ELSEVIER

Jeremy Ramsden

APPLIED NANOTECHNOLOGY

Micro & Nano Technologies Series

William Andrew is an imprint of Elsevier
Linacre House, Jordan Hill, Oxford OX2 8DP, UK
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA

First edition 2009

Library of Congress Cataloging in Publication Data
A catalog record for this book is available from the Library of Congress

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-8155-2023-8

Typeset by: diacriTech, India

Printed and bound in United States of America
09 10 11 12   11 10 9 8 7 6 5 4 3 2 1

No ládd, e nép, mely közt már senki nem hisz,
Ami csodás, hogyan kapkodja mégis.

IMRE MADÁCH

Series Editor's Preface

The possibility of modifying materials using electrical discharges has fascinated mankind ever since he observed the results of lightning striking objects in nature. We do not, of course, know when the first observation took place, but we may be reasonably sure that it was a sufficiently long time ago that many millennia had to pass before electricity was "tamed," and subsequently put to work modifying materials in a systematic, "scientific" way—as exemplified by Humphry Davy's electrolysing common salt to produce metallic sodium at the Royal Institution in London.

But these are essentially faradaic processes (named after Davy's erstwhile assistant Michael Faraday), and such processes are also used extensively today for (micro)machining, as exemplified by electrochemical machining (ECM). They are relatively well known, and are applicable to conducting workpieces. Far less well known is the technology of what is now called spark-assisted chemical engraving (SACE), in which the workpiece is merely placed in the close vicinity of the pointed working electrode, and is eroded by sparks jumping across the gas bubbles that develop around the electrode to reach the electrolyte in which everything is immersed, the circuit being completed by the presence of a large counter-electrode.

This technology can therefore be equally well used for workpieces made from non-conducting materials such as glass, traditionally difficult to machine, especially at the microlevel precision needed for such applications as microfluidic mixers and reactors. The development of attractive machining technologies such as SACE is in itself likely to play a decisive part in the growth of microfluidics-based methods in chemical proccssing and medical diagnostics, to name just two important areas of application.

Since, as the author very correctly points out, knowledge about non-faradaic ECM methods is presently remarkably scanty within the microsystems community, this book is conceived as a comprehensive treatise, covering the entire field, starting with a lucid explanation of the physicochemical fundamentals, continuing with a thorough discussion of the practical questions likely to be asked, and ending with an authoritative exposition of the means to their resolution.

I therefore anticipate that this book will significantly contribute to enabling the rapid growth of micromachining of non-conducting materials, for which there is tremendous hitherto unexploited potential.

Jeremy J. Ramsden
Cranfield University, United Kingdom
December 2008

Preface

This is as much a book about ideas as about facts. It begins (Chapter 1) by explaining—yet again!—what nanotechnology is. For those who feel that this is needless repetition of a well-worn theme, may I at least enter a plea that as more and more people and organizations (latterly the International Standards Organization) engage themselves with the question, the definition is steadily becoming better refined and less ambiguous, and account needs to be taken of these developments.

The focus of this book is nanotechnology in commerce, hence in the first part dealing with basics, Chapter 2 delves into the fascinating relationship between wealth, technology and science. Whereas for millennia we have been accustomed to technology emerging from wealth, and science emerging from technology, nanotechnology exemplifies a new paradigm in which science is in the van of wealth generation.

The emergence of nanotechnology products from underlying science and technology is an instantiation of the process called innovation. The process is important for any high technology; given that nanotechnology not only exemplifies but really epitomizes high technology, the relation between nanotechnology and innovation is of central importance. Its consideration (Chapter 3) fuses technology, economics and social aspects.

Chapter 4 addresses the question "Why might one wish to introduce nanotechnology?" Nanotechnology products may be discontinuous with respect to existing ones in the sense that they are really new, instantiating things that simply did not exist, or were only dreamt about, before the advent of nanotechnology. They may also be a result of *nanification*, decreasing the size of an existing device, or a component of the device, down to the nanoscale. Not every manufactured artefact can be advantageously nanified; this chapter tackles the crucial aspects of when it is technically, and when it is commercially advantageous.

These first four chapters cover Part 1 of this book. Part 2 looks at actual nanotechnology products—in effect, defining nanotechnology ostensively. It is divided into four chapters, the first one (Chapter 5) giving an overview of the entire market. Chapters 7 and 8 deal with, respectively, information technology and healthcare, which are the biggest sectors with strong nanotechnology associations; all other applications, including coatings of various kinds, composite materials, energy, agriculture, and so forth, are included in Chapter 6.

Part 3 deals with more specifically commercial, especially financial aspects, and comprises three chapters. The first one (Chapter 9) is devoted to business models for nanotechnology enterprises. Particular emphasis is placed on the spin-off company, and the role of government in promoting nanotechnology is discussed in some detail. Chapter 10 deals with how demand for nanoproducts can be assessed. The third chapter (11) deals with special problems of designing nano products.

The final part of the book takes a look toward the future. Chapter 12 essentially deals with productive nanosystems; that is, what may happen when molecular manufacturing plays a significant role in industrial production. The implications of this future state are so profoundly different from what we have been used to during the past few centuries that it is worth discussing, even though its advent must be considered a possibility rather than a certainty. There is also a discussion about the likelihood of bottom-up nanofacture (self-assembly) becoming established as an industrial method. Chapter 13 asks how nanotechnology can contribute to the grand challenges currently facing humanity. It is perhaps unfortunate that insofar as failure to solve these challenges looks as though it will jeopardize the very survival of humanity, they must be considered as threats rather than opportunities, with the corollary that if nanotechnology cannot contribute to solving these problems, then humanity cannot afford the luxury of diverting resources into it. The final Chapter 14 is devoted to ethical issues. Whether or not one accepts the existence of a special branch of ethics that may be called "nanoethics", undoubtedly nanotechnology raises a host of issues affecting the lives of every one of us, both individually and collectively, and which cannot be ignored by even the most dispassionate businessperson.

In summary, this book tries to take as complete an overview as possible, not only of the technology itself, but also of its commercial and social context. This view is commensurate with the all-pervasiveness of nanotechnology, and hopefully brings the reader some way toward answering the three questions: What can I know about nanotechnology? What should I do with nanotechnology (how should I deal with it)? What can I hope for from nanotechnology?

Nanotechnology has been and still is associated with a fair share of hyperbole, which sometimes attracts criticism, especially from sober open-minded scientists. But is this hyperbole any different from the exuberance with which Isambard Brunel presented his new Great Western Railway as the first link in a route from London to New York, or Sir Edward Watkin his new Great Central Railway as a route from Manchester to Paris? Moreover, apart from the technology, the nano viewpoint is also an advance in the way of looking at the world which is a worthy successor to the previous advances of knowledge that have taken place over the past millennium. And especially now, when humanity is facing exceptional threats, an exceptional viewpoint coupled with an exceptional technology might offer the only practical hope for survival.

I should like to especially record my thanks to the members of my research group at Cranfield University, with whom our weekly discussions about these issues helped to hone my ideas, my colleagues at Cranfield for many stimulating exchanges about nanotechnology, and to Dr Graham Holt for his invaluable help in hunting out commercial data. It is also a pleasure to thank Enza Giaracuni for having prepared the drawings.

Jeremy J. Ramsden
Cranfield University
January 2009

CHAPTER 1

What is Nanotechnology?

CHAPTER CONTENTS

In the heady days of any new, emerging technology, definitions tend to abound and are first documented in reports and journal publications, then slowly get into books and are finally taken up by dictionaries, which do not prescribe, however, but merely record usage. Ultimately the technology will attract the attention of the International Standards Organization (ISO), which may in due course issue a technical specification (TS) prescribing in an unambiguous manner the terminology of the field, which is clearly an essential prerequisite for the formulation of manufacturing standards.

In this regard, nanotechnology is no different, except that nanotechnology seems to be arriving rather faster than the technologies we might be familiar with from the past, such as steam engines and digital computers. As a reflection of the rapidity of this arrival, the ISO has already set up a Technical

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

Committee (TC 229) devoted to nanotechnologies. Thus, unprecedentedly in the history of the ISO, we shall have technical specifications in advance of a significant industrial sector.

The work of TC 229 is not yet complete, however, hence we shall have to make our own attempt to find a consensus definition. As a start, let us look at the roots of the technology. They are widely attributed to Richard Feynman, who in a now famous lecture at Caltech in 1959 advocated manufacturing things at the smallest possible scale, namely atom by atom—hence the prefix "nano", atoms typically being a few tenths of a nanometre (10^{-9} m) in size. He was clearly envisaging a manufacturing technology, but from the lecture we also have glimpses of a novel viewpoint, namely that of looking at things at the atomic scale—not only artefacts fashioned by human ingenuity, but also the minute molecular machines grown inside living cells.

1.1 NANOTECHNOLOGY AS PROCESS

We see nanotechnology as looking at things—measuring, describing, characterizing and quantifying them, and ultimately reaching a deeper assessment of their place in the universe. It is also making things. Manufacturing was evidently very much in the mind of the actual inventor of the term "nanotechnology", Norio Taniguchi from the University of Tokyo, who considered it as the inevitable consequence of steadily improving engineering precision (Figure 1.1).[1] Clearly, the surface finish of a workpiece achieved by grinding cannot be less rough than atomic roughness, hence nanotechnology must be the endpoint of ultraprecision engineering.

At the same time, improvements in metrology had reached the point where individual atoms at the surface of a piece of material could be imaged, hence visualized on a screen. The possibility was of course already inherent in electron microscopy, which was invented in the 1930s, but numerous incremental technical improvements were needed before atomic resolution became attainable. Another development was the invention of the "Topografiner" by scientists at the US National Standards Institute.[2] This instrument produced a map of topography at the nanoscale by raster scanning a needle over the surface of the sample. A few years later, it was developed

[1] N. Taniguchi, On the basic concept of nano-technology. *Proc. Intl Conf. Prod. Engng Tokyo, Part II* (Jap. Soc. Precision Engng).

[2] R. Young et al., The Topografiner: an instrument for measuring surface microtopography. *Rev. Sci. Instrum.* 43 (1972) 999–1011.

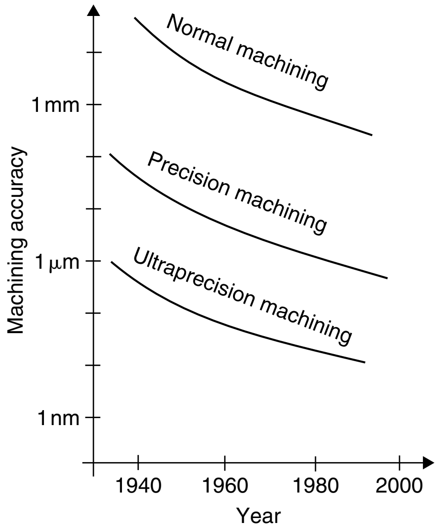


FIGURE 1.1 *The evolution of machining accuracy (after Norio Taniguchi).*

into the scanning tunneling microscope (STM), and in turn the atomic force microscope (AFM) that is now seen as the epitome of nanometrology (collectively, these instruments are known as scanning probe microscopes, SPMs). Hence a little more than 10 years after Feynman's lecture, advances in instrumentation already allowed one to view the hitherto invisible world of the nanoscale in a very graphic fashion. There is a strong appeal in having a small, desktop instrument (such as the AFM) able to probe matter at the atomic scale, which contrasts strongly with the bulk of traditional high-resolution instruments such as the electron microscope, which needs at least a room and perhaps a whole building to house it and its attendant services. Every nanotechnologist should have an SPM in his or her study!

In parallel, people were also thinking about how atom-by-atom assembly might be possible. Erstwhile Caltech colleagues recall Richard Feynman's dismay when William McLellan constructed a minute electric motor by hand-assembling the parts in the manner of a watchmaker, thereby winning the prize Feynman had offered for the first person to create an electrical motor smaller than 1/64th of an inch. Although this is still how nanoscale artefacts are made (but perhaps for not much longer), Feynman's concept was of machines making progressively smaller machines ultimately small enough to manipulate atoms and assemble things at that scale. The most indefatigable champion of that concept was Eric Drexler, who developed the concept of the assembler, a tiny machine programed to build objects atom by atom. It was an obvious corollary of the minute size of an assembler that in order to make anything of a size useful for humans, or in useful numbers, there would have to be a great many assemblers working in parallel. Hence, the first task of the assembler would be to build copies of itself, after which they would be set to perform more general assembly tasks.

This program received a significant boost when it was realized that the scanning probe microscope (SPM) could be used not only to determine nanoscale topography, but also as an assembler. IBM researchers iconically demonstrated this application of the SPM by creating the logo of the company in xenon atoms on a nickel surface at 4 K: The tip of the SPM was used to laboriously push 18 individual atoms into location.[3] Given that the assembly of the atoms in two dimensions took almost 24 hours of laborious manual manipulation, few people associated the feat with steps on the road to molecular manufacturing. Indeed, since then further progress in realizing an assembler has been painstakingly slow;[4] the next milestone was Oyabu's demonstration of picking up (abstracting) a silicon atom from a silicon surface and placing it at somewhere else on the same surface, and then carrying out the reverse operation.[5] Following on in the spirit of Taniguchi, semiconductor processing—the sequences of material deposition and etching through special masks used to create electronic components[6]—integrated circuits—has now achieved feature sizes below the threshold of 100 nm that is usually considered to constitute the upper boundary of the nano realm (the lower boundary being about 0.1 nm, the size of atoms).

Frustration at being unable to apply "top-down" processing methods to achieve feature sizes in the nanometer, or even the tens of nanometers range stimulated the development of "bottom-up" or self-assembly methods. These were inspired by the ability of randomly ordered structures, or mixtures of components, to form definite structures in biology. Well-known examples are proteins (merely upon cooling, a random polypeptide coil of a certain sequence of amino acids will adopt a definite structure), the ribosome, and

---

[3]E.K. Schweizer and D.M. Eigler, Positioning single atoms with a scanning tunnelling microscope. *Nature* (Lond.) 344 (1990) 524–526.

[4]Apart from intensive activity in numerically simulating the steps of molecular manufacturing—e.g., B. Temelso et al., Ab initio thermochemistry of the hydrogenation of hydrocarbon radicals using silicon-, germanium-, tin-, and lead-substituted methane and isobutene. *J. Phys. Chem. A* 111 (2007) 8677–8688.

[5]N. Oyabu, Ó. Custance, I. Yi, Y. Sugawara and S. Morita, Mechanical vertical manipulation of selected single atoms by soft nanoindentation using near contact atomic force microscopy. *Phys. Rev. Lett.* 90 (2003) 176102.

[6]A.G. Mamalis, A. Markopoulos and D.E. Manolakos, Micro and nanoprocessing techniques and applications. *Nanotechnol. Perceptions* 1 (2005) 63–73.

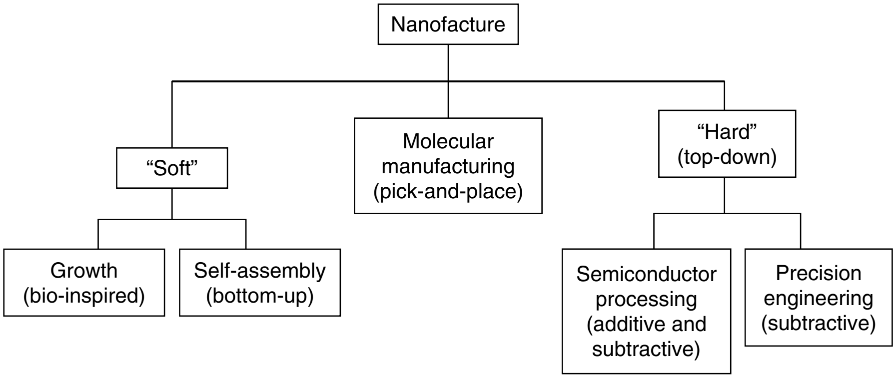


FIGURE 1.2 *Different modes of nanomanufacture (nanofacture).*

bacteriophage viruses—a stirred mixture of the constituent components will spontaneously assemble into a functional final structure.

At present, a plethora of ingeniously synthesized organic and organometallic compounds capable of spontaneously connecting themselves to form definite structures are available. Very often these follow the hierarchical sequence delineated by A.I. Kitaigorodskii as a guide to the crystallization of organic molecules (the Kitaigorodskii Aufbau Principle, KAP)—the individual molecules first form rods, the rods bundle to form plates, and the plates stack to form a three-dimensional space-filling object. Exemplars in nature include glucose polymerizing to form cellulose molecules, which are bundled to form fibrils, which in turn are stacked and glued with lignin to create wood. Incidentally, this field already had a life of its own, as supramolecular chemistry, before nanotechnology focused interest on self-assembly processes.

Molecular manufacturing, the sequences of pick and place operations carried out by assemblers, fits in somewhere between these two extremes. Insofar as a minute object is assembled from individual atoms, it might be called "bottom-up". On the other hand, insofar as atoms are selected and positioned by a much larger tool, it could well be called "top-down". Hence it is sometimes called "bottom-to-bottom". Figure 1.2 summarizes the different approaches to nanofacture (nanomanufacture).

