adaption and adoption by Thomas Aquinas to form his grand synthesis of Greek and Judaeo-Christian thought. Aquinas used these teleological ideas regarding the ordering of Nature to produce a Design Argument for the existence of God. Subsequently, the subject developed into a focal point for both expert and inept comment. The most significant impact upon teleological explanations for the structure of Nature arose not from the work of philosophers but rather from Darwin's *Origin of Species*, first published in 1859. Those arguments that had been used so successfully in the past to argue for the anthropocentric purpose of the natural world were suddenly turned upon their heads to demonstrate the contrary: the inevitable conditioning of organic structures by the local environment via natural selection. Undaunted, some leading scientists sought to retain purpose in Nature by subsuming evolutionary theory within a universal teleology.

We study the role played by teleological reasoning in twentieth-century science and philosophy in Chapter 3. There we show also how more primitive versions of the Anthropic Principles have led in the past to new developments in the physical sciences. In this chapter we also describe in some detail the position of teleology and teleonomy in evolutionary biology and introduce the intimate connection between life and computers. This allows us to develop the striking resemblance between some ideas of modern computer theorists, in which the entire Universe is envisaged as a program being run on an abstract computer rather than a real one, and the ontology of the absolute idealists. The traditional picture of the 'Heat Death of the Universe', together with the pictures of teleological evolution to be found in the works of Bergson, Alexander, Whitehead and the other philosophers of progress, leads us into studies of some types of melioristic world-view that have been suggested by philosophers and theologians.

We should warn the professional historian that our presentation of the history of teleology and anthropic arguments will appear Whiggish. To the uninitiated, the term refers to the interpretation of history favoured by the great Whig (liberal) historians of the nineteenth century. As we shall discuss in Chapter 3, these scholars believed that the history of mankind was teleological: a record of slow but continual progress toward the political system dear to the hearts of Whigs, liberal democracy. The Whig historians thus analysed the events and ideas of the past from the point of view of the present rather than trying to understand the people of the past on their own terms.

Modern historians generally differ from the Whig historians in two ways: first, modern historians by and large discern no over-all purpose in history (and we agree with this assessment). Second, modern historians try to approach history from the point of view of the actors rather than judging the validity of archaic world-views from our own Olympian heights. In the opinion of many professional historians, it is not the job of historians to pass moral judgments on the actions of those who lived in the past. A charge of Whiggery—analysing and judging the past from our point of view—has become one of the worse charges that one historian can level at another; a Whiggish approach to history is regarded as the shameful mark of an amateur.[49]

Nevertheless, it is quite impossible for any historian, amateur or professional, to avoid being Whiggish to some extent. As pointed out by the philosopher Morton White,[51] in the very act of criticizing the long-dead Whig historians for judging the people of the past, the modern historians are themselves judging the work of some of their intellectual forebears, namely the Whig historians. Furthermore, every historian must always select a finite part of the infinitely-detailed past to write about. This selection is necessarily determined by the interests of people in the present, the modern historian if no one else. As even the arch critic of Whiggery, Herbert Butterfield, put it in his *The Whig Interpretation of History*:

> The historian is something more than the mere external spectator. Something more is necessary if only to enable him to seize the significant detail and discern the sympathies between events and find the facts that hang together. By imaginative sympathy he makes the past intelligible to the present. He translates its conditioning circumstances into terms which we today can understand. It is in this sense that history must always be written from the point of view of the present. It is in this sense that every age will have to write its history over again.[50]

This is one of the senses in which we shall be Whiggish: we shall try to interpret the ideas of the past in terms a modern scientist can understand.[55] For example, we shall express the concepts of absolute idealism in computer language, and describe the cosmologies of the past in terms of the language used by modern cosmologists.

But our primary purpose in this book is not to write history. It is to describe the modern Anthropic Principle. This will necessarily involve the use of some fairly sophisticated mathematics and require some familiarity with the concepts of modern physics. Not all readers who are interested in reading about the Anthropic Principle will possess all the requisite scientific background. Many of these readers—for instance, theologians

and philosophers—will actually be more familiar with the philosophical ideas of the past than with more recent scientific developments. The history sections have been written so that such readers can get a rough idea of the modern concepts by seeing the parallels with the old ideas. Such an approach will give a Whiggish flavour to our treatment of the history of teleology.

There is a third reason for the Whiggish flavour of our history: we *do* want to pass judgments on the work of the scientists and philosophers of the past. Our purpose in doing so is not to demonstrate our superiority over our predecessors, but to learn from their mistakes and successes. It is essential to take this approach in a book on a teleological idea like the Anthropic Principle. There is a general belief that teleology is scientifically bankrupt, and that history shows it always has been. We shall show that on the contrary, teleology has on occasion led to significant scientific advances. It has admittedly also led scientists astray; we want to study the past in order learn under what conditions we might reasonably expect teleology to be reliable guide.

The fourth and final reason for the appearance of Whiggery in our history of teleology is that there *are* re-occurring themes present in the history of teleology; we are only reporting them. We refuse to distort history to fit the current fad of historiography.

We are not the only contemporary students of history to discern such patterns in intellectual history. Such patterns are particularly noticeable in the history of science: the distinguished historian of science Gerald Holton[52] has termed such re-occurring patterns *themata*. To cite just one example of a re-occurring thema from the history of teleology, the cosmologies of the eighteenth-century German idealist Schelling, the twentieth-century British philosopher Alexander, and Teilhard de Chardin are quite similar, simply because all of these men believed in an evolving, melioristic universe; and, broadly speaking, there is really only one way to constuct such a cosmology. We shall discuss this form of teleology in more detail in Chapters 2 and 3.