1.2 NANOTECHNOLOGY AS MATERIALS

The above illustrates an early preoccupation with nanotechnology as process—a way of making things. Before the semiconductor processing industry reduced the feature sizes of integrated circuit components to less than 100

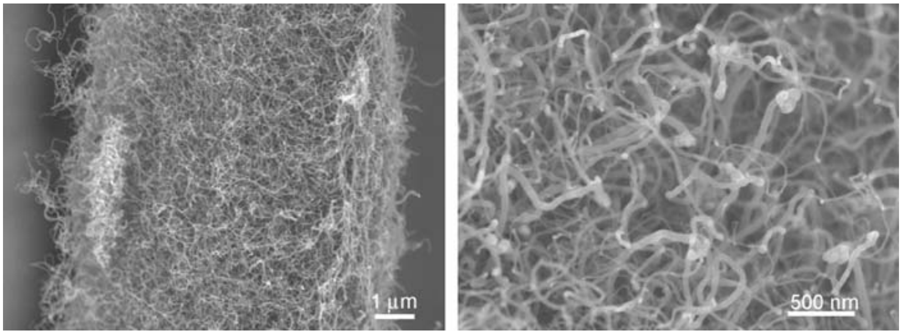


FIGURE 1.3
Scanning electron micrographs of carbon nanotubes grown on the surface of a carbon fiber using thermal chemical vapor deposition. The right-hand image is an enlargement of the surface of the fiber, showing the nanotubes in more detail. Reprinted from B.O. Boscovic, Carbon nanotubes and nanofibers. Nanotechnol. Perceptions *3 (2007) 141–158, with permission from Collegium Basilea.*

nanometers,[7] this however, was no real industrial example of nanotechnology at work. On the other hand, while process—top-down and bottom-up, and we include metrology here—is clearly one way of thinking about nanotechnology, there is already a sizable industry involved in making very fine particles, which, because their size is less than 100 nm, might be called nanoparticles. Generalizing, a nano-object is something with at least one spatial (Euclidean) dimension less than 100 nm; from this definition are derived those for nanoplates (one dimension less than 100 nm), nanofibers (two dimensions less than 100 nm), and nanoparticles (all three dimensions less than 100 nm); nanofibers are in turn subdivided into nanotubes (hollow fibers), nanorods (rigid fibers), and nanowires (conducting fibers).

Although nanoparticles of many different kinds of materials have been made for hundreds of years, one nanomaterial stands out as being rightfully so named, because it was discovered and nanoscopically characterized in the nanotechnology era: graphene and its compactified forms, namely carbon nanotubes (Figure 1.3) and fullerenes (nanoparticles).

A very important application of nanofibers and nanoparticles is in nanocomposites, as described in more detail in Chapter 6.

1.3 NANOTECHNOLOGY AS MATERIALS, DEVICES AND SYSTEMS

One problem with associating nanotechnology exclusively with materials is that nanoparticles were deliberately made for various aesthetic, technological and medical applications at least 500 years ago, and one would therefore

[7]This is a provisional upper limit of the nanoscale. More careful considerations suggest that the nanoscale is, in fact, a property-dependent. See J.J. Ramsden and J. Freeman, The nanoscale, *Nanotechnol. Perceptions* 5 (2009) 3–26.

be compelled to say that nanotechnology began then. To avoid that problem, materials are generally grouped with other entities along an axis of increasing complexity, encompassing devices and systems. A nanodevice, or nanomachine, is defined as a nanoscale automaton (i.e., an information processor), or at least as one containing nanosized components. Responsive or "smart" materials could of course also be classified as devices. A device might well be a system (of components) in a formal sense; it is not generally clear what use is intended by specifying "nanosystem", as distinct from a device. At any rate, materials may be considered as the most basic category, since devices are obviously made from materials, even though the functional equivalent of a particular device could be realized in different ways, using different materials.

1.4 DIRECT, INDIRECT AND CONCEPTUAL NANOTECHNOLOGY

Another axis for displaying nanotechnology, which might be considered as orthogonal to the materials, devices and systems axis, considers direct, indirect and conceptual aspects. Direct nanotechnology refers to nanosized objects used directly in an application—a responsive nanoparticle used to deliver drugs to an internal target in the human body is an example. Indirect nanotechnology refers to a (probably miniature) device that contains a nanodevice, possibly along with other micro or macro components and systems. An example is a cellphone. The internal nanodevice is the "chip"—the integrated electronic information processor circuits with feature sizes less than 100 nm. All the uses to which the cellphone might be put would then rank as indirect nanotechnology. Given the ubiquity of contemporary society's dependency on information processing, nanotechnology is truly pervasive from this viewpoint alone. It is, of course, the very great processing power, enabled by the vast number of components on a small chip, and the relatively low cost (arising from the same reason), both of which increasingly rely on nanotechnology for their realization, that makes the "micro" processor ubiquitous.

Conceptual nanotechnology refers to considering systems or, as we can say even more generally, "phenomena" from the nano viewpoint—trying to understand the mechanism of a process at the atomic scale. Hence, as an example, molecular medicine, which attempts to explain diseases by the actions of molecules, should be classified as conceptual nanotechnology.

1.5 NANOBIOTECHNOLOGY AND BIONANOTECHNOLOGY

These widely used terms are almost self-explanatory. Nanobiotechnology is the application of nanotechnology to biology. For example, the use of semiconductor quantum dots as biomarkers in cell biology research would rank as nanobiotechnology. It encompasses "nanomedicine", which is defined as the application of nanotechnology to human health.

Bionanotechnology is the application of biology—which could be a living cell, or a biomolecule—to nanotechnology. An example is the use of the protein bacteriorhodopsin as an optically switched optical (nanophotonic) switch.

1.6 NANOTECHNOLOGY—TOWARD A DEFINITION

The current dictionary definition of nanotechnology is "the design, characterization, production and application of materials, devices and systems by controlling shape and size at the nanoscale".[8] (The nanoscale itself is at present consensually considered to cover the range from about 1 to 100 nm—see Section 1.7, but also footnote 7.) A slightly different nuance is given by the same source as "the deliberate and controlled manipulation, precision placement, measurement, modeling, and production of matter at the nanoscale in order to create materials, devices, and systems with fundamentally new properties and functions". The International Standards Organization (ISO) also suggests two meanings: (1) understanding and control of matter and processes at the nanoscale, typically, but not exclusively, below 100 nm in one or more dimensions where the onset of size-dependent phenomena usually enables novel applications; and (2) utilizing the properties of nanoscale materials that differ from the properties of individual atoms, molecules, and bulk matter, to create improved materials, devices, and systems that exploit these new properties. Another formulation encountered in reports is "the design, synthesis, characterization and application of materials, devices, and systems that have a functional organization in at least one dimension on the nanometer scale". The US Foresight Institute gives: "Nanotechnology is a group of emerging technologies in which the structure of matter is controlled at the nanometer scale to produce novel materials and devices that have useful and unique properties." The emphasis on control is particularly important: it is this that distinguishes nanotechnology from chemistry, with which it is often

[8] E. Abad et al., *NanoDictionary*. Basel: Collegium Basilea (2005).

compared; in the latter, motion is essentially uncontrolled and random, within the constraint that it takes place on the potential energy surface of the atoms and molecules under consideration. In order to achieve the desired control, a special, nonrandom *eutactic* environment needs to be available. Reflecting the importance of control, a very succinct definition of nanotechnology is simply "engineering with atomic precision"; sometimes the phrase "atomically precise technologies" (APT) is used to denote nanotechnology. However, we should bear in mind the "fundamentally new (or unique) properties" and "novel" aspects that many nanotechnologists insist upon, wishing to exclude ancient or existing artefacts that happen to be small.

1.7 THE NANOSCALE

Any definition of nanotechnology must also incorporate, or refer to, a definition of the nanoscale. As yet, there is no formal definition with a rational basis, merely a working proposal. If nanotechnology and nanoscience regard the atom (with size of the order of 1 ångström, i.e., 0.1 nm) as the smallest indivisible entity, this forms a natural lower boundary to the nanoscale. The upper boundary is fixed more arbitrarily. By analogy with microtechnology, now a well-established field dealing with devices up to about 100 micrometers in size, one could provisionally fix the upper boundary of nanotechnology as 100 nanometers. However, there is no guarantee that unique properties appear below that boundary (see Section 1.6).

The advent of nanotechnology raises an interesting question about the definition of the prefix "micro". An optical microscope can resolve features of the order of 1 micrometer in size. It is really a misnomer to also refer to instruments such as the electron microscope and the scanning probe microscope as "microscopes", because they can resolve features at the nanometer scale. It would be more logical to rename these instruments electron nanoscopes and scanning probe nanoscopes—although the word "microscope" is probably too deeply entrenched by now for a change to be possible.

1.8 NANOSCIENCE

This term is sometimes defined as "the science underlying nanotechnology"—but is this not biology, chemistry and physics—or the "molecular sciences"? It is the *technology* of designing and making functional objects at the nanoscale that is new; *science* has long been working at this scale and below. No one is arguing that fundamentally new physics, in the sense of new elementary forces, for example, appears at the nanoscale; rather it is new combinations of phenomena manifesting themselves at that scale that constitute the new technology. The term "nanoscience" therefore appears to be superfluous if it is used in the sense of "the science underlying nanotechnology", although as a synonym of conceptual nanotechnology it might have a valid meaning as the science of mesoscale approximation.

The molecular sciences, it will have been noted, include the phenomena of life (biology), which do indeed emerge at the nanoscale (although without requiring new elementary laws).

FURTHER READING

M. Arikawa, Fullerenes—an attractive nano carbon material and its production technology. *Nanotechnol. Perceptions* 2 (2006) 114–121.

B.O. Boscovic, Carbon nanotubes and nanofibres. *Nanotechnol. Perceptions* 3 (2007) 141–158.

K.E. Drexler, *Engines of Creation*. New York: Anchor Books/Doubleday (1986).

R. Feynman, There's plenty of room at the bottom. In: H.D. Gilbert (ed.), *Miniaturization*, pp. 282–296. New York: Reinhold (1961).

E. Kellenberger, Assembly in biological systems. In: *Polymerization in Biological Systems*, CIBA Foundation Symposium 7 (new series). Amsterdam: Elsevier (1972).

J.J. Ramsden, What is nanotechnology? *Nanotechnol. Perceptions* 1 (2005) 3–17.

CHAPTER 2

Science, Technology and Wealth

CHAPTER CONTENTS

Our knowledge about the universe grows year by year. There is a relentless accumulation of facts, many of which are reported in scientific journals, but also at conferences (and which may, or may not, be written down in published conference proceedings) and in reports produced by private companies and government research institutes (including military ones) that may never be published—and some work is now posted directly on the internet, in a preprint archive, or in an online journal, or on a personal or institutional website. The printed realm constitutes the scientific literature.[1] Reliable facts, such as the melting temperature of tungsten, count as unconditional knowledge. Such knowledge does not depend on the particular person who carried out the measurement, nor even on human agency (although the actual

[1]In passing, it may be noted that this realm is the only one for which consequential source quality appraisal can be carried out. On this point see, e.g., W. Wirth, The end of the scientific manuscript? *J. Biol. Phys. Chem.* 2 (2002) 67–71.

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

manner of carrying out the experimental determination depends on both). The criterion of reliability is above all repeatability.[2] These facts are discovered in the same way that Mungo Park discovered the upper reaches of the River Niger.

There is also what is called conditional knowledge: inductive inferences drawn from those facts, by creative leaps of human imagination. These are (human) inventions rather than discoveries. Newton's laws (and most laws and theories) fall into this category. They represent, typically, the vast subsuming of pieces of unconditional knowledge into highly compact form. Big tables of data giving the positions of the planets in our solar system can be summarized in a few lines of code—and the same lines can be used to calculate planetary positions for centuries in the past, and to predict them for centuries into the future. Despite the power of this procedure, some people have called this activity of making inferences summarizing data superfluous—the most famous protagonist of this viewpoint probably being William of Ockham, whose proverbial razor was designed to cut off all inductive inferences, all theories, not only overly elaborate ones. However, we must recognize that inductive inference is the heart and soul of

[2]An important aspect of ensuring the reliability of the scientific literature is the peer review to which reports submitted to reputable scientific journals are subjected. Either the editor himself or a specialist expert to whom the task is entrusted ad hoc carefully reads the typescript submitted to the journal, and points out internal inconsistencies, inadequate descriptions of procedures, erroneous mathematical derivations, relevant previous work overlooked by the authors, and so forth. The system cannot be said to be "perfect"—the main weaknesses are: the obvious fact that the reviewer cannot himself or herself actually check the experiments by running them again in his or her laboratory, or verify every step of a lengthy theoretical work, which would take as long as doing the work in the first place; the temptation to undervalue work that contradicts the reviewer's own results; and the pressures imposed by publishers when they are commercial organizations, in which case an additional publishability criterion is whether the paper will sell well, which tends to encourage hyperbole rather than a humbler, more sober and honest style of reporting. Despite these flaws, it would be difficult to overestimate the importance of the tremendous (and entirely honorary) work carried out by reviewers. This elaborate refining process creates a gulf between the quality of work finally published in a printed journal and the web-based preprint archives, online journals and other websites. Conference proceedings are in an intermediate position, some papers being reviewed before being accepted for presentation at a conference, but naturally the criteria are different because the primary purpose of a conference is to report work in progress rather than a completed investigation and the discussions of papers represent a major contribution to their value, yet might not even be reported in the proceedings. As regards the internal reports of companies and government research institutes, although these would not be independently and objectively peer-reviewed in the way that a submission to a journal is, insofar as the report will deal with something of practical value to the institution producing it, it is unlikely to be a repository of uncertain information.

science, and John Stuart Mill and others seem to have been close to the truth when they asserted that only inductive, not deductive, knowledge is a "real" addition.

There is nothing arcane about the actual description of the theories (although the process by which they are first arrived at—what we might call the flash of genius—remains a mystery). In the course of an investigation in physics and its relatives, at any rate, the facts (the primary observations) must first be mapped onto numbers—integer, real, complex or whatever. This mapping is sometimes called modeling. Newton's model of the solar system, for example, maps a planet with all its multifarious characteristics onto a single number representing a point mass. Then, the publicly accepted rules of mathematics, painstakingly established by generations of mathematicians working out proofs, are used to manipulate those numbers and facilitate the perception of new relations between them.

What motivates this growth of knowledge? Is it innate curiosity, as much a part of human nature as growth in physical stature and mental capabilities? Or is it driven by necessity, to solve problems of daily survival? According to the former explanation, curiosity led to discoveries, which in turn led to practical shortcuts (i.e., technology)—for the production of food in the very early era of human existence and later on for producing the artificial objects that came to be seen as indispensable adjuncts to civilization. Many of these practical shortcuts would involve tools, and later machines, hence the accumulation of possessions, in other words wealth. As will be discussed in Part 3, the whole "machinery" of this process constitutes an indivisible system incorporating also libraries and, nowadays, the internet.

This pattern (Figure 2.1) was later promoted by Francis Bacon in his book *The Advancement of Learning* (1605) with such powerful effect that it thereafter became part of the policy of many governments, remaining so to the present. Bacon was struck by the tremendous political power of Spain in his day. It seemed to heavily preponderate over that of Britain. He ascribed it to technology, which directly resulted from scientific discoveries, which were in turn deliberately fostered (as he believed) by the Spanish government. Nearer our own time, in the Germany of Kaiser Wilhelm, a similar policy was followed (as exemplified most concretely by the foundation of the Kaiser Wilhelm institutes). In Bacon's mind, as in that of Kaiser Wilhelm, the apotheosis of technology was military supremacy, perceived as the key to political hegemony, the political equivalent of commercial monopoly. Today, the British government, with its apparatus of research councils funding science that must be tied to definite applications with identifiable beneficiaries, is aiming at commercial rather than political advantage for the nation but the basic idea is the same. Similar policies can be found in the USA, Japan

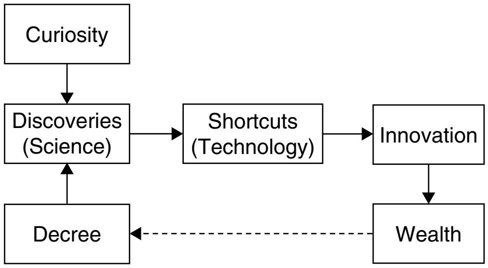


FIGURE 2.1

Sketch of the relationship between science and technology according to the curiosity or decree-driven ("linear") model. According to this view, technology can be considered as applied science. The dashed line indicates the process whereby one state, envious of another's wealth, may seek to accelerate the process of discovery.

and elsewhere. This model is also known as "linear"; because of the link to government it is also known as the "decree-driven" model.

Bacon's work was published 17 years after the failure of the Spanish Armada, which supposedly triggered his thoughts on the matter. Moreover, during that interval, although the threat was almost palpable, the feared Counter-Armada never materialized. This singular circumstance does not seem to have deflected Bacon from his vision, any more than the failure of Germany's adherence to this so-called "linear model" (science leading directly to technology) to deliver victory in the First World War deflected other governments from subsequently adhering to it. Incidentally, these are just two of the more striking pieces of evidence against that model, which ever since its inception has failed to gather solid empirical support.

The alternative model, which appears in much better concord with known facts,[3] is that technology, born directly out of the necessity of survival, enables leisure by enhancing productivity, and a small part of this leisure is used for contemplation and scientific activity (Figure 2.2), which might be described as devising ever more sophisticated instruments to discover ever more abstruse facts, modeling those facts, and inferring theories. The motivation for this work seems, typically, to be a mixture of curiosity *per se* and the desire to enhance man's knowledge of his place in the universe. The latter, being akin to philosophy, is sometimes called natural philosophy, a name still used to describe the science faculties in some universities. Those theories might then be used to enhance technology, probably by others than those who invented the theories, enabling further gains in productivity, and hence yet more leisure, and more science. Note that in this model, the basic step of creative ingenuity occurs at the level of technology; that is, the practical man confronted with a problem (or simply filled with the desire to minimize effort) hits upon a solution in a flash of inspiration.

[3]Not least the fact that technology has existed for many millennia, whereas science—in its modern sense, as used in all the figures in this chapter—only began in the 12th century CE.

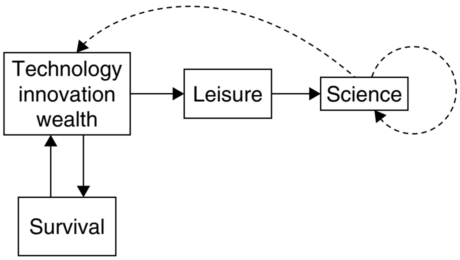


FIGURE 2.2 *Sketch of the relationship between science and technology according to the "alternative" model. Technology-enabled increases in productivity allowed Man to spend less than all his waking hours on the sheer necessities of survival. Some part of each day, or month, could be spent in leisure, and while part of this leisure time would be used simply to recuperate from the strains of labor (and should therefore be counted as part of production, perhaps), part was used in contemplation of the world, and of events, and sometimes this contemplation would lead to inferential leaps of understanding, adding mightily to knowledge. New knowledge leads to further practical shortcuts, more leisure, and so forth, therefore the development is to some degree autocatalytic (sometimes stated as "knowledge begets knowledge"). The dashed lines indicate positive feedback. According to this view, science can be considered as "applied technology".*

A further refinement to this alternative model is the realization that the primary driver for technological innovation is often not linked directly to survival, but is aesthetic. Cyril S. Smith has pointed out, adducing a great deal of evidence, that in the development of civilization decorative ceramic figurines preceded cooking utensils, metal jewellery preceded weapons, and so forth.[4]

Both models adopt the premise that technology leads to wealth. This would be true even without overproduction (i.e., production in excess of immediate requirements), because most technology involves making tools (i.e., capital equipment) that have a relatively permanent existence. Wealth constitutes a survival buffer. Overproduction in a period of plenty allows life to continue in a period of famine. It also allows an activity to be kick-started, rather like the birth of Venus. The Spanish Armada was essentially financed by the vast accumulation of gold and other precious metals from the newly won South American colonies, rather than wealth laboriously accumulated through the working of the linear model, as Bacon imagined.

The corollary is that science cannot exist without wealth. The Industrial Revolution was in full, impressive swing by the time Carnot, Joule and others

[4]C.S. Smith, *A Search for Structure*. Cambridge, MA: MIT Press (1981).

made their contributions to the science of thermodynamics. James Watt had no need of thermodynamics to invent his steam engine, although the formal theoretical edifice built up by the scientists later enabled many improvements to be made to the engine. Similarly, electricity was already in wide industrial use by the time the electron was discovered in the Cavendish Laboratory of Cambridge University.

Of course, in society benefits and risks are spread out among the population. Britain accumulated wealth through many diverse industries (Joule's family were brewers, for example). Nowadays, science is almost entirely carried out by a professional corps of scientists, who in the sense of the alternative model (Figure 2.2) spend all their time in leisure; the wealth of society as a whole is sufficient to enable not only this corps to exist, but also to enable it to be appropriately educated—for unlike the creative leaps of imagination leading to practical inventions, the discovery of abstruse facts and the theories inferred from them requires many years of hard study and specialized training.

2.1 NANOTECHNOLOGY IS DIFFERENT

We can, then, safely assert that all technological revolutions that had such profound effects on our civilization (steam engines, electricity, radio, and so forth) began with the technology, and the science (enabled by the luxury of leisure that the technologies enabled) followed later—until the early decades of the 20th century. The discovery of radioactivity and atomic (nuclear) fission were purely scientific discoveries, and their technological offshoot, in the form of the atomic pile (the first one of which was constructed around 1942), was devised by Enrico Fermi, one of the leading nuclear theoreticians, and his colleagues working on the Manhattan project. The rest—nuclear bombs and large-scale electricity generating plants—is, as they say, history. This "new model", illustrated in Figure 2.3, represents a radical departure from the previous situation. In the light of what we have said above, it begets the question "how is the science paid for?", since it is not linked to any existing wealth-generating activity. The answer appears to be twofold. Firstly, wealth has been steadily accumulating on Earth since the dawn of civilization, and beyond a certain point there is simply enough wealth around to allow one to engage in luxuries such as the scientific exploration of wholly new phenomena without any great concern about affordability. We conjecture that this point was reached at some time early in the 20th century. Secondly, governments acquired (in Britain, largely due to the need to pay for participation in the First World War) an unprecedented ability to gather large quantities

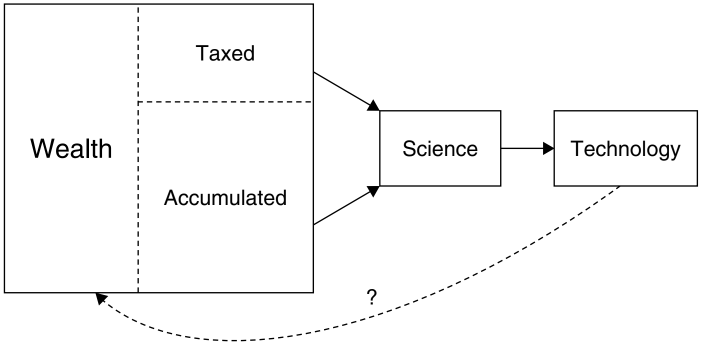


FIGURE 2.3 *The "new model" relating wealth, science and technology, applicable to the nuclear industry and nanotechnology. Note the uncertainty regarding the contribution of these new industries to wealth. There are probably at least as many opponents of the nuclear industry (who would argue that it has led to overall impoverishment; e.g., due to the radioactive waste disposal problem) as supporters. In this respect the potential of nanotechnology is as yet unproven.*

of money from the general public through taxation. Since governments are mostly convinced of the validity of the "linear model", science thereupon enjoyed a disproportionately higher share of "leisure wealth" than citizens had shown themselves willing to grant freely in the preceding century. The earlier years of the 20th century also saw the founding of a major state (the USSR) organized along novel lines. As far as science was concerned, it probably represented the apotheosis of the linear model (qua "scientific socialism"). Scientific research was seen as a key element in building up technical capability to match that of the Western world, especially the USA. On the whole, the policy was vindicated by a long series of remarkable achievements, especially and significantly in the development of nuclear weapons, which ensured that the USSR acquired superpower status.[5]

We propose that this "new model" applies to nanotechnology. Several reasons can be adduced in support. One is the invisibility of nanotechnology. Since atoms can only be visualized and manipulated using sophisticated nanoscopes, and hence do not form part of our everyday experience, they are not likely to form part of any obvious solution to a problem.[6] Another

[5]See D. Holloway, *Stalin and the Bomb*. New Haven: Yale University Press (1994).

[6]It should, however, be borne in mind that these nanoscopes are themselves products of a highly sophisticated *technology*, not science (one may also note that the motivation for developing electron microscopes included a desire to characterize the fine structure of materials used in technological applications).

reason is the very high worldwide level of expenditure, and corresponding activity, in the field, even though there is as yet no real nanotechnology industry.[7] In accordance with our insistence (Section 1.6) upon the novelty element needed by any technology wishing to label itself "nano", we do not include the silver halide-based photographic industry and the carbon black (automotive) industry (as will be elaborated in Chapter 5, neither truly fulfills the idea of atomically precise manufacturing).

2.2 THE EVOLUTION OF TECHNOLOGY

Human memory, especially "living memory", is strongly biased towards linearity. By far the most common mode of extrapolation into the future is a linear one. Unfortunately for this manifestation of "common" sense, examination of technology developments over periods longer than the duration of a generation shows that linearity is a quite erroneous perception. Nowadays, there should be little excuse for persistence in holding the linear viewpoint, since most of us have heard about Moore's law, which states that the number of components (transistors, etc.) on an integrated circuit chip doubles every 18 months. This remarkably prescient statement (more an industry prediction than a law) has now held for several decades. But, as Ray Kurzweil has shown, exponential technology development applies to almost every technology—until, that is, some kind of saturation sets in.[8]

Of course, any exponential law looks linear provided one examines a short enough interval; that is probably why the linear fallacy persists. Furthermore, at the beginning of a new technology, an exponential function increases very slowly—and we *are* at the beginning of nanotechnology. Progress—especially in atom-by-atom assembly—is almost painfully slow at present. On the other hand, progress in information-processing hardware, which nowadays counts as indirect nanotechnology (cf. Section 1.3), is there for all to see. The ENIAC computer (*circa* 1947) contained of the order of 10^4 electronic components and weighed about 30 tonnes. A modern high-performance computer capable of 5–10 TFLOPS[9] occupies a similar volume. Formerly, for carrying out large quantities of simple additions, subtractions, multiplications and

[7] Apart from an appreciable industry, with a global turnover of around $750 million (electron microscopes and atomic force microscopes), servicing the needs of those developing nanotechnology.

[8] R. Kurzweil, *The Singularity is Near*. New York: Viking Press (2005).

[9] 1 TFLOPS is 10^{15} floating-point operations per second.

divisions, as required in statistics, for example, one might have used the Frieden electromechanical calculator that cost several thousand dollars and weighed several tens of kilograms; the same performance can nowadays be achieved with a pocket electronic calculator costing one dollar and weighing a few tens of grams.[10]

The improvements in the performance (speed, energy consumption, reliability, weight and cost) of computer hardware are remarkable by any standards. If similar improvements could have been achieved with motorcars, they would nowadays move at a speed of 3000 kilometers per hour, use one liter of petrol to travel a 100,000 kilometers, last 10,000 years, weigh 10 milligrams, and cost about 10 dollars! Some of these attributes, at least—or their functional equivalents—might be achievable with nanotechnology.

Kurzweil (loc. cit.) elaborates on the exponential growth model applicable to a single technology to place technology as a whole in the context of the evolution of the universe, in which it occupies one of six epochs:

Epoch 1: Physics and chemistry are dominant; the formation of atomic structures (as the primordial universe, full of photons and plasma, expands and cools).

Epoch 2: Biology emerges; DNA is formed (and with it, the possibility of replicating and evolving life forms; as far as we know today, this has only occurred on our planet, but there is no principal reason why it could not occur anywhere offering favourable conditions).

Epoch 3: Brains emerge and evolve; information is stored in neural patterns (both in a hard-wired sense and in the soft sense of neural activity; living systems thereby enhance their short-term survivability through adaptability, and hence the possibility of K-selection[11]).

Epoch 4: Technology emerges and evolves; information is stored in artificial hardware and software designs.

Epoch 5: The merger of technology and human intelligence; the methods of biology, including human intelligence, are integrated into the exponentially expanding human technology base. This depends on technology mastering the methods of biology (including human intelligence).

---

[10] Assertion of the "same performance" neglects psychology—human factors—the existence of which provide one of the reasons why design is so important.

[11] See Section 3.1.

Epoch 6: The awakening of the universe; patterns of matter and energy become saturated with intelligent processes and knowledge; vastly expanded human intelligence, predominantly nonbiological, spreads throughout the universe.

The beginning of Epoch 6 is what Kurzweil calls the singularity, akin to a percolation phase transition.

2.3 THE NATURE OF WEALTH AND VALUE

Wealth is defined as accumulated value. A wealthy country is one possessing impressive infrastructure—including hospitals, a postal service, railways, and huge and sophisticated factories for producing goods ministering to the health and comforts of the inhabitants of the country. It also possesses an educated population, having not only universal literacy and numeracy, but also a general interest in intellectual pursuits (as might be exemplified by a lively publishing industry, active theaters and concert halls, cafés scientifiques[12] and the like) and a significant section of the population actively engaged in advancing knowledge; libraries, universities and research institutes also belong to this picture. Thus, wealth has both a tangible, material aspect and an intangible, spiritual aspect.

This capital—material and spiritual—is, as stated, *accumulated* value. Therefore, we could replace "wealth" in Figures 2.1–2.3 by "value (part of which is refined and accumulated in a store)". We should therefore inquire what is value.

Past political economists (such as John Stuart Mill and Adam Smith) have distinguished between value in use and value in exchange (using money)."Value in use" is synonymous with usefulness or utility, perhaps the most fundamental concept in economics, and defined by Mill as the capacity to satisfy a desire or serve a purpose. It is actually superfluous to distinguish between value in use and value in exchange, because the latter, equivalent to the monetary value of a good (i.e., its price) is simply a quantitative measure of its value in use. A motivation for making a distinction might have been the "obvious" discrepancies, in some cases, between price and perceived value. But as soon as it is realized that we are only talking about averages, and that the distributions might be very broad, the need for the distinction vanishes. For some individual, a good might seem cheap—to him it is undervalued and

[12]These have become important forum for debating science issues. They began in Leeds in 1998, modeled on the café philosophique started in Paris in 1992.

a bargain—and for another the converse will be the case. Indeed it might be hard to find someone who values something at exactly the price at which it is offered for sale in the market. A difficulty arises in connection with human life, because there are some ethical grounds for placing infinite value upon it, which might be hard to accommodate in sums. But the insurance industry has solved the problem adequately for the purposes of political economy—it can be equated to anticipated total earnings over a lifetime.[13] A further difficulty arises regarding the possible additional stipulation that for something to have value, there must be some difficulty in its attainment. But here too the difficulty appears to be artificial. Gravity would be more valuable on the Moon than on Earth, where it has, apparently, zero value because it is omnipresent. But perhaps it has zero *net* value: for aviation it is a great nuisance but for motoring it is essential. Air is easily attainable but clean air is a different matter, and even in antiquity whole cities were abandoned because of insufferably bad air. Confusion may arise here because the mode of paying for air is different from that customary for commodities such as butter or sugar. Intrinsically, however, there is nothing terribly arcane about value, which heuristically at any rate we can equate with price, and there is not even any need for Pareto's ingenious and more general concept of ophelimity. It should be emphasized that value is always shifting. Certain components of a particular type of aircraft might be very expensive to manufacture, but once that aircraft is no longer in service anywhere in the world, stocks of spare parts become valueless. Mill erred when he tried to determine value relative to some hypothetical fixed standard. The value of almost everything is conditional on the presence of other things, and organized in an exceedingly complicated web of interrelationships.

If utility is considered as the most fundamental concept in economics, the relationship between supply and demand is considered to be the most fundamental law. According to this law, the supply of a good will increase if its price increases, and demand will increase if its price falls, the actual price corresponding to that level of supply exactly matching that of demand—considered to represent a kind of equilibrium. Demand for necessities is stated to be inelastic, because it diminishes rather slightly with increasing price, whereas demand for luxuries is called elastic, because it falls steeply as the price increases. However, this set of relationships has little predictive

[13]The reader may also recall King James V of Scotland's question "How much am I worth?", that was wittily answered by the miller of Middle Hill as "29 pieces of silver—one less than the value of our Saviour." (A. Small, *Interesting Roman Antiquities Recently Discovered in Fife*. Edinburgh: printed for the author and sold by John Anderson & Co. (1823)).

value. Most suppliers will fix the price of their wares based on a knowledge of the past, and adjustments can be and are constantly being made on the basis of feedback (numbers of units sold).[14] Because there is a finite supply of many goods (since we live on a finite planet), their supply cannot increase with increasing price indefinitely; on the other hand, the supply of services could in principle be increased indefinitely *pari passu* with demand.[15]

There have, of course, been numerous attempts to elaborate the simple law of supply and demand. One interesting decomposition of demand is that of Noritaki Kano, into three components: basic, performance, and excitement. For example, the basic needs of the prospective buyer of a motor-car are that it is safe, will self-start reliably, and so forth. Even if the supplier fulfills them to the highest possible degree, the customer will merely be satisfied in a rather neutral fashion, but any deficiency will evoke disappointment. In other words, these attributes are essentially privative in nature. Performance (e.g., fuel consumption per unit distance traveled) typically increases continuously with technological development; customer satisfaction will be neutral if performance is at the level of the industry average; superior performance will evoke positive satisfaction. Finally, if no special effort has been made to address excitement needs (which are not always explicitly expressed, and may indeed only be felt subconsciously), customer satisfaction will be at worst neutral, but highly positive if the needs are addressed. These three components clearly translate directly into components of value.

2.4 THE SOCIAL VALUE OF SCIENCE

Francis Bacon argued in his *Advancement of Learning* (1605) that science discovery should be driven not just by the quest for intellectual enlightenment, but also for the "relief of man's estate". This view is, naturally enough, closely associated with Bacon's "linear" model of wealth creation (Figure 2.1), and forms the basis of the notion (nowadays typically promulgated by state funders of scientific research) that feeding into technological development and wealth creation is an official duty incumbent upon those scientists in receipt of state funds for their work. According to the "alternative model" (Figure 2.2) on the other hand, a scientist voluntarily devotes a part of his or her leisure to research, and there is no especial duty to explicitly promote wealth creation. However, the modern situation of a professional corps of

[14] One of the problems faced by commercial operators is the difficulty of "reading" feedback (let alone responding to it).

[15] Not least since the suppliers of the services mostly themselves require the same services.

scientists who are in effect paid by society to devote their whole time to leisure (which they in turn typically wholly devote to research) would appear unarguably to give society the right to demand a specific contribution to the creation of wealth on which, ultimately, the continuation of this arrangement depends.

When seeking to analyze the present situation and attempting to present a reasonable recommendation, shifting perspectives during the last few hundred years must be duly taken into account. The Industrial Revolution and the immense wealth it generated managed very well without (or with very little) science feeding into it, but during the last hundred years or so science has become increasingly associated with obtaining mastery over nature. A survey of the papers published in leading scientific journals indeed shows that a majority is directly concerned with that. However, this work was in general undertaken in a piecemeal fashion. For example, I gather that H.E. Hurst's seminal work on the analysis of irregular time series was undertaken at his own initiative while he was engaged as Scientific Consultant to the Ministry of Public Works in Egypt, when he was confronted with the need to make useful estimates of the required capacities of the dams at that time being proposed for construction on the Nile. In some cases scientific results were made use of with excellent results; in others with disastrous results;[16] there are many other examples of both excellent results and disasters obtained without any scientific backing. Hence, historical evidence does not allow us to conclude that a scientific research backing guarantees success in a technological endeavor, but rather shows that many other factors, most prominently political ones, intervene. One very positive aspect is that at least this decoupling of science from technology prevented the growth of distortions in the unfettered, disinterested pursuit of objective truth, which almost inevitably becomes a casualty if wealth instead is pursued.

But when it comes to the "new model" (Figure 2.3), we have technology wholly dependent upon science; in other words, without the science there would be no technology and, as already stated, nanotechnology seems to fall into this category. Further implications will be explored in Chapters 3 and 9.

We cannot usefully turn to historical evidence on this point because too little has accumulated. It follows that any extrapolation into the future is likely to be highly speculative. Nevertheless, we cannot rule out the advent of a new era of highly effective science-based handling of affairs that would hopefully yield excellent results. Although the economies and especially the

[16]The Kongwa (Tanganyika) groundnut scheme of the Overseas Food Corporation serves as an example of a disastrous outcome.

banking sectors of most countries of the world are now rather fragile, to which the response in many circles is rather conservative retrenchment, this is just the wrong kind of response. The whole system of the planet (ecological, social, industrial, financial, and so forth) has been driven so hard to such extremes that mankind can scarcely afford to make more mistakes, in the sense that there is practically no buffering capacity left. Hence in a very real sense survival will depend on getting things right. The delicacy of judgment required from the decision-making process is further exacerbated by globalization, thanks to which we now in effect only have one "experiment" under way, and failure means collapse of everything, not just a local perturbation.