In Chapter 4 we shall describe in detail how the modern form of the Anthropic self-selection principle arose out of the study of the famous Large Number Coincidences[29] of cosmology. Here the Anthropic Principle was first employed in its modern form to demonstrate that the observed Large Number Coincidences are necessary properties of an observable Universe. This was an important observation because the desire for an explanation of these coincidences had led Dirac[30] to conclude that Newton's gravitation constant must decrease with cosmic time. His suggestion was to start an entirely new sub-culture in gravitation research. We examine then in more detail the idea that there may exist ensembles of different universes in which various coincidences between the values of fundamental constants deviate from their observed values. One of the earliest uses of the Anthropic self-selection idea was that of Whitrow[31] who invoked it as a means of explaining why space is found to possess three dimensions, and we develop this idea in the light of modern ideas in theoretical physics. One of the themes of this chapter is that the recognition of unusual and suggestive coincidences between the numerical values of combinations of physical constants can play an important role in framing detailed theoretical descriptions of the Universe's structure.

Chapter 5 shows how one can determine the gross structure of all the principal constituents of the physical world as equilibrium states between competing fundamental forces. We can then express these characteristics solely in terms of dimensionless constants of Nature aside from inessential geometrical factors like 2π. Having achieved such a description one is in a position to determine the sensitivity of structures essential to the existence of observers with respect to small changes in the values of fundamental constants of Nature. The principal achievement of this type of approach to structures in the Universe is that it enables one to identify which fortuitous properties of the Universe are real coincidences and distinguish them from those which are inevitable consequences of the particular values that the fundamental constants take. The fact that the mass of a human is the geometric mean of a planetary and an atomic mass while the mass of a planet is the geometric mean of an atomic mass and the mass of the observable Universe are two striking examples.[32] These apparent 'coincidences' are actually *consequences* of the particular numerical values of the fundamental constants defining the gravitational and electromagnetic interactions of physics. By contrast the fact that the disks of the Sun and Moon have virtually the same angular size (about half a degree) when viewed from Earth is a pure coincidence and it does not appear to be one that is necessary for the existence of observers. The ratio of the Earth's radius and distance from the Sun is another pure coincidence, in that it is not determined by fundamental constants of Nature alone, but were this ratio slightly different from what it is observed to be, observers could not have evolved on Earth.[33]

The arguments of Chapter 5 can be used to elucidate the inevitable sizes and masses of objects spanning the range from atomic nuclei to stars. If we want to proceed further up the size-spectrum things become more complicated. It is still not known to what extent properties of the whole Universe, determined perhaps by initial conditions or events close the Big Bang, play a role in fixing the sizes of galaxies and galaxy clusters. In Chapter 6 we show how the arguments of Chapter 5 can be extended into the cosmological realm where we find the constants of Nature joined by several dimensionless cosmological parameters to complete the description of the Universe's coarse-grained structure. We give a detailed

overview of modern cosmology together with the latest consequences of unified gauge theories for our picture of the very early Universe. This picture enables us to interrelate many aspects of the Universe once regarded as independent coincidences. It also enables us to highlight a number of extraordinarily finely tuned coincidences upon which the possible evolution of observers appears to hinge. We are also able to show well-known Anthropic arguments regarding the observation that the Universe is isotropic to within one part in ten thousand are not actually correct.[17]

In order to trace the origin of the Universe's most unusual large scale properties, we are driven closer and closer to events neighbouring the initial singularity, if such there was. Eventually, classical theories of gravitation become inadequate and a study of the first instants of the Universal expansion requires a quantum cosmological model. The development of such a quantum gravitational theory is the greatest unsolved problem in physics at present but fruitful approaches towards effecting a marriage between quantum field theory and general relativity are beginning to be found. There have even been claims that a quantum wave function for the Universe can be written down.[34]

Quantum mechanics involves observers in a subtle and controversial manner. There are several schools of thought regarding the interpretation of quantum theory. These are described in detail in Chapter 7. After describing the 'Copenhagen' and 'Many Worlds' interpretations we show that the latter picture appears to be necessary to give meaning to any wave function of the entire Universe and we develop a simple quantum cosmological model in detail. This description allows the Anthropic Principle to make specific predictions.

The Anthropic Principles seek to link aspects of the global and local structure of the Universe to those conditions necessary for the existence of living observers. It is therefore of crucial importance to be clear about what we mean by 'life'. In Chapter 8 we give a new definition of life and discuss various alternatives that have been suggested in the past. We then consider those aspects of chemical and biochemical structures that appear necessary for life based upon atomic structures. Here we are, in effect, extending the methodology of Chapter 5 from astrophysics to biochemistry with the aim of determining how the crucial properties of molecular structures are related to the invariant aspects of Nature in the form of fundamental constants and bonding angles. To complete this chapter we extend some recent ideas of Carter[35] regarding the evolution of intelligent life on Earth. This leads to an Anthropic Principle prediction which relates the likely time of survival of terrestrial life in the future the number of improbable steps in the evolution of intelligent life on Earth via a simple mathematical inequality.

In Chapter 9 we discuss the controversial subject of extraterrestrial life and provide arguments that there probably exists no other intelligent species with the capability of interstellar communication within our own Milky Way Galaxy. We place more emphasis upon the ideas of biologists regarding the likelihood of intelligent life-forms evolving than is usually done by astronomers interested in the possibility of extraterrestrial intelligence. As a postscript we show how the logic used to project the capabilities of technologically advanced life-forms can be used to frame an Anthropic Principle argument against the possibility that we live in a Steady-State Universe. This shows that Anthropic Principle arguments can be used to winnow-out cosmological theories. Conversely, if the theories which contradict the Anthropic Principle are found to be correct, the Anthropic Principle is refuted; this gives another test of the Anthropic Principle.