FURTHER READING

J.D. Bernal, *The Social Function of Science*. London: Routledge (1939).

J. Pethica, T. Kealey, P. Moriarty and J.J. Ramsden, Is public science a public good? *Nanotechnol. Perceptions* 4 (2008) 93–112.

J.J. Ramsden, S. Aida and A. Kakabadse (eds), *Spiritual Motivation: New Thinking for Business and Management*. Basingstoke: Palgrave Macmillan (2007).

CHAPTER 3

Innovation

CHAPTER CONTENTS

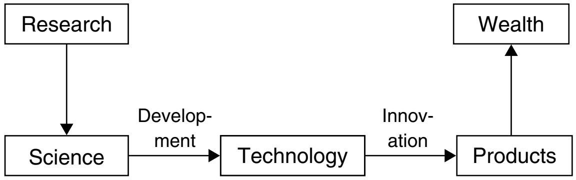
Although the dictionary definition of “innovation” is simply “the bringing in of novelties”, it has in recent years become a more narrowly defined concept much beloved especially by government ministries and their agencies charged with animating economic activity in their countries. Indeed in 2007 the UK government, which has been in the van of this process, created a new Department of Innovation, Universities and Skills, revealingly linking innovation with universities. In this usage, innovation has come to mean specifically the process whereby new products are introduced into the commercial sphere: “The technical, designing, manufacturing, management and commercial activities involved in the marketing of a new (or improved) product or the first commercial use of a new (or improved) process or equipment”.[1] It implies not only the commercialization of a major advance in

[1] C. Freeman, *The Economics of Industrial Innovation*. London: Frances Pinter (1982).

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

FIGURE 3.1

Detail of the transformation of science to wealth, applicable to both the "linear" and "new" models.



the technological state of the art, but also "includes the utilization of even small-scale changes in technological know-how".[2] Thomas Alva Edison was not only a brilliant inventor but also a masterful innovator (who is reputed to have said "it's 1% inspiration and 99% perspiration"); however, the inventor is very often not the innovator. Suction sweepers are associated not with Spengler, their inventor, but with Hoover; similarly the sewing machine is associated with Isaac Merrit Singer, not with Elias Howe, and still less with Barthélemy Thimonnier or Thomas Saint.[3]

The innovator is thus crucial to the overall process of wealth creation. The concept of innovation can be naturally entrained in the "linear model" (Figure 2.1). If we define "high technology" as "technology emerging from science", then nanotechnology is clearly a high technology, according to the "new model" (Figure 2.3) outlined in the previous chapter, and the process of innovation, in its new constrained usage (of introducing novel products into the commercial sphere), is likely to be highly relevant. Figure 3.1 shows more explicitly how science can be transformed into wealth via innovation.

It is not hard to find reasons for the flurry of official interest in the topic. Governments have noticed that a great deal of research, financed from the public purse, appears to be of very little strategic importance.[4] Even though

---

[2] R. Rothwell, Successful industrial innovation. *R & D Management* 22 (1992) 221–239.

[3] Thimonnier, indeed, narrowly escaped with his life when tailors smashed his machines in fear of losing their livelihoods.

[4] This is, at root, an indictment of the system of scientific research funded on the basis of the so-called competitive grant proposals. Scientists learn as undergraduates to select problems of importance to work on (I well remember this advice from Sir Peter Medawar given in his lecture "Advice to a Young Scientist"), and left to themselves that is what they will do during their research careers. Unfortunately the "competitive grant proposal" system does not leave them to themselves. If resources beyond the brain, the hand and pencil and paper are required, funding must be secured by submitting a research proposal to a research council. The research must be described in great detail, and if the proposal is accepted and the research is funded, the scientist is then required to follow his submitted plans to the letter. They may have seemed reasonable at the time the proposal was written, but new knowledge is constantly being discovered and invented, and the scientist is assimilating some of that and having new thoughts of his own, not least while undertaking the research that had been

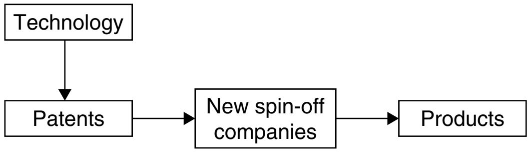


FIGURE 3.2 *Detail of the transformation of technology to products, applicable to all the models. New technologies could of course be patented by large established companies as well, but nowadays it is more typical for such companies to buy spin-offs in order to acquire a new technology.*

the funding and execution of scientific research is not, in most countries, prominent in the public mind, nevertheless governments feel that they have to justify public spending, even minor portions of the total.[5] The justification of funding scientific research therefore becomes its capacity to generate wealth through innovation, and the partial convergence of the "new model" with the "linear model" (although they are not isomorphous) allows the old tradition of Baconian thinking to continue.[6]

Innovation, in the sense of the implementation of discovery, or how research results are turned into products, is a theme at the heart of this book (cf. Chapter 9). Governments have become particularly wedded to the path shown in Figure 3.2. Given that the granting of a patent—in other words the

---

proposed. The weakness of the system is exacerbated by the slowness of the submission and approval process—roughly two years elapses between having the initial idea and starting the research council-funded research. And, inevitably with this way of organizing things, the report of the completed work to the research council becomes an end in itself. The scientist must be seen to have fulfilled what was contractually required of him, not least since the success of future proposals may depend on the quality of his file maintained by the research council. The actual result is of little consequence (cf. C.N. Parkinson, *In-Laws and Outlaws*, pp. 134–135. London: John Murray (1964)). I have myself noticed that at meetings convened to review projects, scientists reporting on their work increasingly refer merely to the number of "deliverables" that have been produced—such as the number of papers published, patents applied for or the amount of additional funding secured, with the most cursory attention being given to the actual content of those outputs—what one really wants to know is whether new insight and understanding have been generated. When these outputs *are* subjected to closer scrutiny, it often turns out that the basic ideas were published decades ago, and then simply forgotten or overlooked.

[5]For example, the funds disbursed by the UK Biotechnology and Biological Sciences Research Council amount to about one hundredth of government expenditure on health; and the UK's annual contribution to the facilities at CERN, impressive and costly as they are, amounts to a few pounds per head of the population—in other words, a pint or two of beer.

[6]Cf. I. Gibson and S.R.P. Silva, Harnessing the full potential of nanotechnology for wealth creation. *Nanotechnol. Perceptions* 4 (2008) 87–92.

right to monopolistically exploit an invention for a certain number of years—is a clear prerogative of governments, it is perhaps not surprising to find they have a vested interest in promoting patenting, regardless of the presence or absence of any overall economic benefit to the country (cf. Section 9.9).

Economists, especially J.A. Schumpeter, have noticed that established technologies sometimes die out, creating space for new ones. This phenomenon came to be called creative destruction. The man in the street expresses it through proverbs such as "you cannot make an omelette without breaking an egg", and biologists are also familiar with the idea, a good example being the death of about half the neurons at a certain epoch in the development of the brain of the embryonic chicken (and, I dare say, of other embryonic animals). At the time Schumpeter was putting forward the notion, it was widely believed that epochs of rapid multiplication of new species were preceded by mass destruction of existing ones.[7] It is not difficult to see why preceding destruction is an unnecessary condition for the occurrence of creative construction. Obviously a literally empty potential habitat has space for colonization (by so-called r-selection—see Section 3.1)—although if it is truly devoid of life initial colonization might be quite difficult. On the other hand, an apparently crowded habitat may be very rich in potential niches for new species capable of imaginatively exploiting them (the so-called K-selection—see Section 3.1). The scientist specializing in biomolecular conformation will be familiar with the fact that for ribonucleic acid (RNA) polymers to adopt their final stable structure, intramolecular bonds formed while the polymer is still being synthesized have subsequently to be broken.[8]

Figure 3.1 omits details about the process whereby the new products are transformed into wealth. Evidently, in order to do so people must want to buy the products—in other words, there must be a market for them. For incremental technologies, demand for novelty typically comes from buyers of existing products. Directly or indirectly, manufacturers receive feedback from buyers (including the manufacturers' own employees), which can more or less straightforwardly be worked into a steadily improving product. This situation is referred to as "market pull". Disruptive technologies, by definition, are qualitatively different from those in existence at the moment of

[7] Subsequent, more detailed knowledge of the palaeontological record makes this belief untenable; a striking example is the fact that the "Cambrian explosion", perhaps the most remarkable emergence of new species known, was not preceded by a mass extinction event.

[8] A. Fernández, Kinetic assembling of the biologically active secondary structure for *CAR*, the target sequence for the *Rev* protein of HIV-1. *Arch. Biochem. Biophys.* 280 (1990) 421–424.

their emergence. Any user of an existing technology sufficiently farsighted to imagine a qualitatively different solution to his problem is likely himself to be the innovator. Therefore, market pull is inapplicable; one refers to technology push, or the technological imperative. The development of technology is considered to be autonomous, and the emergence of new technologies determines the desire for goods and services.[9]

3.1 THE TIME COURSE OF INNOVATION

By analogy with biological growth, a good guess for the kinetics would be the sigmoidal logistic equation

$$Q(t) = K/\{1 + \exp[-r(t-m)]\} \tag{3.1}$$

where Q is the quantity under observation (the degree of innovation, for example), K is the carrying capacity of the system (the value to which Q tends as time $t \to \infty$), r is the growth rate coefficient, and m is the time at which $Q = K/2$ and $\mathrm{d}Q/\mathrm{d}t = r$. The terms r-selection and K-selection can be explained by reference to this equation: the former operates when a niche is relatively empty and everything is growing as fast as it can, therefore the species with the biggest r will dominate; the latter operates when an ecosystem is crowded, and dominance must be achieved by increasing K. This is perhaps more easily seen by noting that equation (3.1) is the solution to the differential equation

$$\mathrm{d}Q/\mathrm{d}t = rQ(1 - Q/K). \tag{3.2}$$

The application of this equation to innovation implies, perhaps a little surprisingly, that innovation grows autonomously; that is, it does not need any adjunct (although, as written, it cannot start from zero). Perhaps, indeed, the lone innovator is still a leading figure. Hirooka has gathered some evidence for this time course, the most extensive being for the electronics industry.[10] He promulgates the view that innovation comprises three successive logistic curves: one each for technology, development and diffusion. "Development" is used by Hirooka in a sense different from that of Figure 3.1, in which research leads to science (i.e., the accumulation of scientific knowledge),

[9]See J. Hodgkinson et al., Gas sensors 2: The markets and challenges, *Nanotechnol. Perceptions* 5 (2009) 83–107 for further discussion of this point.

[10]M. Hirooka, Complexity in discrete innovation systems. *E:CO* 8 (2006) 20–34.

and development of that science leads to technology, out of which innovation creates products such as the personal computer. There seems to be no need to have separate "development" and "diffusion" trajectories: these taken together constitute innovation. In Hirooka's electronics example, the technology trajectory begins with the point contact transistor invented in 1948, and m is reached in about 1960 with the metal oxide-semiconductor transistor and the silicon-based planar integrated circuit. This evidence is not, however, wholly satisfactory, not least because there seems to be a certain arbitrariness in assigning values of Q. Furthermore, why the trajectory should end with submicron lithography in 1973 is not clear. The continuation of Moore's law up to the present (and it is anticipated to continue for at least several more years) implies that we are still in the exponential phase of technological progress. The "development" trajectory is considered to begin with the UNIX operating system in 1969 and continues with other microprocessors (quantified by the number of components on the processor chip, or the number of memory elements) and operating systems, with m reached in about 1985 with the Apple Macintosh computer; the diffusion trajectory is quantified by the demand for integrated circuits (chips).

Perhaps Hirooka's aim was only to quantify the temporal evolution; at any rate, he does not offer a real explanation of the law that he promulgates, but seems to be more interested in aligning his ideas with those of the empirical business cycles of Kondratiev and others.[11] For insight into what drives the temporal evolution of innovation, one should turn to consideration of the noise inherent in a system (whether socio-economic, biological, mechanical, etc.).[12] Some of this noise (embodied in random microstates) is amplified up to macroscopic expression,[13] and provides a potent source of microdiversity.

---

[11] However, a fundamental critique of the cycles is that they fail to take the steady accumulation of knowledge into account. Although the colorful phrase "creative destruction" carries with it the innuendo of *tabula rasa*, of course things are not really like that; although many firms (considered as the basic units of innovation) are destroyed, the hitherto accumulated knowledge remains virtually intact, because of which history cannot really repeat itself, certainly not to the extent of driving a cycle with unchanging period and amplitude, unless some very special regulatory mechanism is operating (but this is not what is being suggested). In a similar fashion, even though past mass extinctions destroyed up to 90% of all living species, the records of the entire past remained encoded in the DNA of the survivors.

[12] P.M. Allen, M. Strathern and J.S. Baldwin, Evolutionary drive. *E:CO* 8 (2006) 2–19.

[13] R. Shaw, Strange attractors, chaotic behaviour, and information flow. *Z. Naturforsch.* 36a (1981) 80–112.

Equation (3.2) should therefore be replaced by

$$dQ/dt = rQ(1 - Q/K) + \xi(t), \tag{3.3}$$

where ξ is a random noise term (a more complete discussion than we have space for here would examine correlations in the noise). This modification also overcomes the problem that equation (3.2) cannot do anything if Q is initially zero.

The amplification of the noise up to macroscopic expression is called by Allen "exploration and experiment". Any system in which mechanisms of exploration and experiment are suppressed is doomed in any environment other than a fixed, unchanging one, although in the short term exploration and experiment are expensive (they could well be considered as the price of long-term survival).

Recognition of microdiversity as the primary generator of novelty does not in itself provide clues to its kinetics. It may, however, be sufficient to argue from analogy with living systems. By definition, a novelty enters an empty (with respect to the novelty) ecosystem; growth is only limited by the intrinsic growth rate coefficient (the r-limited régime in ecology). Inevitably as the ecosystem gets filled up, crowding constraints prevent exponential growth from continuing.

One may legitimately ask whether the first positive term in equation (3.3) should be proportional to Q. Usually innovation depends on other innovations occurring concurrently. Kurzweil comments that technology can sometimes grow superexponentially. Equation (3.1) should therefore only be taken as a provisional starting point. We need to consider that technological growth, dQ/dt, is proportional to Q^n, and carefully examine whether n is, in fact, greater than unity. For this we also need to work out how to place successive entities in a developing technology on a common scale of the degree of development. How much more developed is the MOS transistor than the p–n junction transistor? Possibly the complexity of the object especially the notion of thermodynamic depth,[14] might provide a practicable quantification. These matters remain to be investigated further.

During the post-m stage we enter the K-limited régime: survival is now ensured not through outgrowing the competition, but through ingenuity in exploiting the highly ramified ecosystem. The filling will itself create some new niches, but eventually the system will become saturated. Even factors

[14] S. Lloyd and H. Pagels, Complexity as thermodynamic depth. *Ann. Phys.* 188 (1988) 186–213.

such as the fatigue of university professors training the researchers and developers through repeatedly having to expound the same material plays a rôle.

3.2 CREATIVE DESTRUCTION

The development of the electronics industry is perhaps atypically smooth. Successive technologies usually overlapped the preceding ones, and the industry was generally at pains to ensure compatibility of each technological step with the preceding one. But, innovation is often a much more turbulent affair; one may characterize it using words like discontinuity, disruption, or revolution.

An early example of disruptive innovation is the stirrup, which seems to have diffused westwards from China around the 8th century CE, and in Europe was taken up on a large scale by the Franks led by Charles Martel. At a stroke it enabled the horse to be used far more effectively in warfare: with stirrups a knight could hold a heavy lance, the momentum of the horse would be added to his own, and the lance would be virtually unstoppable by any defenses then current. It also enabled arrows to be fired from a longbow by a mounted rider. This is a good example of technology push—there is no actual evidence that a group of knights sat down and decided this was what they wanted to enable them to fight more effectively in the saddle. It was a push that was to have far-reaching social consequences. Other armies adopted the innovation, and there was a concomitant increase in defensive technology, including armor and fortified castles. Warfare rapidly became significantly more expensive than hitherto. White has argued that this triggered a revolutionary social change[15]—to support the expense, land was seized by leaders like Martel and distributed to knights in exchange for military service, which they then fulfilled at their own expense. The knights in turn took control over the peasants who lived on the land, cultivating it and raising livestock. In other words, the stirrup led to the introduction of feudalism, a far greater revolution (in the sense that it affected far more people) than that of the technology *per se*.

The classification of innovations as either technology push (typically associated with disruptive innovation: by definition, the market cannot demand something it does not know about) or market pull (for incremental

---

[15] L. White, *Medieval Technology and Social Change*. New York: Oxford University Press (1962).

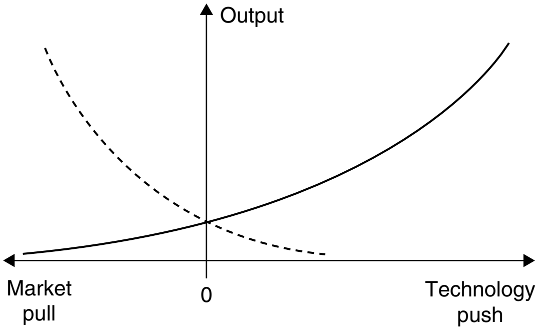


FIGURE 3.3 *Proposed quasi-equilibrium between technology push (solid line) and market pull (dashed line). The ideal level of output occurs where they exactly match each other. This diagram neglects consideration of possible temporal mismatch between push and pull.*

innovations, whereby technology responds to customer feedback) does not seem to cover all cases, however. There is currently no real demand for new operating systems for personal computers, for example, yet Microsoft, the market leader (in terms of volume), is constantly launching new ones. The innovation is incremental, yet customers complain that each successive one is worse than its predecessors (e.g., "Vista" compared with "XP"). Simple economic theory suggests that such products should never be introduced; presumably only the quasimonopolistic situation of Microsoft allows it to happen.

An extension to the basic push-pull concept is the idea of "latent demand". It can be identified post hoc by unusually rapid takeup of a disruptive innovation. By definition, latent demand is impossible to identify in advance; its existence can only be verified by experiment.

I suggest that in analogy to supply and demand, push and pull also (under certain circumstances, the special nature of which needs further inquiry) "equilibrate", as illustrated in Figure 3.3.

As already pointed out near the beginning of this chapter, the term "creative destruction" was introduced by Joseph Schumpeter, but it is in itself incomplete and inadequate for understanding disruptive innovation. It would be more logical to begin with the "noise", which at the level of the firm is represented by the continuous appearance of new companies—after all, nothing can be destroyed before it exists. Noise can, however, be both positive and negative. If the commercial *raison d'être* of the company disappears, then the company will presumably also disappear, along with others that depended on it. This process can be modeled very simply: if the companies are all characterized by a single parameter F, which we can call "fitness", and time advances in discrete steps (as is usual in simulations), then at each time step the least fit company is eliminated along with a certain number of its "neighbors" (in the sense of being linked by some kind of commercial dependence)

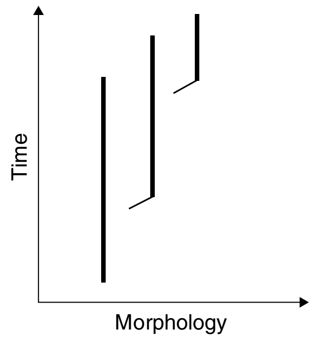


FIGURE 3.4 *Sketch of speciation according to the punctuated equilibrium concept. Thick vertical lines correspond to incremental innovations, and thin diagonal lines to disruptive innovations resulting in a change of technological "morphology".*

regardless of their fitnesses, and all are replaced by new companies with randomly assigned fitnesses.[16] This model was introduced as a description of the evolution of living species, and has a number of interesting properties such as a critical (self-organized?) fitness threshold, the height of which depends on the number of neighbors affected by an extinction. Furthermore, if the proper time of the model (successive time steps) is mapped onto real (universal or sidereal) time by supposing that the real waiting time for an extinction is proportional to $\exp(F)$, extinctions occur in well-delineated bursts ("avalanches") in real time, and the sizes of the bursts follow a power law distribution. Palaeontologists call this kind of dynamics "punctuated equilibrium" (Figure 3.4).[17]

The Bak-Sneppen model emphasizes the interdependence of species. Companies do not exist in isolation, but form part of an ecosystem. Ramsden and Kiss-Haypál have argued that the "economic ecosystem" (i.e., the economy) optimizes itself such that human desires are supplied with the least effort—a generalization of Zipf's law, whence it follows that the distribution of company sizes obeys:[18]

$$s_k = P(k + \rho)^{-1/\theta} \tag{3.4}$$

where s_k is the size of the kth company ranked according to size such that k is the rank, $k = 1$ being the largest company, P is a normalizing parameter, and ρ and θ are the independent parameters of the distribution, called, respectively, the competitive exclusion parameter and the cybernetic temperature.

[16] P. Bak and K. Sneppen, Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.* 71 (1993) 4083–4086.

[17] It is especially associated with the names of Ruzhnetsev, Gould and Vrba.

[18] J.J. Ramsden and Gy. Kiss-Haypál, Company size distributions in different countries. *Physica* A 277 (2000) 220–227.

Competitive exclusion means that in any niche, ultimately one player will dominate (this is, in fact, a simple consequence of general systems theory). The Dixit-Stiglitz model of consumer demand is another application.[19]

3.3 WHAT DRIVES DEVELOPMENT?

Is there any deeper underlying mechanism behind the dynamics represented by equation (3.3)? Since invention and innovation are carried out by human beings, one should perhaps look at their underlying motivations. An important principle would appear to be that everyone wants to do a good job—if they get the chance to do so.[20] It is only natural for technologists to respond to feedback with incremental innovation. Natural curiosity and the energy to follow it up is sufficient to provide a basis for ξ in equation (3.3). But the further growth of innovation, including the actual values of the parameters r and K, depends on other factors, including retarding ones such as inertia and friction. It depends on the fiscal environment, because much innovation requires strong concentration of capital. It also depends on "intellectual capital"—knowledge and skills—that must depend in some way on public education. Globalization means that, even more so than ever before, "no country is an island", and in consequence it becomes difficult to discern what elements of national policy favor innovation.

3.4 CAN INNOVATION BE MANAGED?

Given that the average lifetime of a firm is a mere 12 years,[21] one might suppose that the directors of even the largest and best-managed companies are constantly afraid of the possibility of sudden extinction. The management literature abounds with exhortations for companies to "remain agile" in order to be able to adapt and survive. But there is little that can be offered in the way of specific advice.

[19]The relationship (3.4) first emerged with a proper derivation in Mandelbrot's work on the analysis of the frequency of word usage in texts. Later it was applied to systems as diverse as the expression of proteins in bacteria (see B.B. Mandelbrot, Contribution à la théorie mathématique des jeux de communication, *Publ. Inst. StatistÚniv. Paris* 2 (1952) 1–124).

[20]P.A. Hunter, Creating sustained performance improvement. In: J.J. Ramsden, S. Aida and A. Kakabadse (eds), *Spiritual Motivation: New Thinking for Business and Management*, Ch. 15, pp. 185–206. Basingstoke: Palgrave Macmillan (2007).

[21]R. Foster and S. Kaplan, *Creative Destruction*. New York: Doubleday (2001).

A fruitful approach would appear to be to start with an empirical examination of whether ξ can be correlated with factors such as the percentage of personal income that is saved and, hence, available for concentrating in large capital enterprises; and the organization of public education in a country.

Such empirical examination can yield surprising results. A striking result of this nature is that although the number of patents granted to a company does correlate with its spending on research and development, there is no simple relationship between the spending and corporate performance: in other words money cannot simply buy effective innovation.[22]

New, truly disruptive technologies may require special attention paid to public acceptance. It is widely considered that the failure of companies developing genetically modified (GM) cultivated plant technology to foster open discussion with the public was directly responsible for the subsequent mistrust of foodstuffs derived from GM plants, mistrust that is especially marked in Europe, with huge commercial consequences. Problems associated with the lack of discussion were further exacerbated by the deplorable attitude of many supposedly independently thinking scientists, who often unthinkingly sided with the companies, unwarrantedly (given the paucity of evidence) assuming that ecosystems would not be harmed. Insofar as many scientists working in universities are nowadays dependent on companies for research funding, this attitude came close to venality and did nothing to enhance the reputation of scientists as bastions of disinterested, objective appraisal. Nanotechnologists are now being exhorted to pay heed to those mistakes and ensure that the issues surrounding the introduction of the technology are properly debated openly. There is, in fact, an unprecedented level of public dialog on nanotechnology and, perhaps as a direct consequence, a clear majority of the population seems to be well disposed towards it.

3.5 THE EFFECT OF MATURITY

One of the greatest discouragements to the introduction of innovation is plainly a high degree of maturity of the technology. This corresponds to the logistic curve (3.1) asymptotically approaching $Q = K$. A good example is the market for prostheses (hip, femur, etc.) Although many surgeons implanting them in patients have innovative ideas about how to improve their design, and new materials are emerging all the time, especially nanocomposites and materials with nanostructured surfaces promoting better assimilation with

[22] B. Jaruzelski, K. Dehoff and R. Bordia, *Smart Spenders: The Global Innovation 1000*. McLean, VA: Booz Allen Hamilton (2006).

the host tissue, the existing technology is already at such a high level in general it is extraordinarily difficult to introduce novelty into clinical practice. Unlike electronics, the biomedical field is heavily regulated. In many countries, onerous animal trials must be undertaken before a product is permitted to even be tested on humans. Furthermore, after years of consolidation (cf. Figure 9.5), global supply is dominated by two very large companies, both headquartered in the USA.

Another way of looking at this is to consider the market as a complex dynamical system with multiple basins of attraction. As proved by Ashby,[23] such a system will inevitably end up stuck in one of its basins and exploration will cease. This is the phenomenon of habituation. The system can only be reset if it receives a severe external shock. This is what Schumpeter was presumably trying to convey with his notion of "creative destruction".[24]

FURTHER READING

P.M. Allen and M. Strathern, Complexity, stability and crises. In: J.J. Ramsden and P.J. Kervalishvili (eds). *Complexity and Security*, pp. 71–92. Amsterdam: IOS Press (2008).

J.J. Ramsden, *Bioinformatics: an Introduction* 2nd edn. London: Springer (2009).

J.J. Ramsden, The representation of complexity. In: J.J. Ramsden and P.J. Kervalishvili (eds). *Complexity and Security*, pp. 93–102. Amsterdam: IOS Press (2008).

[23]W.R. Ashby, The mechanism of habituation. In: *NPL Symposium No 10, Mechanization of Thought Processes*. London: HMSO (1960).

[24]The possibility of intrinsic bursts of destruction should not be neglected, cf. the Bak-Sneppen model (footnote 16).

CHAPTER 4

Why Nanotechnology?

CHAPTER CONTENTS

With almost every manufactured product, if the same performance can be achieved by using less material, there will be a cost advantage in doing so. A well-known example is the metal beverage can. Improvements in design—including the formulation of the alloy from which it is made—have led to significantly less material being used for the same function (containing a beverage). In this example, there are concomitant, secondary advantages of miniaturization (e.g., because the can is also lighter in weight, it costs less to move around). There may be additional issues related to recyclability.

Note that in this case the site for miniaturization was the thickness of the wall of the can. The basic functional specifications of the can include the volume of beverage that must be contained. This cannot be miniaturized. On the other hand, if the wall could be made of a nanoplate, and still fulfill all requirements for strength and impermeability, it would have become a nanoproduct.

In the case of engineering products fulfilling a structural or mechanical purpose, their fundamental scale of size and strength is set by the human

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

body. The standard volume of beverage in a can is presumably based on what a human being likes to drink when quenching his thirst. Perhaps the innovator carried out experiments, much as George Stephenson determined the gauge standard for his railways by measuring the distance between the wheels for a hundred or so farm carts in the neighborhood of the Stockton and Darlington Railway and taking the mean, which happened to be $4'8\frac{1}{2}''$.[1]

The length and mechanical strength of a walking stick must be able to support the person using it. Miniaturization of such products therefore generally implies the use of thinner, stronger materials, which might well be nanocomposites, but nevertheless the length of the stick and the dimensions of the hand grip cannot be miniaturized.

Another major class of product deals with processing and displaying information. The venerable example is that of the clock, which (in a sense) computes, and displays, the time of day. Human eyesight places a lower limit on the useful size of the display and other input and output man/machine interfaces. In the case of mechanical clocks there is a fabrication issue: although a tiny wristwatch uses less material than a standard domestic interior clock, it is more expensive to make, both because the parts must be finished with higher precision, and because it is more troublesome to assemble them.

But what is the intrinsic lower limit of the physical embodiment of one bit of information (presence or absence)? Single electron electronics and Berezin's proposal for isotopic data storage[2] suggest that it is, respectively, one electron or one neutron; in other words one quantum, considered as the irreducible minimum size of matter. But a quantum is absolutely small, in the sense that observing its state will alter it[3]—which seems to suggest that

---

[1] It is said that Edward Pease, who led the consortium of businessmen promoting the railway, ordered him to make the width of the track equal to that of local country carts. The story is quite instructive as an example of how *not* to proceed. Railways represented a discontinuity with respect to the technology of farm carts. This was recognized by Stephenson's rival Isambard Brunel, who chose his gauge standard of 7 feet by considering the intrinsic possibilities of the new technology. Despite its technical superiority, the reputedly indomitable will of the Stephenson brothers ultimately prevailed, even rejecting the Rennie brothers' reasonable compromise of 5′6″—not only in Britain, but also in much of the rest of the world. It is surprising that the new high-speed railways in Japan were constructed using the "standard" $4'8\frac{1}{2}''$ gauge, since the national railway system had anyway a different gauge of 3′6″; a broader gauge would have allowed even greater speed, stability and on-board luxury.

[2] A.A. Berezin, Stable isotopes in nanotechnology. *Nanotechnol. Perceptions* 5 (2009) 27–36.

[3] P.A.M. Dirac, *The Principles of Quantum Mechanics*, 4th edn. Oxford: Clarendon Press (1958).

it is useless for the intended purpose. Only in the quantum computer is the possibility exploited that the quantum object can exist in a superposition of states (observation generally forces the elimination of the superposition). Quantum logic therefore implies virtually unlimited parallelism (superposition). Although intensive research work is currently being undertaken to develop quantum computers, this development has yet to bear fruit in the shape of a working device and therefore, strictly speaking, falls outside the scope of this book, which is focused on actual products.

Conventional logic, in which something is either present or absent and in which the superposition of both presence and absence does not exist must therefore be embodied in objects larger than individual quanta. The lower size limit of this physical embodiment seems to be a single atom.[4] In principle, therefore, it seems that information storage (memory) could be based on cells capable of containing a single atom, provided what is being observed is not a quantum state, without any loss of functionality.

The most dramatic progress in miniaturization has therefore occurred in information processing.[5] In this case, the fabrication technology has undergone a qualitative change since Jack Kilby's first integrated circuit. Making large-scale integrated circuitry in the same way that the first integrated groups of components were made—the mode of the watchmaker—would be prohibitively expensive for a mass-market commodity. Semiconductor processing technology, however, combines miniaturization with parallelization. Not only have the individual components become smaller, but the area processed simultaneously has dramatically increased (measured by the standard diameter of the silicon wafers, which has increased from 3 inches up to 12 inches).

Within the processor, miniaturization means not only having to use a smaller quantity of costly material, but also shorter distances between components. Since information processing speed is limited by the time taken by the information carriers—electrons—to traverse a component, processing has become significantly faster as a result. Furthermore, since information processing is irreversible, heat is dissipated, and miniaturization also miniaturizes the quantity of heat dissipated per logical operation. The miniaturization has therefore gone beyond maintaining the same performance using less material, but has actually enhanced performance.

---

[4] Cf. Berezin's proposal for data storage, footnote 2.

[5] Nevertheless, there is still a long way to go—a memory cell $100 \times 100 \times 100$ nm in size still contains of the order of 10^9 atoms.

Nevertheless, regardless of the actual sizes of the circuits in which the information processing takes place, the computer/human interface has perforce had to remain roughly the same size. But nevertheless, the nature of the interface has undergone a profound change. Formerly, the processing units were contained in a large room maintained at a fixed temperature and humidity. Job requests were typically handed to an operator, who would load them onto the computer and in due course collect a printout of the results, leaving them for collection by the requester. Miniaturization of the processing units has revolutionized computing in the sense that it has enabled the creation of the *personal* computer. The owner directly feeds instructions into it, and the results are displayed as soon as the computation has finished. (The largest parts of the personal computer are typically the keyboard with which instructions are given, and the screen on which results are displayed.) The miniature processor-enabled personal computer has made computing pervasive. In fact, it would be hard to overestimate the social effects of this pervasiveness. It is an excellent example of the qualitative results of miniaturization.

Another issue is accessibility, which is very size dependent. In an earlier epoch, children were much in demand as chimney sweeps, because they were small enough to clamber up domestic chimneys wielding a broom. The complexity of the circuits required for cellular telephony are such that a handheld device containing them only became possible with the development of miniaturized, very large-scale integrated circuitry. A similar consideration applies to the development of swallowable medical devices equipped with light sources, a camera, and perhaps even sensors and actuators for drug release.

The minute size of integrated circuit components also enables circuits of greater complexity to be devised and realized than would otherwise be possible. In addition, qualitatively different functions may emerge from differently sized devices. There are also secondary advantages of smallness, such as a requirement for smaller test facilities.

4.1 FABRICATION

Provided performance can be maintained, the smaller a device, the less material is used, leading to cost savings. It may also be easier to devise massively parallel fabrication procedures—indeed great use has been made of this possibility. Together, these innovations may enable single-use (disposable) devices to be introduced, with obvious advantages in applications such as medicine, avoiding the extra work of sterilization and the risks of cross-patient infection.

4.2 PERFORMANCE

Performance may be enhanced by reducing the size. If the reason for the size reduction is accessibility or ease of fabrication, the scaling of performance with size must be analyzed to ensure that performance specifications can still be achieved. It is worth noting that the performance of many microsystems (microelectromechanical systems, i.e. MEMS) devices actually degrades with miniaturization,[6] and the currently available sizes reflect a compromise between performance and other desired attributes.

If vast quantities of components can be made in parallel very cheaply, devices can be designed to incorporate a certain degree of redundancy, immunizing the system as a whole against malfunction of some of its components.[7] Low unit cost and low resources required for operation make redundancy feasible, hence high system reliability. A more advanced approach is to design the circuit such that it can itself detect and switch out faulty components. However, for malfunctions that depend on the presence of at least one defect (e.g., an impurity atom) in the material constituting the component, if the defects are spatially distributed at random, the smaller the component, the smaller the fraction that are defective.[8]

4.3 AGILE MANUFACTURING

The nearer nanotechnology approaches the ultimate goal of productive nanosystems (see Section 12.1), the more flexible manufacturing should be. One aspect that needs careful consideration is the fact that agile (adaptive) production systems are necessarily rooted in algorithms: hence, agile factories must necessarily be computer controlled, due to the large volume of information that has to be processed very rapidly (i.e., in real time) during production. The computer must (at least with present technology) run according to a preloaded program, which is necessarily closed and hence cannot but reflect present knowledge. The control center of the factory is therefore intrinsically ill-equipped to adapt to the ever-unfolding events that constitute the course of existence, which is largely constituted by the unknowable part of

[6]C. Hierold, From micro- to nanosystems: mechanical sensors go nano. *J. Micromech. Microengng* 14 (2004) S1–S11.

[7]See J. von Neumann, Probabilistic logics and the synthesis of reliable organisms from unreliable components. In: C.E. Shannon and J. McCarthy (eds), *Automata Studies*, pp. 43–98. Princeton: University Press (1956).