Finally, in Chapter 10, we attempt to predict the possible future histories of the Universe in the light of known physics and cosmology. We describe in detail the expected evolution of both open and closed cosmological models in the far future and also stress a number of global constraints that exist upon the structure of a universe consistent with our own observations today. In our final speculative sections we investigate the possibility of life surviving into the indefinite future of both open and closed universes. We define life using the latest ideas in information and computer theory and determine what the Universe must be like in order that information-processing continue indefinitely; in effect, we investigate the implications for physics of the requirement that 'life' never becomes extinct. Paradoxically, this appears to be possible only in a *closed* universe with a very special global causal structure, and thus the requirement that life never dies out—which we define precisely by a new '*Final Anthropic Principle*'—leads to definite testable predictions about the global structure of the Universe. Since indefinite survival in a closed universe means survival in a high-energy environment near the final singularity, the Final Anthropic Principle also leads to some predictions in high-energy particle physics.

Before abandoning the reader to the rest of the book we should make a few comments about its contents. Our study involves detailed mathematical investigations of physics and cosmology, studies of chemistry and evolutionary biology as well as a considerable amount of historical description and analysis. We hope we have something new to say in all these areas. However, not every reader will be interested in all of this material. Our chapters have, in the main, been constructed in such a way that they can be read independently, and the notes and references are collected together accordingly. Scientists with no interest in the history of ideas can just skip the chapters in which they are discussed. Likewise, non-scientists can avoid mathematics altogether they wish. One last word: the authors are cosmologists, not philosophers. This has one very important consequence which the average reader should bear in mind. Whereas philosophers and theologians appear to possess an emotional attachment to their theories and ideas which requires them to believe them, scientists tend to regard their ideas differently. They are interested in formulating many logically consistent possibilities, leaving any judgement regarding their truth to observation. Scientists feel no qualms about suggesting different but mutually exclusive explanations for the same phenomenon. The authors are no exception to this rule and it would be unwise of the reader to draw any wider conclusions about the authors' views from what they may read here.

1.2 Anthropic Definitions

> Definitions are like belts. The shorter
> they are, the more elastic they need to be.
> S. Toulmin

Although the Anthropic Principle is widely cited and has often been discussed in the astronomical literature, (as can be seen from the bibliography to this chapter alone), there exist few attempts to frame a precise statement of the Principle; rather, astronomers seem to like to leave a little flexibility in its formulation perhaps in the hope that its significance may thereby more readily emerge in the future. The first published discussion by Carter[1] saw the introduction of a distinction between what he termed 'Weak' and 'Strong' Anthropic statements. Here, we would like to define precise versions of these two Anthropic Principles and then introduce Wheeler's Participatory Anthropic Principle[6] together with a new Final Anthropic Principle which we shall investigate in Chapter 10.

The Weak Anthropic Principle (WAP) tries to tie a precise statement to the notion that any cosmological observations made by astronomers are biased by an all-embracing selection effect: our own existence. Features of the Universe which appear to us astonishingly improbable, *a priori*, can only be judged in their correct perspective when due allowance has been made for the fact that certain properties of the Universe are necessary if it is to contain carbonaceous astronomers like ourselves.

This approach to evaluating unusual features of our Universe first re-emerges in modern times in a paper of Whitrow[31] who, in 1955, sought an answer to the question '*why does space have three dimensions?*'. Although unable to explain why space actually has, (or perhaps even why it must have), three dimensions, Whitrow argued that this feature of the World is not unrelated to our own existence as observers of it. When formulated in three dimensions, mathematical physics possesses many unique properties that are necessary prerequisites for the existence of rational information-processing and 'observers' similar to ourselves. Whitrow concluded that only in three-dimensional spaces can the dimensionality of space *be* questioned. At about the same time Whitrow also pointed out that the expansion of the Universe forges an unbreakable link between its overall size and age and the ambient density of material within it.[36] This connection reveals that only a very 'large' universe is a possible habitat for life. More detailed ideas of this sort had also been published in Russian by the Soviet astronomer Idlis.[37] He argued that a variety of special astronomical conditions must be met if a universe is to be habitable. He also entertained the possibility that we were observers merely of a tiny fraction of a diverse and infinite universe whose unobserved regions may not meet the minimum requirements for observers that there exist hospitable temperatures and stable sources of stellar energy.

Our definition of the WAP is motivated in part by these insights together with later, rather similar ideas of Dicke[13] who, in 1957, pointed out that the number of particles in the observable extent of the Universe, and the existence of Dirac's famous Large Number Coincidences '*were not random but conditioned by biological factors*'. This motivates the following definition:

Weak Anthropic Principle (*WAP*): *The observed values of all physical and cosmological quantities are not equally probable but they take on values restricted by the requirement that there exist sites where carbon-based life can evolve and by the requirement that the Universe be old enough for it to have already done so.*

Again we should stress that this statement is in no way either speculative or controversial. It expresses only the fact that those properties of the Universe we are able to discern are self-selected by the fact that they must be consistent with our own evolution and present existence. WAP would not necessarily restrict the observations of non-carbon-based life but *our* observations are restricted by our very special nature.

As a corollary, the WAP also challenges us to isolate that subset of the Universe's properties which are *necessary* for the evolution and continued existence of our form of life. The entire collection of the Universe's laws and properties that we now observe need be neither necessary nor sufficient for the existence of life. Some properties, for instance the large size and great age of the Universe, do appear to be necessary conditions; others, like the precise variation in the distribution of matter in the Universe from place to place, may not be necessary for the development of observers at some site. The non-teleological character of evolution by natural selection ensures that *none* of the observed properties of the Universe are sufficient conditions for the evolution and existence of life.