[8]This is an elementary application of the Poisson distribution. See A. Rényi, *Probability Theory*, pp. 122–125. Budapest: Akadémiai Kiadó (1970).

the future. The financial turbulence of 2008, which is now starting to have serious industrial consequences, is an all-too-obvious illustration of this truism. Different sectors seem to be intrinsically more or less sensitive to future uncertainty—rational planning demands that this sensitivity be quantified to determine the sectorial appropriateness of agile manufacturing. We need to anticipate that real-world contexts will raise challenges of increasing uncertainty and diversity and so require agility to be achieved by means that are suitably resilient and adaptable to change ("agile agility" or "adaptable adaptability"); that is, agility needs to be explicitly designed for the unknown and unexpected, not merely to cope with well-understood tempos and boundaries. Naturally this will have a cost: presumably the more adaptable an industrial installation, the more expensive, hence it will be appropriate to determine a suitable limit to the necessary adaptability for a particular sector.

FURTHER READING

P.M. Allen, Complexity and identity: the evolution of collective self. In: J.J. Ramsden, S. Aida and A. Kakabadse (eds), *Spiritual Motivation: New Thinking for Business and Management*, pp. 50–73. Basingstoke: Palgrave Macmillan (2007).

CHAPTER 5

The Nanotechnology Business

CHAPTER CONTENTS

This chapter addresses the questions: What nanotechnology is already commercialized? How big is the actual market? How big is the potential market? Coverage of the actual technologies takes place in the three remaining chapters of this part. Here, the main purpose is to put the whole in perspective.

Figure 5.1 summarizes the current situation. Note that the higher upstream the nanotechnology, the more indirect the final product. The more indirect, the harder it is to introduce a radical technology, since much more needs to be overturned. Until now, nanotechnology has been most prominent as a substitutional indirect technology (e.g., the introduction of 65 nm lithography in computer chips), and as an incremental quasidirect technology (carriers for active ingredients in cosmetics).

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

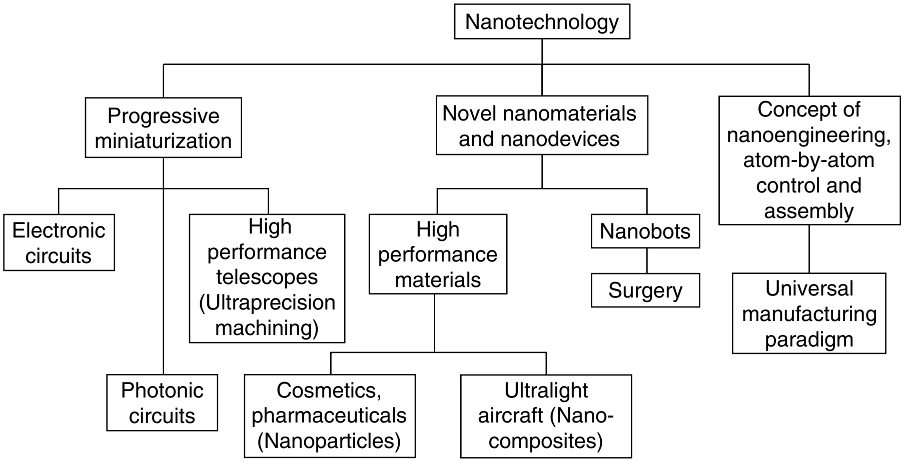


FIGURE 5.1

Indirect, direct and conceptual branches of nanotechnology (from left to right), with examples.

5.1 NANOTECHNOLOGY STATISTICS

A general caveat is in order here. There are a huge number of statistics about nanotechnology floating around the world. Websites, electronic newsletters and reports of commercial research are the main secondary sources. Hullman has compiled a summary of some of the secondary sources to create a tertiary report,[1] which well highlights the two main (related) problems: the huge variety of numerical estimates for most quantities ("indicators") and the difficulty of defining categories. The main reason for the huge discrepancies appears to be the wide variety of definitions of the indicators that are employed. The more easily accessible secondary sources (e.g., electronic newsletters) rarely, if ever, carefully define how they arrive at the quantities given. Reports that are supposed to be based on primary research might be more reliable, but this cannot be established without scrutinizing them in detail, and since they are rather expensive (typically costing several thousand US dollars) few people or organizations acquire a number of different ones and critically compare them. The best solution is probably to undertake the primary research oneself. The nanotechnology industry is still small enough to make this feasible at a cost still reasonable compared with that of acquiring the commercial reports, and with a considerable gain in reliability.

Adding to the confusion surrounding the so-called quantitative indicators is the fact that two of the most widely used terms in commercial predictions,

[1] A. Hullman, *The Economic Development of Nanotechnology—An Indicators-Based Analysis*. Brussels: European Commission, Directorate-General for Research, Nano Science and Technology Unit (2006).

Table 5.1 Definitions of Commonly Used Words for Large Numbers.

	Meaning in:		S.I. terminology[c]	
Word	**Europe**[a]	**USA**	**prefix**	**symbol**
Million	10^6	10^6	mega	M
Milliard	10^9	*b*	giga	G
Billion	10^{12}	10^9	tera	T
Trillion	10^{18}	10^{12}	exa	E

[a] *The same word, with the same meaning, is used in French and German.*
[b] *Rarely used.*
[c] *Corresponding to the European meanings.*

"billion" and "trillion", have parallel definitions differing respectively by three and six orders of magnitude from one another. Although the geographical origin of the number and its context usually allow one to reliably decide what is meant, it is regrettable that this ambiguity has been allowed to persist. Usage in the UK is currently the most confusing because while located in Europe, it shares the same language as the USA. The definitions are summarized in Table 5.1. To avoid confusion, in this book we shall mostly write the numbers out explicitly.

5.2 THE TOTAL MARKET

Figure 1 of the Hullman report (loc. cit.) shows the predicted evolution of the world nanotechnology market (presumably defined as sales). Predictions for the year 2010 range from about \$100 milliard to over \$1400 milliard—in other words, roughly the same as the entire present manufacturing turnover in the USA (\$$1.1 \times 10^{12}$ in 2007, one (US) trillion in round numbers; the Taylor report predicts a global nanotechnology market exceeding \$$2 \times 10^{12}$ around 2012).

A major ambiguity is whether the entire value of a consumer product containing upstream nanotechnology is counted towards the market value; often the nanotechnology only constitutes a small part of the total. Another major ambiguity (see Figure 2 of the Hullman report) is the possibility of double counting. Frequently, the nanotechnology market is divided into different sectors without a clear indication of the criteria for belonging to each division. For example, much of nanobiotechnology is concerned with medical devices, and the devices themselves may contain nanomaterials, yet these are all given separate categories; most aerospace applications involve materials, yet these are two separate categories.

Yet another problem is that it is rarely clear to what extent "old" nanotechnology is included. The world market forecast for nanotechnology given in Figure 1 of the Hullman report starts from zero in the year 2001 (but other data given in the same report suggests that in 1999 the world market was already of the order of $\$10^{12}$!). This report is, in fact, unusual insofar as nanotools or nanobiotechnology are given as the dominant sectors (e.g., Figure 3 of the Hullman report) whereas elsewhere it is generally accepted that the overwhelming part of the nanomarket is at present constituted from nanomaterials, with nanodevices and nanotools occupying an almost negligible part.[2] A more reasonable estimate is that the background level of nanomaterials valid at least up to about 2005 is a turnover of $\$5 \times 10^9$ per annum. It is then relatively minor—for comparison, the annual sales of Procter & Gamble in 2007 were $\$75 \times 10^9$. The nanomaterials market is dominated by nanoparticles, which includes (i) a very large volume of carbon black (2006 revenue was approximately $\$1.25 \times 10^9$), chiefly used as an additive for the rubber used to make the tires for motor vehicles; (ii) silver halides used in the photographic industry to sensitize the emulsions with which photographic film is coated; and (iii) titanium dioxide used as a white pigment in paint—in other words, "old" nanotechnology (which should therefore not be considered as nanotechnology at all—see Chapter 1).

Furthermore, one almost never sees any uncertainties associated with these estimates. They must often be of the same order of magnitude as the estimates themselves. For example, Figure 12 of the Hullman report shows the distribution of company sizes (measured by turnover) in different countries; but in the USA and the UK the overwhelming majority of companies do not reveal their sizes.

Another criticism is that in many cases per-capita comparison would be more relevant than an absolute one; for example, Figure 14 of the Hullman report compares the numbers of institutes active in nanotechnology for European countries—at first glance the graph simply seems to follow the populations of the countries. Even when normalized by population, though, the distribution of institute sizes might vary widely from country to country.

Given these deficiencies, we can only repeat what was already stated above—the best recommendation that we can give in this book is that if a company wishes to forecast the market in its particular niche, it had better

[2] The Hullman report merely compiles secondary sources, without any criticism or even highlighting discrepancies.

attempt it by itself—the results are likely to be more reliable than those taken from elsewhere, and the assumptions used in compiling the data can be clearly stated and hence will be transparent and accessible.

5.3 THE CURRENT SITUATION

To recapitulate, total global demand for nanoscale materials, tools and devices was estimated at $5–8 milliard in 2003 and forecast to grow at an average annual growth rate (AAGR) of around 30%, implying that it would have reached almost $30 milliard in 2008. Whether this figure was actually reached is not yet known. Hence, comparing nanotechnology to other key emerging technologies, the global nanotechnology market is roughly comparable in size to the biotechnology sector, but far smaller than the $800 milliard global informatics market. However, the nanotechnology market is believed to be growing more than twice as fast as either of the other two.

As stressed in the preceding section, figures of this nature are necessarily somewhat approximate, not least because of the lack of uniformity regarding the criteria for inclusion; that is, the answer to the question, what is nanotechnology? The global personal computer market is presently worth about $20 milliard and as Moore's law continues its march, more and more of this market could reasonably be included in the nanotechnology sector—we do not know precisely what proportion of it is at present.

As already stated, nanoscale metal oxide nanoparticles are already very widely used. Typical current applications include sunscreens (titanium oxide and zinc oxide), abrasion-resistant coatings, barrier coatings (especially coatings resistant to gas diffusion) and antimicrobial coatings. Applications of fullerenes and carbon nanotubes, which are "true" nanotechnology products, still constitute essentially niche markets. However, the fastest-growing nanomaterials segments are nanotubes (with an amazing projected AAGR of 170–180% over the next five years) and nanocomposites (about 75% AAGR).

The nanomaterials segment, which, as already mentioned, includes several long-established markets ("old" nanotechnology) such as carbon black (used as filler for rubber), catalytic converter materials, and silver nanoparticles used in photographic films and papers, until very recently accounted for almost all (i.e., in excess of 95%) of global nanotechnology sales. By 2008, however, it is anticipated that the nanomaterials share of the market will have shrunk to around 75% of total sales. Nanotools are now estimated to have increased their share to around 5% (about $1 milliard), and nanodevices are supposed to have established a major presence in the market, with a 20–25% share (equivalent to $6 milliard annually).

The projections naturally depend on a great many imponderables, including the general level of economic activity and economic growth. One highly debatable matter is the influence of government spending. Opinions range from unfavorable (i.e., government spending hinders rather than facilitates technical progress), to favorable (i.e., it makes an indispensable contribution to national competitiveness). We return to this theme in Chapter 9.

In terms of tonnage, total global consumption of all types of nanomaterials was estimated at having surpassed 9 million metric tons in 2005 and is predicted to reach 10 million tons by 2010, at an AAGR of almost 10%.[3] Nonpolymer organic materials account for the largest share of total nanomaterials consumption, the bulk of which are carbon black fillers, which is of course a relatively simple traditional material. The share of simple oxide nanomaterials is expected to double from currently about 8% to 16% in 2010. Metal nanomaterials are the second-largest segment, with more than 20% of the market at present.[4]

Regarding product morphology, by 2010 nanoparticulates' share of the market is projected to shrink somewhat to just over 50%, while thin films, monolithics and composites are expected to grow to 25%, 20% and 3% respectively.

Currently hundreds of kinds of nanomaterials are in use or under development, both in their pure form and as composites. Examples include carbon in novel forms, tungsten, titanium and cobalt, as well as many technical ceramics such as new forms of aluminum oxide, silicon carbide and their composites. Many of these are candidates for adding to paper-based products.

The range of applications for nanomaterials is growing rapidly. Whereas until now nanomaterials have tended to be associated with niche segments such as bouncier tennis balls, a new trend of serious large-scale applications is emerging. These applications currently include tires and other rubber products, pigments, synthetic bone and automotive components. Tomorrow's applications include automotive coatings, medical devices and filtration media, to name just a few.

To repeat, one of the difficulties in gathering statistics and appraising the collections that are published almost every month is that it is often not clear exactly what is included as nanotechnology and what is not. For example, carbon black is a traditional material that does indeed have some

[3] The basis of these predictions, gleaned from a wide variety of sources, is almost nonexistent, amounting in most cases to a crude extrapolation of the trends of the past few years.

[4] The numbers given in this and the following paragraphs represent a considered consensus among a great variety of publicly accessible sources, too numerous to list individually.

nano attributes, but it does not belong conceptually to nanotechnology, strictly speaking, because it is not novel. Sometimes what is included within nanotechnology is in effect merely a relabeled traditional product.[5]

Very often these near-nano products enhance the performance of the materials to which they are added to close to the theoretical limit, in which case the almost inevitably higher expense associated with substituting them by real nanomaterials would not result in any significantly increased added value, and hence there is no driver to make the substitution.

5.4 CONSUMER PRODUCTS

The Woodrow Wilson Center has made a list of consumer products containing nanotechnology that continues to be curated.[6] Some of these are given in the next three tables. This set of data provides a useful snapshot of the current commercial market.

Regarding Table 5.2, it is not always clear what the exact criteria for inclusion are, especially for products that could fit in multiple categories. For example, would a household appliance be included under "Appliances" or under "Home and garden"? Most appliances include some electronics, I would imagine. And does an automobile (which may contain some onboard information processors with nanoscale features in their chips) count as a single product? Spray paint containing nanoparticles for use by owners to repair minor scratches presumably ranks as an automotive product, but does each available color count as a separate product? Furthermore, the compilers of the data have not themselves verified whether the manufacturers' claims are correct. Moreover, one has no indication of the volumes sold. Cellphones probably outrank all the other members of its category.

It is perhaps surprising that there are already so many food products at least containing, if not based on, nanotechnology—these, incidentally, might well have been included in the "Health and fitness" category. The list is anyway dominated by health and fitness products, which are further subdivided in Table 5.3. Presumably medical products not available to uncontrolled consumers (e.g., prescription drug delivery nanomaterials) are not included in the list at all—or else none are currently available.

[5] E.g., J. Harris and D. Ure, Exploring whether 'nano-' is always necessary. *Nanotechnol. Perceptions* 2 (2006) 173–187.

[6] *Project on Emerging Nanotechnologies: Consumer Products Inventory*. Washington (DC): Woodrow Wilson International Center for Scholars (project began in March 2006).

Table 5.2 Numbers of Consumer Products in Different Categories (Status in January 2009).[a]

Category	Number	%
Health and fitness	502	58
Home and garden	91	10
Food	80	9
Electronics	56	6
Automotive	43	5
Appliances	31	4
Other	70	8
Total	873	100

[a] *Source: see footnote 6.*

Table 5.3 Numbers of Consumer Products in the "Health and Fitness" Category (Status in January 2009).[a]

Subcategory	Number	%
Personal care	153	28
Cosmetics	126	23
Clothing	115	21
Sporting goods	82	15
Filtration	40	7
Sunscreen	33	6
Total	549	100

[a] *Source: see footnote 6.*

Finally, it is interesting to look at which elements dominate nanotechnology applications (Table 5.4). Presumably these are mostly in the form of nanoparticles. Carbon presumably means fullerenes. Silicon, titanium and zinc are presumably nanoparticles of their oxides. Since the database is of consumer products, presumably silicon-based integrated circuits are not included.

Table 5.4 Numbers of Consumer Products Categorized According to the Elements Declared as Constituting the Nanocomponent (Status in January 2009).[a]

Element	Number of Products	%
Silver	235	56
Carbon	71	17
Titanium	38	9
Silicon	31	7
Zinc	29	7
Gold	16	4
Total	420	100

[a] *Source: see* *footnote 6*.

The consumer market is of course extremely fickle. The epithet "nano" is often used as a marketing ploy, even if the product contains no nanomaterials at all.[7]

Furthermore, it is evolving with amazing rapidity. A camera in 1960 contained no electronics, but now contains probably 80% or more, much of which is heading towards the nanoscale. A similar trend has occurred regarding personal calculators, the functional equivalent of which would have been a slide rule or a mechanical device in 1960. The personal computer did not even exist then. A motor-car typically contained about 10% (in value) of electronics in 1960; this figure is now between 30% and 50% and much of it is already, or fast becoming, nano.

A crucial point regarding consumer market volume is the renewal cycle. Whereas in other markets technical considerations dominate—for example, in many European cities the underground railway trains and trams might be of the order of 50 years old and still in good working order—psychosocial factors dominate the decision whether to replace a consumer product. It seems remarkable that those who have a mobile phone (i.e., the majority of the population) typically acquire a new one every 6 months (many are anyway lost or stolen). Other consumer electronics items such as a personal computer, video recorder or television set might be renewed every 1–2 years. Even

[7] See, e.g., D.M. Berube, The magic of nano. *Nanotechnol. Perceptions* 2 (2006) 249–255.

a motor-car is likely to be changed at least every 5 years, despite the many technological advances that ensure that it is still in perfect working order at that age.

Here, deeper issues are raised. Without the frenetic pace of renewal, the hugely expensive infrastructure (e.g., semiconductor processing plants) supporting present technology could not be sustained, and though rapid "planned obsolescence" seems wasteful, without it innovation might grind to a halt, with possibly deleterious consequences for mankind's general ability to meet future challenges (including those associated with global warming).

5.5 THE SAFETY OF NANOPRODUCTS

One issue that has not so far received much prominence is that of safety, especially regarding nanoparticles in products brought into contact with the skin, if not actually ingested. Compared with the furore over genetically modified food crops, leading to widespread prohibition of their cultivation, at least in Europe, nanoparticle-containing products have generally had a favorable reception, perhaps because of the considerable care taken by the industry to inform members of the public about the technological developments that led to them.

Nevertheless, there is no doubt that nanoparticles have significant biological effects. An extensive literature already exists.[8] A member of the public might wish to take the following widely known facts into account:

1. Workers, especially miners, exposed to fine particles suffer occupational diseases such as silicosis and asbestosis. Tumors typically first appear after many years of exposure, and are painful, incurable and fatal.
2. Widespread use of coal for domestic heating (e.g., in London up to the 1950s and in Germany up to the 1990s) led to severe atmospheric pollution and widespread respiratory complaints.
3. On the other hand, restricted exposure to dusts (speleotherapy, e.g., as practiced in the "Rehabilitation" Scientific-Medical Center of the Ukrainian Health Ministry in Uzhgorod (Ungvár)) is considered to be therapeutic; from 1864 to 1905 (when electric traction was introduced) people suffering from respiratory complaints were

[8]See, e.g., P.A. Revell, The biological effects of nanoparticles. *Nanotechnol. Perceptions* 2 (2006) 283–298 and references therein.

encouraged to travel on the Metropolitan and District Railways in London, in the days when their trains were still steam-hauled, and hence the tunnels through which they passed were rich in sulfur fumes.

4. Cigarette smoking in public places is now subject to draconian restrictions (in Europe and the USA)—even though the original epidemiological studies of Richard Doll purporting to link smoking with disease have been shown to be flawed.

5. The increase of motor traffic in major cities, coupled with official encouragement of diesel engines, which emit large quantities of nanoparticulate carbon in their exhaust, have made air pollution as bad nowadays as it was in the days when domestic heating using coal was widespread (in cities such as Athens, where there was very little heating anyway, the pollution nowadays is incomparably worse than anything previously experienced in history).

This list could be prolonged, but the point is made that no coherent policy can be discerned at present; the situation is full of paradoxes. Items 1 and 2 can doubtless be resolved by recalling Paracelsus's dictum "The poison is in the dose", but in other cases doubtless economic and political factors took precedence over scientific and medical ones. The British government now seems to be resolved to bring some order into this chaos, and has commissioned a report prescribing how studies to determine the biological hazards of nanoparticles ought to be carried out.[9] The dispassionate observer of the field will find it remarkable that despite decades of investigations, most reported studies have failed to carry out requisite controls, or are deficient in other regards. A very great difficulty of the field is the extremely long incubation time (decades) of some of the diseases associated with exposure to particles. The effects of long-term chronic exposure might be particularly difficult to establish. At the same time, ever since Prometheus man has been exposed to smoke, an almost inevitable accompaniment to fire, and doubtless the immune system has developed the ability to cope with many kinds of particles.[10]

[9]C.L. Tran et al., *A Scoping Study to Identify Hazard Data Needs for Addressing the Risks Presented by Nanoparticles and Nanotubes*. London: Institute of Occupational Medicine (2005).

[10]This is very clearly not the case with highly elongated particles such as blue asbestos fibers, however; parallels with long carbon nanotubes are already exciting concern.

The most appropriate response is indeed to make good the deficiencies of previous work, as recommended by Tran et al. (loc. cit.).[11] The question remains, what are we to do meanwhile, since it might be many years before reasonably definitive answers are available. Most suppliers of nanomaterials would, naturally enough, prefer the status quo to continue until there is clear evidence for acting otherwise; pressure groups are active in promulgating the opposite extreme, advocating application of the precautionary (or "White King") principle (do nothing unless it is demonstrably safe) and an innovation-stifling regulatory régime. The latter is anyway supported by governments (probably, even doing the research required to establish the safety or otherwise of nanoparticles contravenes existing health and safety legislation) and supergovernmental organizations such as the European Commission.

Hence, the most sensible course that can be taken by the individual consumer is to apply the time-honored principle of *caveat emptor*. But, the consumer will say, we are not experts, how can we judge? But, the expert may respond, we all live in a technologically advanced society, and we all have a corresponding responsibility to acquaint ourselves with the common fund of knowledge about our world in order to ensure a long and healthy life. Naturally we have a right to demand that this knowledge is available in accessible and intelligible form.

5.6 GEOGRAPHICAL DISTRIBUTION

How is nanotechnology activity spread around the world? According to the Civilization Index (CI),[12] the countries of the world can be grouped into four categories:

I. High per-capita income, high level of scientific activity (e.g., Canada, Germany, Japan, Switzerland, UK, USA).

[11] It is perhaps a little surprising, given the weighty expertise that went into this report, that the outcome—the recommendations—are almost trivial. For example, it is considered that the top priority is "the formation of a panel of well-characterized, standardized nanoparticles for comparison of data between different projects and laboratories", and "the development of short-term *in vitro* tests aimed at allowing toxicity to be predicted from the physicochemical characteristics of the particles" is recommended. Why, one may ask, was this not done before? Many scientists have worked on these problems already; it might have been more appropriate to examine why such a poor standard of experimentation has been accepted with so little criticism for so long.

[12] *A New Index for Assessing Development Potential*. Basel: Collegium Basilea (2008).

II. Low per-capita income, high level of scientific activity (e.g., Argentina, China, Georgia, Hungary, India, Russia).

III. High per-capita income, low level of scientific activity (e.g., Brunei, Kuwait, Libya, Saudi Arabia).

IV. Low per-capita income, low level of scientific activity (e.g., Angola, Indonesia, Thailand, Zambia).

Category I comprises the wealthiest countries of the world. They are active in nanotechnology, have a high level of scientific research and technical development in most areas, and have a good level of higher education. We would expect that countries in this category are leading in at least one branch, both scientifically and in developing innovative products.

Category II mostly comprises countries of the former Soviet Union, which had highly developed scientific research activities for much of the 20th century but since 1991 have fallen on hard times, together with countries that historically had strong traditions of technical innovation (for example, until around the 17th century China was well ahead of Europe) but failed to sustain past momentum (for reasons that are not understood) and, perhaps more significantly, failed to develop a strong science to parallel their technology; this subgroup within the category also has a large rural, barely educated population.

Category III comprises countries with arguably a lower level of civilization that have acquired vast riches in recent decades through the export of raw materials found in their territories, especially oil. They have manifested little interest in supporting the global scientific community; what technology they have is mostly imported.

Category IV comprises countries with a lower level of civilization that might require centuries of development before reaching the attainments of Category I (see also Section 5.6.2).

5.6.1 The fiscal environment for nanotechnology

The three major poles of economic activity (the EU, Japan and the USA) are quite sharply distinguished regarding expenditure on nanotechnology (research and technical development):

- Category I (Japan): two-thirds private, one-third public
- Category II (USA): one-half private, one-half public
- Category III (EU): one-third private, two-thirds public.

Although the total expenditure in each of these three poles is roughly the same (around 4×10^9 CHF; again, the validity of this statement depends on what is included under "nanotechnology"), its effectiveness differs sharply. There can be no doubt that the Japanese model is the most successful. Solidly successful companies (without any magic immunity from the vagaries of the market) with immense internal resources of expertise have impressive track records in sustainable innovation according to the alternative model (Figure 2.2), but are well placed to develop nanotechnology according to the new model (Figure 2.3). Category II has several successful features, not least the highly effective Small Business Innovative Research (SBIR) grant scheme for funding innovative starting companies, and benefits from enormous military expenditure on research, much of which is channeled into universities. Category III is decidedly weak. There is an overall problem in that the fraction of GDP devoted to research and development in the EU is less than half that found in Japan or the USA. Moreover, what is spent is not well used. Many companies have been running down their own formerly impressive research facilities for decades (the clearest evidence for this is the paucity of top-ranking scientific papers nowadays emerging from European companies). Government policy has tended to encourage these companies to collaborate with universities, enabling them to reduce the level of public funding. Within Europe, there are immense differences between countries, however. Among the leading countries (Britain, France, Germany) France is in the weakest position. Traditionally anyway weak in the applied sciences, without a strong tradition of university research, and with its admirable network of state research institutes (the CNRS) in the process of being dismantled, there is little ground for optimism. In the UK, the level of innovation had become so poor that the government has virtually forced the universities to become commercial organizations, patenting inventions and hawking licenses to companies, and insisting on commercial outcomes from projects funded by the state research councils. Although university research is ostensibly much cheaper than company research, most companies seem to have unrealistic expectations of how much they can expect to get from a given expenditure, in which they are anyway typically ungenerous to a fault. The British government avows the linear model (Figure 2.1), and is fond of emphasizing the importance of the research "base" as the foundation on which industrial innovation rests, but paradoxically is extremely mean about paying for that base, whose funds are cut at the slightest excuse, hence we cannot be optimistic about the future. In Germany, there is a strong *Mittelstand* of medium-sized engineering firms with many of the characteristics of Japanese companies. Furthermore, the state Fraunhofer institutes of applied sciences, along with the Max-Planck institutes (the equivalent of the French CNRS) are flourishing centers of real competence. If the EU were only Germany, we would be optimistic. The European Commission (the central administrative service of the European Union) seems to be aware of the problems, and has initiated a large supranational program of research and technical development, but the outcome is remarkably meager in comparison with the money and effort put into it. The main instrument is the "Framework" research and technical development program, but this is rather bureaucratic, easily influenced by dubious lobbying practices, and hence generally unpopular. The bureaucracy is manifested by excessive controls and reporting requirements, brought in as a result of the generally deplorably high level of fraud in the overall EU budget (including agrigulture, regional funds, etc.), which dwarfs the scientific activity *per se*, but all expenditure is subject to the same rules. There is little wonder that it has been concluded that the "Framework" program actually hinders innovation in European industry.[13]

What should be done seems clear; whether it will be done is another matter. It should, however, be emphasized that it will be ineffective to merely increase spending on scientific research without rethinking some of the premises according to which it is carried out; in particular, a much more critical approach to its prosecution and the outcomes is needed.

5.6.2 Nanotechnology in the developing world

Nanotechnology, it has been proposed, is very attractive for poor, technology-poor countries to embrace because it seems to require less investment before yielding returns. Furthermore, nanotechnology offers more appropriate solutions to current needs than some of the sophisticated Western technologies available for import. Water purification using sunlight-irradiated titanium dioxide nanoparticles would be a characteristic example. Are these propositions reasonable?

Alas, the answer probably has to be "no". One of the greatest handicaps countries in Category IV face is appallingly ineffectual government; precisely where direction would be needed to focus local talent there is none, and most of these governments are mired in seemingly ineradicable venality. The situation nowadays in many African countries is apparently considerably worse than half a century ago when, freshly independent, they were ruled with enthusiasm and a great desire to develop a worthy autonomy. Zimbabwe offers a very sobering illustration: the country had a good legacy of physical and educational infrastructure from the colonial era, but today, after the

[13] House of Lords Select Committee on the European Communities, Session 1993–94, 12th Report, Financial Control and Fraud in the Community (HL paper 75). London: HMSO, 1994.

government has bent over backwards to distribute land to the landless, they have shown themselves incapable of stewardship and agricultural output has plummeted.

The doubtless well-meaning efforts that have resulted in the foundation of institutions such as the new Library of Alexandria and the Academy of the Third World also seem doomed to failure, for they are rooted in an uncritical admiration of sterile "pure" sciences in the Western tradition which, while superficially glamorous in a narrow academic sense, are incapable of taking root and growing—nor would such growth be useful to their environment.

Having said that, if a country wished to focus all its resources on one area, nanotechnology would probably be the best choice, because its interdisciplinary nature would ensure that the knowledge base had to be broad, while the immediacy of applications would ensure rapid returns. The criterion of success will be for a country to achieve leadership in some subfield of the technology: this will show it has crossed the sustainability threshold, which is unlikely to be achievable by merely imitating leading Western scholarship.

It would be greatly encouraging if any country launching a focused nanotechnology effort would avoid the pitfalls of "standard empiricism" (see Chapter 13) and make a fresh start with aim-oriented science, and from the beginning encourage healthy, open criticism. At the same time, good use should be made of the global (for such may it be considered) scientific legacy, by sending scholars to a variety of foreign centers of excellence to learn. It would be futile to await handouts from international funds (IMF, World Bank and the like) for such a purpose—they are not interested in promoting independent science. In most countries, the leaders could well afford to fund appropriate scholarships for undertaking doctoral degrees (for example) abroad. Why should not Bokassa scholarships become as important and valuable as Rhodes scholarships have been in the past? These returning scholars would represent seeds of immense growth potential.

Despite their problems, these countries have two great advantages compared with the developed world. One is that they have practically nothing to dismantle first, which represents such a big obstacle to the introduction of new ways of thinking.[14] The other is that their natural resources are relatively unexplored and unexploited; looking at them from the bottom up is almost certain to yield new knowledge, leading to new avenues for wealth creation.

[14]Cf. the introduction of mobile telephony: penetration exceeds that in the developed world, because the fixed line infrastructure is so poor.

CHAPTER 6

Miscellaneous Applications

CHAPTER CONTENTS

The remaining chapters of Part 2 survey commercial and near-commercial applications. Because of their importance, information technology and health applications are placed in separate chapters; all remaining applications are

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

covered here. The order of coverage is firstly upstream technologies, notably materials, carbon-based materials and ultraprecision engineering, followed by downstream technologies in alphabetical order.

6.1 NONCARBON MATERIALS

The main raw materials manufactured on a large scale are nanoparticles. As already pointed out in Chapter 5, the bulk of these are traditional particles, notably carbon black, silver halide emulsion crystals, and pigments, which are in no sense engineered with atomic precision; some of them merely happen to fall within the accepted range of nano-objects.[1] These products are, in fact, typically quite polydisperse. Attempts to synthesize monodisperse populations on a rational basis have a considerable history.[2] Natural materials such as clays also provide a significant source of nanoparticles. Key parameters of nanoparticles are size (and size distribution), chemical composition, and shape (including porosity).

6.1.1 Composites

Nanoparticles have relatively few direct uses; mostly their applications are in composites (i.e., a mixture of component A added to a matrix of component B, the latter usually being the majority component)—a nanocomposite differs from a conventional composite only insofar as the additive is nanosized and better dispersed in the matrix. The purpose of adding materials to a polymer matrix is to enhance properties such as stiffness, heat resistance, fire resistance, electrical conductivity, gas permeability, and so forth; the object of any composite is to achieve an advantageous combination of properties. If the matrix is a metal, then we have a metal-matrix composite (MMC). A landmark was Toyota's demonstration that the incorporation of a few

[1] This remark does not do justice to the extraordinary sophistication of a fabricated photographic emulsion crystal, achieved over more than a century of intensive research. For a historical overview, see R.J. Hercock and G.A. Jones, *Silver by the Ton*. London: McGraw-Hill (1977).

[2] Uniform grain size was an important goal of emulsion manufacturers (see footnote 1). For other semiconductors see, e.g., J.J. Ramsden, The nucleation and growth of small CdS aggregates by chemical reaction, *Surf. Sci.* 156 (1985) 1027–1039; and T. Graham, On liquid diffusion applied to analysis, *J. Chem. Soc.* 15 (1862) 216–269 for an example of much earlier work that achieved nanoparticle monodispersity.

weight percent of a nanosized clay into a polyamide matrix greatly improved the thermal, mechanical and gas permeability (barrier) properties of the polymer.[3]

There is no general theory suggesting that the advantage scales inversely with additive size; whether a nanocomposite is commercially viable depends on all the parameters involved. There is such a huge variety of materials that it is perhaps futile to attempt a generalization. However, the very small size of individual nanoparticles would make it feasible to incorporate a greater variety of materials within the matrix for a given additive weight percent. Very often ensuring wetting of the particle by the matrix presents a significant technological hurdle. Most successful composites require the additive to be completely wetted by the matrix. Wetting behavior can be predicted using the Young–Dupré approach;[4] if, however, the particle becomes very small, the surface tension will exhibit a curvature-dependent deviation from the bulk value appropriate for a planar particle-matrix interface.

The chief manufacturing routes for polymeric nanocomposites are blending preformed nanoparticles with the molten matrix; dispersing preformed nanoparticles in the monomer precursor to the matrix and polymerizing it; and synthesizing the nanoparticles *in situ* within the matrix.

6.1.2 Coatings

Many properties of materials essentially only concern their surfaces; if so, it is much more cost-effective to apply a thin film to the bulk material in order to achieve the desirable interfacial attribute (e.g., a low coefficient of friction). Furthermore, desirable bulk properties are not compromised. Alternatively, the surface can be modified using a technique such as ion implantation.[5]

---

[3] A much older composite is paint, which consists of a pigment (quite possibly made of nanoparticles) dispersed in a matrix of varnish. Paint can be said to combine the opacity of the pigment with the film-forming capability of the varnish. Another mineral-polymer composite is the material from which many natural seashells are constructed: platelets of aragonite dispersed in a protein matrix. In this case, however, the “matrix” only constitutes a few percent of the volume of the composite.

[4] For an introduction, see M.G. Cacace et al., The Hofmeister series: salt and solvent effects on interfacial phenomena. *Q. Rev. Biophys.* 30 (1997) 241–278.

[5] See Fraunhofer Gesellschaft, *Produktionstechnik zur Erzeugung funktionaler Oberflächen. Status und Perpecktiven.* Braunschweig (2008); J.J. Ramsden et al., The design and manufacture of biomedical surfaces. *Annals CIRP* 56/2 (2007) 687–711.

6.2 CARBON-BASED MATERIALS

Fullerenes and carbon nanotubes are true children of the Nano Revolution: they did not exist as commercial commodities before. Although carbon black and diamond-like carbon thin films have some nanofeatures, they are not atomically engineered and, moreover, existed before the era of nanotechnology; we do not propose to recruit them retrospectively.

Industrial problems associated with large-scale fullerene manufacture have been solved;[6] to date, their applications remain niche, however. Far more interest is associated with carbon nanotubes.[7] Some of their extraordinary properties include: a very high aspect ratio (their diameter can be less than 1 nm, but they can be many micrometers long), making them suitable for field emission applications and as conducting additives to polymer matrices with a very low percolation threshold; very high electron mobility and the highest current density of any known material, approximately $10^9\,\mathrm{A\,cm^{-2}}$ (cf. copper with $10^6\,\mathrm{A\,cm^{-2}}$, and aluminum 10 times less); ballistic electron transport; the highest Young's modulus of any known material (approximately 1 TPa along the axis); the highest thermal conductivity of any known material (approximately $4000\,\mathrm{W\,m^{-1}\,K^{-1}}$). Manufacturing problems seem to be on the way to being solved (note that some key applications use only very small quantities); the main issue today is probably purity. Nanotubes grown from a hydrocarbon feedstock such as acetylene using chemical vapor deposition require a metal catalyst (usually iron or nickel), which can be troublesome to remove afterwards; preparations are frequently contaminated with amorphous carbon.