Carter,[35] and others, have pointed out that as a self-selection principle the WAP is a statement of Bayes' theorem. The Bayesian approach[38] to inference attributes *a priori* and *a posteriori* probabilities to any hypothesis before and after some piece of relevant evidence, E, is taken into account. In such a situation we call the before and after probabilities p_B and p_A, respectively. The fact that for any particular outcome O, the probability of observing O before the evidence E is known equals the probability of observing O given the evidence E, after E was accounted for, is expressed by the equation,

$$p_B(O) = p_A(O/E) \tag{1.1}$$

where / denotes a conditional probability. Bayes' formula[38] then gives the relative plausibililty of any two theories α and β in the face of a piece of evidence E as

$$\frac{p_E(\alpha)}{p_E(\beta)} = \frac{p_A(E/\alpha)p_A(\alpha)}{p_A(E/\beta)p_A(\beta)} \tag{1.2}$$

Thus the relative probabilities of the truth of α or β are modified by the conditional probabilities $p_A(E/\alpha)$ and $p_A(E/\beta)$ which account for any bias of the experiment (or experimenter) towards gathering evidence that favours α rather than β (or *vice versa*). The WAP as we have stated it is just an application of Bayes' theorem.

The WAP is certainly not a powerless tautological statement because cosmological models have been defended in which the gross structure of the Universe is predicted to be the same on the average whenever it is observed. The, now defunct, continuous creation theory proposed by Bondi, Gold and Hoyle is a good example. The WAP could have been used to make this steady-state cosmology appear extremely improbable even before it came into irredeemable conflict with direct observations. As Rees points out,[12]

> the fact that there is an epoch when [the Hubble time, t_H, which is essentially equal to the age of the Universe] is of order the age of a typical star . . . . . is not surprising in any 'big bang' cosmology. Nor is it surprising that we should ourselves be observing the universe at this particular epoch. In a steady-state cosmology, however, there would seem no *a priori* reason why the timescale for stellar evolution should not be *either* [much less than] t_H (in which case nearly all the matter would be in dead stars or 'burnt-out' galaxies) or [much greater than] t_H (in which case only a very exceptionally old galaxy would look like our own). Such considerations could have provided suggestive arguments in favour of 'big bang' cosmologies . . .

We can also give some examples of how the WAP leads to synthesizing insights that deepen our appreciation of the unity of Nature. Observed facts, often suspected at first sight to be unrelated, can be connected by examining their relation to the conditions necessary for our own existence and their explicit dependence on the constants of physics. Let us reconsider, from the Bayesian point of view, the classic example mentioned in section 1.1, relating the size of the Universe to the period of time necessary to generate observers. The requirement that enough time pass for cosmic expansion to cool off sufficiently after the Big Bang to allow the existence of carbon ensures that the observable Universe must be relatively old and so, because the boundary of the observable Universe expands at the speed of light, very large. The nuclei of carbon, nitrogen, oxygen and phosphorus of which we are made, are cooked from the light primordial nuclei of hydrogen and helium by nuclear reactions in stellar interiors. When a star nears the end of its life, it disperses these biological precursors throughout space. The time required for stars to produce carbon and other bioactive elements in this way is roughly the lifetime of a star on the 'main-sequence' of its evolution, given by

$$t_\star \sim \left(\frac{Gm_N^2}{hc}\right)^{-1} \frac{h}{m_N c} \sim 10^{10}\,\text{yrs} \tag{1.3}$$

where G is Newton's gravitation constant, c is the velocity of light, h is Planck's constant and m_N is the proton mass. Thus, in order that the Universe contain the building-blocks of life, it must be at least as old as $t_\star$ and hence, by virtue of its expansion, at least $ct_\star$ (roughly ten billion light years) in extent. No one should be surprised to find the Universe to be as large as it is. We could not exist in one that was significantly smaller. Moreover, the argument that the Universe should be teeming with civilizations on account of its vastness loses much of its persuasiveness: the Universe has to be as big as it is in order to support just one lonely outpost of life. Here, we can see the deployment of (1.2) explicitly if we let the hypothesis that *the large size of the Universe is superfluous for life on planet Earth* be α and let hypothesis β be that *life on Earth is connected with the size of the Universe.* If the evidence E is that the Universe is observed to be greater than ten billion light years in extent then, although $p_B(E/\beta) \ll 1$, the hypothesis is not necessarily then improbable because we have argued that $p_A(E/\beta) \simeq 1$.

We also observe the expansion of the Universe to be occurring at a rate which is irresolvably close to the special value which allows it the smallest deceleration compatible with indefinite future expansion. This feature of the Universe is also dependent on the epoch of observation. And again, if galaxies and clusters of galaxies grow in extent by mergers and hierarchical clustering,[2] then the characteristic scale of galaxy clustering that we infer will be determined by the cosmic epoch at which it is observed.

Ellis[39] has stressed the existence of a spatial restriction which further circumscribes the range of observed astronomical phenomena. What

amounts to a universal application of the principle of natural selection would tell us that observers may only exist in particular regions of a *spatially inhomogeneous* universe. Since realistic mathematical models of inhomogeneous universes are extremely difficult to construct, various unverifiable cosmological 'Principles' are often used by theoretical cosmologists to allow simple cosmological models to be extracted from Einstein's general theory of relativity. These Principles invariably make statements about regions of the Universe which are unobservable not only in practice but also in principle (because of the finite speed of light). Principles of this sort need to be used with care. For example, Principles of Mediocrity like the Copernican Principle or the Principle of Plenitude (see Chapter 3) would imply that if the Universe did possess a preferred place, or centre, then we should not expect to find ourselves positioned there. However, general relativity allows possible cosmological models to be constructed which not only possess a centre, but which also have conditions conducive to the existence of observers only near that centre. The WAP would offer a good explanation for our central position in such circumstances, whilst the Principles of Mediocrity would force us to conclude that we do not exist at all!

According to WAP, it is possible to contemplate the existence of many possible universes, each possessing different defining parameters and properties. Observers like ourselves obviously can exist only in that subset containing universes consistent with the evolution of carbon-based life.

This approach introduces necessarily the idea of an ensemble of possible universes and was suggested independently by the Cambridge biologist Charles Pantin in 1965. Pantin had recognized that a vague principle of amazement at the fortuitous properties of natural substances like carbon or water could not yield any *testable* predictions about the World, but the amazement might disappear if[40]

> we could know that our Universe was only one of an indefinite number with varying properties, [so] we could perhaps invoke a solution analogous to the principle of Natural Selection; that only in certain universes which happen to include ours, are the conditions suitable for the existence of life, and unless that condition is fulfilled there will be no observers to note the fact

However, as Pantin also realized, it still remains an open question as to why *any* permutation of the fundamental constants of Nature allows the existence of life, albeit a question we would not be worrying about were such a fortuitous permutation not to exist.

If one subscribes to this 'ensemble interpretation' of the WAP one must decide how large an ensemble of alternative worlds is to be admitted. Many ensembles can be imagined according to our willingness to speculate—different sets of cosmological initial data, different numerical values of fundamental constants, different space-time dimensions, different laws of physics—some of these possibilities we shall discuss in later chapters.

The theoretical investigations initiated by Carter[1] reveal that in some sense the subset of the ensemble containing worlds able to evolve observers is very 'small'. Most perturbations of the fundamental constants of Nature away from their actual numerical values lead to model worlds that are still-born, unable to generate observers and become cognizable. Usually, they allow neither nuclei, atoms nor stars to exist.

Whatever the size and variety of permutations allowed within a hypothetical ensemble of 'many worlds', one might introduce here an analogue of the Drake equation[41] often employed to guess the number of extraterrestrial civilizations in our Galaxy. Instead of expressing the probability of life existing *elsewhere* as a product of independent probabilities for the occurrence of processes like planetary formation, protocellular evolution and so forth, one could express the probability of life existing *anywhere* as a product of probabilities that encode the fact that life is only possible if parameters like the fine structure constant or the strong coupling constant lie in a particular numerical range.[42,43]

The existence of the fundamental cosmic timescale like (1.3), fixed only by invariant constants of Nature, c, h, G, and m_N, was exploited by Dicke[13] to produce a powerful WAP argument against Dirac's conclusion[30] that the Newtonian gravitation constant, G, is decreasing with time. Dirac had noticed that the dimensionless measure of the strength of gravity

$$\alpha_G \equiv \frac{Gm_N^2}{hc} \sim 10^{-39} \tag{1.4}$$

is roughly of order the inverse square root of the number of nucleons in the observable Universe, $N(t)$, at the present time $t_0 \sim 10^{10}$ yrs. At any time, t, the quantity $N(t)$ is simply

$$N(t) \equiv \frac{M_U}{m_N} = \frac{4\pi\rho_U(ct)^3}{3m_N} \sim \frac{c^3t}{Gm_N} \sim 10^{78}\left(\frac{t}{10^{10}\text{ yrs}}\right) \tag{1.5}$$

if we use the cosmological relation that the density of the Universe, ρ_U, is related to its age by $\rho_U \sim (Gt^2)^{-1}$. (The present age of roughly 10^{10} yrs is displayed in the last step.) Dirac argued that it is very unlikely that these two quantities should possess simply related dimensionless magnitudes which are both so vastly different from unity and yet be independent. Rather, there must exist an approximate *equality* between them of the form

$$N(t) \sim \alpha_G^{-2} \tag{1.6}$$

However, whereas α_G is a *time-independent* combination of constants, $N(t)$ increases linearly with the time of observation, t, which for us is the present age of the Universe. The relation (1.6) can only hold for all times if one component of α_G is time-varying and so Dirac suggested that we must have $G \propto t^{-1}$ so that $N(t) \propto \alpha_G^{-2} \propto t^2$. The quantities $N(t)$ and α_G^{-2} are now observed to be of the same magnitude because (as a result of some unfound law of Nature) they are actually *equal*, and furthermore, they are of such an enormous magnitude because they both increase linearly in time and the Universe is very old—although this 'oldness' can presumably only be explained by the WAP even in this scheme of 'varying' constants for the reasons discussed above in connection with the size of the Universe.

However, the WAP shows Dirac's radical conclusion of a time-varying Newtonian gravitation constant to be quite unnecessary. The coincidence that today we observe $N \sim \alpha_G^{-2}$ is necessary for our existence. Since we would not expect to observe the Universe either before stars form or after they have burnt out, human astronomers will most probably observe the Universe close to the epoch $t_\star$ given by (1.3). Hence, we will observe the time-dependent quantity $N(t)$ to take on a value of order $N(t_\star)$ and, by (1.3) and (1.4), this value is necessarily just

$$N(t_\star) \sim \frac{t_\star}{Gm_N} \sim \alpha_G^{-2} \tag{1.7}$$

where the second relation is a *consequence* of the value of $t_\star$ in (1.3). If we let δ be Dirac's hypothesis of time-varying G, while γ is the hypothesis that G is constant while the 'evidence', E, is the coincidence (1.6); then, although the *a priori* probability that we live at the time when the numbers $N(t)$ and α_G^{-2} are equal is very low, ($p_B(E/\gamma) \ll 1$), this does not render hypothesis γ (the constancy of G) implausible because there is an anthropic selection effect which ensures $p_A(E/\gamma) \simeq 1$. This selection effect is the one pointed out by Dicke. We should notice that this argument alone explains why we must observe $N(t)$ and α_G^{-2} to be of equal magnitude, but not why that magnitude has the extraordinarily large value $\sim 10^{79}$. (We shall have a lot more to say about this problem in Chapters 4, 5 and 6).

As mentioned in section 1.1, Carter[1] introduced the more speculative Strong Anthropic Principle (SAP) to provide a 'reason' for our observation of large dimensionless ratios like 10^{79}; we state his SAP as follows:

Strong Anthropic Principle (SAP): The Universe must have those properties which allow life to develop within it at some stage in its history.

An implication of the SAP is that the constants and laws of Nature must be such that life can exist. This speculative statement leads to a number of quite distinct interpretations of a radical nature: firstly, the most obvious is to continue in the tradition of the classical Design Arguments and claim that:

(A) *There exists one possible Universe 'designed' with the goal of generating and sustaining 'observers'.*

This view would have been supported by the natural theologians of past centuries, whose views we shall examine in Chapter 2. More recently it has been taken seriously by scientists who include the Harvard chemist Lawrence Henderson[44] and the British astrophysicist Fred Hoyle, so impressed were they by the string of 'coincidences' that exist between particular numerical values of dimensionless constants of Nature without which life of any sort would be excluded. Hoyle[45] points out how natural it might be to draw a teleological conclusion from the fortuitous positioning of nuclear resonance levels in carbon and oxygen:

> I do not believe that any scientist who examined the evidence would fail to draw the inference that the laws of nuclear physics have been deliberately designed with regard to the consequences they produce inside the stars. If this is so, then my apparently random quirks have become part of a deep-laid scheme. If not then we are back again at a monstrous sequence of accidents.

The interpretation (A) above does not appear to be open either to proof or to disproof and is religious in nature. Indeed it is a view either implicit or explicit in most theologies.

This is all we need say about the 'teleological' version of the SAP at this stage. However, the inclusion of quantum physics into the SAP produces quite different interpretations. Wheeler[6] has coined the title '*Participatory Anthropic Principle*' (PAP) for a second possible interpretation of the SAP:

(B) *Observers are necessary to bring the Universe into being.*

This statement is somewhat reminiscent of the outlook of Bishop Berkeley and we shall see that it has physical content when considered in the light of attempts to arrive at a satisfactory interpretation of quantum mechanics.[46] It is closely related to another possibility:

(C) *An ensemble of other different universes is necessary for the existence of our Universe.*

This statement receives support from the 'Many-Worlds' interpretation of quantum mechanics and a sum-over-histories approach to quantum gravitation because they must unavoidably recognize the existence of a whole class of *real* 'other worlds' from which ours is selected by an optimizing principle.[47] We shall express this version of the SAP

mathematically in Chapter 7, and we shall see that this version of the SAP has consequences which are potentially testable.

Suppose that for some unknown reason the SAP is true and that intelligent life must come into existence at some stage in the Universe's history. But if it dies out at our stage of development, long before it has had any measurable non-quantum influence on the Universe in the large, it is hard to see why it *must* have come into existence in the first place. This motivates the following generalization of the SAP:

Final Anthropic Principle (*FAP*): *Intelligent information-processing must come into existence in the Universe, and, once it comes into existence, it will never die out.*

We shall examine the consequences of the FAP in our final chapter by using the ideas of information theory and computer science. The FAP will be made precise in this chapter. As we shall see, FAP will turn out to require the Universe and elementary particle states to possess a number of definite properties. These properties provide observational tests for this statement of the FAP.

Although the FAP is a statement of physics and hence *ipso facto*[53] has no ethical or moral content, it nevertheless is closely connected with moral values, for the validity of the FAP is the physical precondition for moral values to arise and to continue to exist in the Universe: no moral values of any sort can exist in a lifeless cosmology. Furthermore, the FAP seems to imply a melioristic cosmos.

We should warn the reader once again that both the FAP and the SAP are quite speculative; unquestionably, neither should be regarded as well-established principles of physics. In contrast, the WAP is just a restatement, albeit a subtle restatement, of one of the most important and well-established principles of science: that it is essential to take into account the limitations of one's measuring apparatus when interpreting one's observations.

References

1. B. Carter, in *Confrontation of cosmological theories with observation*, ed. M. S. Longair (Reidel, Dordrecht, 1974), p. 291.
2. See for example P. J. Peebles, *The large scale structure of the universe*, (Princeton University Press, Princeton, 1980).
3. A. Sandage and E. Hardy, *Astrophys. J.* **183,** 743 (1973). S. Weinberg, *Gravitation and cosmology* (Wiley, NY, 1972).
4. We examine some of these claims in Chapters 2 and 3.
5. E. Hubble, *Proc. natn. Acad. Sci., U.S.A.* **15,** 169 (1929).
6. J. A. Wheeler, in *Foundational problems in the special sciences*, ed. R. E.

Butts and J. Hintikka (Reidel, Dordrecht, 1977), p. 3; in *The nature of scientific discovery*, ed. O. Gingerich (Smithsonian Press, Washington, 1975), pp. 261–96 and pp. 575–87.

7. D. Clayton, *Principles of stellar evolution and nucleosynthesis* (McGraw-Hill, NY, 1968). R. J. Tayler, *The stars: their evolution and structure* (Wykeham, London, 1970).
8. For reviews of a number of examples see, in particular, B. J. Carr and M. J. Rees, *Nature* **278,** 605 (1979), V. Weisskopf, *Science* **187,** 605 (1975); we shall investigate this in detail in Chapters 5 and 6.
9. For an interesting overview of constants see *The constants of nature*, ed. W. H. McCrea and M. J. Rees (Royal Society of London, London, 1983). This book was originally published as the contents of *Phil. Trans. R. Soc.* Vol. A **310** in 1983 See also J. M. Levy-Leblond, *Riv. nuovo Cim.* **7,** 187 (1977).
10. P. Candelas and S. Weinberg, *Nucl. Phys.* B **237,** 397 (1984).
11. P. C. W. Davies, *J. Phys.* A **5,** 1296 (1972).
12. M. J. Rees, *Comm. Astrophys. Space Phys.* **4,** 182 (1972).
13. R. H. Dicke, *Rev. Mod. Phys.* **29,** 355 and 363 (1977); *Nature* **192,** 440 (1961).
14. P. L. M. Maupertuis, *Essai de cosmologie* (1751), in *Oeuvres*, Vol. 4, p. 3 (Lyon, 1768). We discuss these developments in detail in sections 2.5 and 3.4.
15. J. D. Barrow, 'Cosmology, the existence of observers and ensembles of possible universes', in *Les voies de la connaissance* (Tsukuba Conference Proceedings, Radio-France Culture, Paris 1985).
16. J. D. Barrow, *Quart. J. R. astron. Soc.* **23,** 146 (1983).
17. C. B. Collins and S. W. Hawking, *Astrophys. J.* **180,** 317 (1973); J. D. Barrow, *Quart. J. R. astrom. Soc.* **23,** 344 (1982). For a popular discussion see Chapter 5 of J. D. Barrow and J. Silk, *The left hand of creation* (Basic Books, NY, 1983 and Heinemann, London, 1984).
18. This is an old argument applied to cosmology by G. F. R. Ellis and G. B. Brundrit, *Q. J. R. astron. Soc.* **20,** 37 (1979). See F. J. Tipler, *Quart. J. R. astron. Soc.* **22,** 133 (1981) for a discussion of the history of this argument.
19. Notice the infinity alone is not a sufficient condition for this to occur; it must be an exhaustively random infinity in order to include all possibilities.
20. If the visible part of the Universe is accurately described by Friedman's equation without cosmological constant (as seems to be the case, see ref. 3) then a density exceeding about $2 \,.\, 10^{-29}$ gm cm^{-3} is required.
21. A. Guth, *Phys. Rev.* D **23,** 347 (1981); K. Sato, *Mon. Not. R. astron. Soc.* **195,** 467 (1981); A. Linde, *Phys. Lett.* B **108,** 389 (1982). For an overview see G. Gibbons, S. W. Hawking, and S. T. C. Siklos, *The very early universe* (Cambridge University Press, Cambridge, 1983).
22. A. Linde, *Nuovo Cim. Lett.* **39,** 401 (1984).
23. J. D. Barrow, *Phil. Trans. R. Soc.* A **310,** 337 (1983); E. Witten, *Nucl. Phys.* B **186,** 412 (1981).
24. R. D. Reasonberg, *Phil. Trans. R. Soc.* A **310,** 227 (1983).
25. H. B. Nielsen, *Phil. Trans. R. Soc.* A **310,** 261 (1983); J. Iliopoulos, D. V. Nanopoulos, and T. N. Tamvaros, *Phys. Lett.* B **94,** 141 (1980); J. D. Barrow and A. C. Ottewill, *J. Phys.* A **16,** 2757 (1983).
26. J. A. Wheeler and W. H. Zurek, *Quantum theory and measurement*, (Prince-

ton University Press, Princeton, 1982); H. Everett, *Rev. Mod. Phys.* **29,** 454 (1957); F. J. Tipler, 'Interpreting the wave function of the universe', *Phys. Rep.* (In press.) B. Espagnet, *Scient. Am. Nov* (1979), p. 128.

27. V. Trimble, *Am. Scient.* **65,** 76 (1977); F. Dyson, *Scient. Am.* **224,** No. 3, pp. 50–9 (Sept 1971); J. Leslie, *Am. Phil. Quart.* **19,** 141 (1982), *Am. Phil. Quart.* **7,** 286 (1970), *Mind* **92,** 573 (1983); in *Scientific explanation and understanding: essays on reasoning and rationality in science*, ed. N. Rescher (University Press of America, Lanham, 1983), pp. 53–83; in *Evolution and creation*, ed. E. McMullin (University of Notre Dame Press, Notre Dame 1984); P. J. Hall, *Quart. J. R. astron. Soc.* **24,** 443 (1983); J. Demaret and C. Barbier, *Revue des Questions Scientifique* **152,** 181, 461 (1981); E. J. Squires, *Eur. J. Phys.* **2,** 55 (1981); P. C. W. Davies, *Accidental universe* (Cambridge University Press, Cambridge, 1982); R. Breuer, *Das anthropische Prinzip* (Meyster, München, 1981); J. D. Barrow and J. Silk, *Scient. Am.*, April (1980), p. 98; A. Finkbeiner, *Sky & Telescope*, Aug. (1984), p. 107; J. D. Barrow and F. J. Tipler, *L'homme et le cosmos* (Imago-Radio France, Paris, 1984); J. Eccles, *The human mystery* (Springer, NY, 1979); B. Lovell, *In the centre of immensities*, (Harper & Row, NY, 1983); J. A. Wheeler, *Am. Scient.* **62,** 683 (1974); G. Gale, *Scient. Am.* **245** (No. 6, Dec.), 154 (1981); M. T. Simmons, Mosaic (March-April 1982) p. 16; G. Wald, *Origins of Life* **5,** 7 (1974); S. W. Hawking, *CERN Courier* **21** (1), 3 (1981); G. F. R. Ellis, *S. Afr. J. Sci.*, **75,** 529 (1979); S. J. Gould, *Natural History* **92,** 34 (1983); J. Maddox, *Nature* **307,** 409 (1984); P. C. W. Davies, *Prog. Part. Nucl. Phys.* **10,** 1 (1983); F. J. Dyson, *Disturbing the universe* (Harper & Row, NY, 1979).
28. For a representative general bibliography of Design Arguments see: H. Baker, *The image of man* (Harper, NY, 1947); P. Bertocci, *An introduction to the philosophy of religion*, (Prentice Hall, NY, 1951); *The cosmological arguments*, ed. D. R. Burnill (Doubleday, NY, 1967); E. A. Burtt, *The metaphysical foundations of modern physical science* (Harcourt Brace, NY, 1927); C. Hartshorne, *A natural theology for our time* (Open Court, NY, 1967); L. E. Hicks, *A critique of Design Arguments* (Scribners, NY, 1883); R. H. Hurlbutt III, *Hume, Newton and the Design Argument*, (University Nebraska Press, Lincoln, Nebraska, 1965); P. Janet, *Final causes* (Clark, Edinburgh, 1878); D. L. LeMahieu, *The mind of William Paley* (University Nebraska Press, Lincoln, Nebraska, 1976); A. O. Lovejoy, *The Great Chain of Being: a study in the history of an idea*, (Harvard University Press, Cambridge, 1936); J. D. McFarland, *Kant's concept of teleology*, (University Edinburgh Press, Edinburgh, 1970); T. McPherson, *The argument from design* (Macmillan, Edinburgh, 1972); L. Stephen, *English thought in the eighteenth century*, Vol. 1 (Harcourt Brace, NY, 1962); R. G. Swinburne, *Philosophy* **43,** 164 (1968); F. R. Tennant, *Philosophical theology*, 2 vols (Cambridge University Press, Cambridge, 1930); A. Woodfield, *Teleology* (Cambridge University Press, Cambridge, 1976); L. Wright, *Teleological explanations* (University of California Press, Berkeley, 1976).
29. J. D. Barrow, *Quart. J. R. astron. Soc.* **22,** 388 (1981).
30. P. A. M. Dirac, *Nature* **139,** 323 (1937).
31. G. Whitrow, *Br. J. Phil. Sci.* **6,** 13 (1955).
32. B. J. Carr and M. J. Rees, *Nature* **278,** 605 (1979).
33. M. H. Hart, *Icarus* **33,** 23 (1978).
34. J. Hartle and S. W. Hawking, *Phys. Rev.* D **28,** 2960 (1983); F. J. Tipler, preprint (1984).

35. B. Carter, *Phil. Trans. R. Soc.* A **310,** 347 (1983).
36. Whitrow, cited in E. Mascall, *Christian theology and natural science: 1956 Bampton lectures* (Longmans Green, London, 1956).
37. G. Idlis, *Izv. Astrophys. Inst. Kazakh. SSR* **7,** 39 (1958), in Russian.
38. See, for example, P. L. Meyer, *Introductory probability and statistical applications* (Addison-Wesley, NY, 1971).
39. G. F. R. Ellis, *Gen. Rel. Gravn.* **11,** 281 (1979); G. F. R. Ellis, R. Maartens, and S. D. Nel, *Mon. Not. R. Soc.* **184,** 439 (1978).
40. C. F. A. Pantin, in *Biology and personality*, ed. I. T. Ramsey (Blackwell, Oxford, 1965), pp. 103–4.
41. I. S. Shklovskii and C. Sagan, *Intelligent life in the universe* (Dell, NY, 1966).
42. J. D. Barrow, in ref. 23.
43. T. L. Wilson, *Quart. J. R. astron. Soc.* **25,** 435 (1984).
44. L. J. Henderson, *The fitness of the environment* (Smith, Gloucester, Mass., 1913; reprinted Harvard University Press, Cambridge, Mass., 1970) and *The order of Nature* (Harvard University Press, Cambridge, Mass., 1917).
45. F. Hoyle, in *Religion and the scientists* (SCM, London, 1959).
46. M. Jammer, *The philosophy of quantum mechanics* (Wiley, NY, 1974).
47. R. P. Feynman and A. R. Hibbs, *Quantum mechanics and path integrals* (McGraw-Hill, New York, 1965); L. S. Schulman, *Techniques and applications of path integration* (Wiley, NY, 1981).
48. C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), Chapter 44.
49. For a lucid recent expression of this view from a professional historian see 'Whigs and professionals', by C. Russell in *Nature* **308,** 777 (1984).
50. H. Butterfield, *The Whig interpretation of history* (G. Bell, London, 1951), p. 92. This book was first published in 1931. Butterfield later toned down his opposition to Whiggery. See, for example, his book *The Englishman and his history* (Cambridge University Press, Cambridge, 1944).
51. M. White, unpublished lecture at Harvard University (1957); quoted by W. W. Bartley III, in *The retreat to commitment* (Knopf, NY, 1962), pp. 98–100.
52. G. Holton, *Thematic origins of scientific though* (Harvard University Press, NY, 1973).
53. Physical assumptions and moral assumptions belong to different logical categories. Physical assumptions, like all scientific statements, are statements of matters of fact: syntactically, they are declarative sentences. Moral assumptions, on the other hand, are statements concerning moral obligation: syntactically, they are imperative sentences, which contain the word 'ought' or its equivalent. For further discussion of this point see any modern textbook on moral philosophy, for instance H. Reichenbach, *The rise of scientific philosophy* (University of California Press, Berkeley, 1968).
54. For further discussion of self-reference arguments in mathematics, see D. Hofstadter, *Gödel, Escher, Bach: an eternal golden braid* (Basic Books, NY, 1979).
55. For a more recent defence of the idea that the past has to be discussed in terms the present can understand, see D. Hull, *History & Theory* **18,** 1 (1979). We are grateful to Professor S. G. Brush for this reference.