The good field emission characteristics (including ultrahigh brightness and small energy dispersion) make carbon nanotubes (CNTs) outstandingly good electron sources for scanning electron microscopy—although the world market is very small. They are also useful in residual applications of vacuum tubes (e.g., high-power microwave amplifiers). There is obviously an enormous potential application as field emission displays, although parallel innovations are required here, including a convenient means of positioning them. Furthermore, there is intense competition from organic light-emitting devices. The electronic properties (very high current densities) make CNTs attractive for the vertical wires (VIAs) connecting layers of integrated circuits,

[6] M. Arikawa, Fullerenes—an attractive nano carbon material and its production technology. *Nanotechnol. Perceptions* 2 (2006) 114–121.

[7] B.O. Boscovic, Carbon nanotubes and nanofibres. *Nanotechnol. Perceptions* 3 (2007) 141–158.

although exactly how their fabrication will be integrated into existing semiconductor processing technology has not been sorted out. They may also be used as gates in field effect transistors (note that they can be prepared as semiconductors or metals).

The very low percolation threshold of these extremely elongated objects (a few volume percent) enables the preparation of electrically conducting polymers with such a low volume fraction of CNTs that the composite is visually unaffected by their presence. The main product of Hyperion (Section 9.7) is conducting paint suitable for the mass-production lines of the automotive industry. Other applications include antistatic coatings and electromagnetic screening films. Of great interest is the possibility of preparing transparent conducting films that could be used as the counterelectrode in displays. At present, indium-doped tin oxide (ITO) is the main material used for this purpose, but not only is it too brittle to be usable in flexible displays, but the world supply of indium is expected to be exhausted within 2–3 years at present rates of consumption!

CNTs can also be used as the charge storage material in supercapacitors. All the atoms of single-wall nanotubes are on the surface, hence they have the highest possible specific surface area ($1.5 \times 10^3\ m^2\ g^{-1}$), suggesting a theoretical upper limit for energy density of $20\ W\ h\ kg^{-1}$. Supercapacitors rated at 1000 farads are commercially available. Nevertheless, carbon black, which is much cheaper, is almost as good, hence it is doubtful whether this application will be commercially viable.

The very small size of a single nanotube makes it an attractive electrode material in electrochemical applications for which microelectrodes have already been shown to diminish transport-related overpotentials.

As far as composites are concerned, despite the extraordinarily high Young's modulus, mechanical performance has not been demonstrated to be superior to that already attainable with carbon fibers. The problem is how to disperse them in the matrix.

The variable valence and the availability of electrons makes CNTs attractive potential catalysts for certain reactions, for example in the petrochemical industry.

6.3 ULTRAPRECISION ENGINEERING

The market for ultraprecision machine tools is relatively small, amounting to a few tens of millions of dollars annually. The USA has just two companies in this business, each one selling a few dozen machines a year; the machines themselves cost hundreds of thousands to a million dollars apiece.

6.4 AEROSPACE AND AUTOMOTIVE INDUSTRIES

The dominant goal is to reduce vehicle weight without compromising the chemical and other attributes. For spacecraft, launch is one of the highest cost factors and is directly related to mass, but aircraft and even road vehicles benefit from reduced weight—less fuel is required to accelerate them. Hence there is much activity in seeking to replace the heavy metals used in components by lightweight polymers strengthened by nanoparticulate or nanofibrous additives (see Section 6.1.1). Other more specific aims include formulating lightweight electrically conducting materials for use in fuel lines to avoid the build up of static electricity, ultrahard (abrasion-resistant) paint, low-friction finishes and so forth. A significant difference between aerospace and automotive is that the lead time for the introduction of an innovation is typically 10 times longer in the former than in the latter (in which it is about 3 years), due to the more stringent needs for testing. Since sports equipment has many similar requirements, but is not usually safety-critical, it offers an interesting path for materials development to manufacturers in these sectors.

6.5 CATALYSIS

It has long been recognized that the specific activity of heterogeneous catalysts increases with increasing state of division. This is of course an old market that has long been a very significant part of the chemical industry. The world market amounts to about 15 milliard USD. However, even though many catalysts use nanosize metal clusters (for example), they cannot be called examples of atomically precise engineering. Indeed, the whole field is remarkable for the high degree of empirical knowledge prevailing. In the future, nanotechnology offers the chance to assemble catalysts atom by atom. There is a general feeling in industry that there is still considerable potential for increasing the activity of catalysts (that is, through both more effective acceleration of the desired reaction and more effective suppression of undesired side reactions).

About a quarter of the world market is accounted for by oil refining, and over half is nowadays accounted for by automotive exhaust catalysts.

6.6 CONSTRUCTION

The main application for nanotechnology in this sector is currently in materials, especially concrete enhanced using nanoparticles. Even though superior properties can be demonstrated, however, market penetration of nano innovations can be expected to be slow, because of the traditional low-tech attitudes prevailing in much of the industry.

The current penchant of architects for designing large buildings predominantly covered in glass has provided a welcome counter-tendency, however, because glass offers many possibilities for nanotechnological enhancement. In particular, nanostructured superhydrophobic surfaces imitating those of the leaves of plants such as the lotus enable raindrops to scavenge dirt and keep the surfaces clean. Nanoparticles of wide band-gap semiconductors such as titanium dioxide can be incorporated into the surface of the glass, where they absorb ultraviolet light, generating highly oxidizing (or reducing) species able to decompose pollutants adsorbed from the atmosphere. Ultrathin film coatings, even of metals, can be applied to the surface of glass, in order to control light and heat transmittance and reflectance. Sophisticated glasses with electrically switchable transmittance are now available. "Anti-graffiti" paint, from which other paint sprayed on can be easily removed, has also gained a certain popularity (although a social, rather than a technological, solution might be more effective at eliminating unwanted graffiti).

Ultimately, the availability of dramatically new nano-engineered materials (e.g., ultrastrong and ultralight diamondoid panels) may usher in a totally new era of architecture.

6.7 ENERGY

A field as diverse as energy is potentially affected in many ways by nanotechnology. For example, improved catalysts enhancing yields of petroleum-based fuel affect energy supplies. We can only hint at a few possible directions in this section.

As far as production and storage are concerned, the Holy Graal is mimicry of natural photosynthesis.

6.7.1 Production

Areas where nanotechnology might make significant impact are photovoltaic cells and fuel cells. To repeat, regarding the former, the Holy Graal is natural photosynthesis, which achieves the necessary photoinduced charge separation by extraordinarily sophisticated structure at the nanoscale. Regarding the latter, a major difficulty is the complex set of conflicting attributes that the materials constituting the fuel cell, especially the most important solid oxide type, must fulfill. Since nanocomposite materials are able to combine diverse attributes more effectively than conventional materials, there is some hope that more robust designs may emerge through a more systematic application of rational design.

6.7.2 Storage

High energy and power densities have become particularly relevant as a consequence of the search for replacements for the fossil fuel-powered internal combustion engine. The main contenders are supercapacitors and accumulators. The proliferation of portable electronic devices has also greatly increased demand for small and lightweight power sources for goods such as laptop computers and cellphones.

6.7.3 Lighting

Lighting appears here because it is an indispensable part of civilization, and is so widely used that improvements (e.g., more light output for the same input of electrical energy) have the potential to make a significant impact on energy consumption. We need only mention light-emitting diodes (LEDs) here; organic LEDs (OLEDs) are usually included in the domain of nanotechnology.

6.8 ENVIRONMENT

Environment is an even more amorphous concept than energy in a commercial context. Here, we only consider the remediation of contaminated soils and groundwater by the addition of nanoparticles.[8] If a source of ultraviolet light is available (sunlight is adequate), titanium dioxide is a useful material; absorption of light creates electron-hole pairs acting as strong reducing-oxidizing agents for a large variety of organic compounds adsorbed on the nanoparticle surface. By this means many recalcitrant potential pollutants can be destroyed.[9] Until now attention has been mainly concentrated on the actual science of the photoassisted chemical decomposition rather than devising a complete process in which the nanoparticles, having done their work, would be collected and possibly regenerated for further use.

---

[8] D. Rickerby and M. Morrison, Prospects for environmental nanotechnologies. *Nanotechnol. Perceptions* 3 (2007) 193–207.

[9] E.g., H. Hidaka et al., Photoassisted dehalogenation and mineralization of chloro/fluorobenzoic acid derivatives in aqueous media. *J. Photochem. Photobiol.* A 197 (2008) 115–123.

Soil remediation is also mainly concerned with eliminating pollution. The decomposition of chlorinated hydrocarbons is catalyzed by magnetite (Fe_3O_4). Hence the addition of nanoparticulate iron oxide to soil is a possible remediation method. Unfortunately there is minimal documented experience to guide the would-be practitioner. These environmental applications would have to operate on a large-scale in order to be effective. The effects of releasing large numbers of nanoparticles into the biosphere are not known. Iron is generally presumed to be a rather benign element, but nanoparticles may be able to penetrate within the microbial organisms ubiquitous in cells with unknown effects on their vitality and on interspecies interactions.

6.9 FOOD

The most useful application of nanotechnology to the food industry is in enhancing packaging. A major problem of the industrial food industry is keeping processed food fresh. If packaging can more effectively act as a barrier (e.g., to oxygen), food can be kept fresher for longer prior to sale and opening the package. This functional enhancement is essentially achieved by transforming the polymer packaging film into a nanocomposite (see §6.1.1), by adding plate-like nanoparticles (e.g., certain clays) that enormously increase the tortuosity of diffusion pathways through the film. Worldwide sales of nanotechnology products to the food and beverage packaging sector increased to almost $900 million in 2004 from $150 million in 2002—again, the significance of this figure depends on exactly what is included. There were less than 40 identifiable nanopackaging products in the market 3 years ago, whereas there are about 250 at present. Considering that the global food packaging business is currently worth around $120 milliard, there is obviously considerable growth potential for introducing nanotechnology into this sector.

Major market trends in the food and beverage sector include: improving the performance of packaging materials in a passive sense (e.g., by increasing their transparency); prolonging the shelf life of the contents (e.g., by selectively managing the gas permeability of the packaging); improving sterility (e.g., by immobilizing antibiotics that kill microbes on contact within the packaging material); indicator packaging (which changes color if the package has been subjected to a deleteriously high temperature, for example, which renders the contents unfit for consumption but otherwise leaves no visible traces); and interactive packaging (which might respond to a potential customer touching it by changing color).

Nanotechnology allows molecular-scale structural alterations of packaging materials. With different nanostructures, plastics can gain various gas and water vapor permeabilities to fit the requirements of preserving fruit, vegetables, beverages, wines and other foods. By adding nanoparticles, manufacturers can produce bottles and packages with more light resistance and fire resistance, stronger mechanical and thermal performance and less gas absorption. Such nano-tweaking can significantly increase the shelf life of foods and preserve flavor and color. Nanostructured films coating the package can prevent microbial invasion and ensure food safety. With embedded nanosensors in the packaging, consumers will be able to determine whether food has gone bad or estimate its nutritional content.

Conceptual nanotechnology would cover attempts to understand nutrition from the molecular viewpoint, not only confining attention to the elemental composition of micronutrients, but also their material state. There is also some interest in "molecular gastronomy", a term coined by Nicolas Kurti and Hervé This in 1992 signifying the application of scientific laboratory techniques to cooking, and since enthusiastically taken up by a variety of chefs around the world, although it remains very much a niche activity.

Indirect nanotechnology is dominated by powerful microprocessors enabling computation to be all-pervasive. The farmer using a geographical information system to drive robots in his fields is making good use of that. In the future, butchers may routinely employ tomography on carcasses to determine the optimal dissection. The tomography itself requires heavy computations; nanotechnology-enabled processing power may become powerful enough to enable the optimal dissection to be automatically determined. Cold storage systems—and indeed the logistics of the entire global distribution system—are controlled by microprocessors.

Direct nanotechnology would cover the use of nanoscale sensors, cheap and unobtrusive enough to be ubiquitous, to monitor the state of food, including possible contamination with pesticides, or infectious agents acquired in the factory, or deficiencies arising through improper operations in a restaurant kitchen. Driven by a plethora of scandals leading to food poisoning, sometimes on quite a large scale, this is perceived to be a very welcome development by the general public, significantly offsetting some of the disadvantages of the modern agro-industrial complex. Benefits of a similar nature are already resulting from the use of nanocoatings for packaging materials, enhancing their desirable gas permeability characteristics, and sometimes incorporating an indicator function able to respond to (e.g.) the premature leakage of oxygen into a sealed package.[10]

[10] J.J. Ramsden, *Nanotechnology in Coatings, Inks and Adhesives*. Leatherhead: Pira International (2004).

The contentious aspect of nanotechnology concerns the possibility of including nanoscale nutritional additives in food, another manifestation of direct nanotechnology. Additives to enhance the nutritional value of food are already widespread in the processed food industry (a very common example is the addition of vitamin C to fruit juices). The idea behind using nanotechnology is to enhance the functionality and hence effectiveness of these additives—for example, encapsulating the vitamin C in minute hollow spheres made from calcium carbonate, so that the vitamin does not oxidize and become nutritionally valueless while the juice is standing in the air waiting to be drunk, but will only be released in the acidic environment of the stomach. Inasmuch as these additives are already becoming more and more sophisticated, introducing nanotechnology seems to be merely a continuation of an existing trend. Since nanotechnology is typically associated with achieving higher added value for a product, it is natural that it is of particular interest to the rather young field of "neutraceuticals"—foodstuffs deliberately enhanced with substances that would rank as pharmaceuticals. This development has in itself not been free of controversy—probably the best-known example is the addition of fluoride to drinking water.

The fundamental argument against this kind of thing is that our physiology is not adapted to such novelties, and may not be able to adapt before some harm is done. This constitutes the basic objection to ingesting genetically modified foods. It is quite difficult to find the right level at which to address the problem. Clearly DNA as such is not in general toxic—we are eating it all the time. On the other hand, *certain sequences*, e.g. those of a virus, are demonstrably harmful, at least under certain circumstances. The situation recalls the debates over the quality of drinking water in London in the middle of the 19th century—certain experts likened the inadvertent consumption of microorganisms in the water supply as being no more dangerous than eating fish. Given the state of knowledge at the time, it would perhaps have been difficult to adduce irrefutable evidence and arguments against that viewpoint.

The only way to proceed is to build up knowledge that can then be applied to calculate risks, and weighed against possible benefits. Provided the knowledge is there, this can be done quite objectively and reliably (see Section 14.3), but gaining the knowledge is likely to be a laborious task, especially when it comes to assessing the chronic effects resulting from many years of low-level exposure. There is particular anxiety regarding the addition of small metallic or metal oxide nanoparticles to food. Although a lot about their biological effects is indeed already known,[11] the matter is complex enough

[11] P.A. Revell, The biological effects of nanoparticles. *Nanotechnol. Perceptions* 2 (2006) 283–298.

for the ultimate fates of such particles in human bodies to be still rather poorly understood, and new types of nanoparticles are being made all the time.[12] On the other hand, it is also worth bearing in mind that some kinds of nanoparticles have been around for a long time—volcanoes and forest fires generate vast quantities of dust and smoke, virus particles are generally within the nano-range, comestible biological fluids such as milk contain soft nanoparticles, and so forth, merely considering natural sources. Anthropogenic sources include combustion in many forms, ranging from candles, oil lamps, tallow dips etc. used for indoor lighting, internal combustion engines—this is a major source of nanoparticle pollution in cities, along with the dust generated from demolishing buildings—cooking operations, and recreational smoking. The occupational hazards from certain industries, especially mining and mineral processing (silicosis, asbestosis), are well recognized, and the physicochemical and immunological aspects of the hazards of the nanoparticles reasonably well understood.[13]

A good example of how nanoscale knowledge has led to a profoundly new understanding of previously unsuspected hazards is provided by the discovery, using the black lipid membrane (BLM) technique,[14] that certain cyclic polyunsaturated compounds synthesized by bacteria used for the biotechnological production of artificial sweeteners are able to form ion channels in our cell membranes. Trace quantities of these compounds remain in the so-called "high energy" and other soft drinks that seem to enjoy a growing popularity, and may be responsible for the growing incidence of heart problems among teenagers in societies where the consumption of these beverages has become the norm. Knowing this is one thing; it is another matter to diffuse the knowledge among the general public, in order that they may weigh the risks against the supposed enjoyment.

An important aspect of the current debate on the matter concerns the possibility of choice. Ideally, if knowledge is insufficient for it to be clear

[12] It is actually quite inadequate to refer generically to nanoparticles. It is already known that their toxicity depends on size, shape and chemical constitution, and very possibly on their state of crystallinity (think of the problems of polymorphism of active ingredients in the pharmaceutical industry!). Therefore, at the very least some information on these characteristics should be provided when referring to "nanoparticles".

[13] C.J. van Oss, J.O. Naim, P.M. Costanzo, R.F. Giese Jr., W. Wu and A.F. Sorling. Impact of different asbestos species and other mineral particles on pulmonary pathogenesis. *Clays Clay Minerals* 47 (1999) 697–707.

[14] P.A. Grigoriev, Unified carrier-channel model of ion transfer across lipid bilayer membranes. *J. Biol. Phys. Chem.* 2 (2002) 77–79.

whether benefits or risks are preponderant, a product should be available both with and without the nanoadditive. Then the consumer can make his or her choice—*caveat emptor*.[15] In reality, it is well known that choice tends to contract. For example, nearly all the world's soybeans come from a certain (genetically modified) variety; 99% of tomatoes grown in Turkey are no longer indigenous varieties.[16] It appears to be empirically well established that mysterious "market forces" drive matters to this result, and the presence or absence of choice needs to be taken into account when it is discussed whether nanoadditives should be permitted or not. We are familiar with the state of affairs in traditional (non-nano) food processing. For example, it is possible to buy dairy products, such as cheese, made from either raw or pasteurized milk. For some consumers, avoiding the risk of contracting tuberculosis is the preponderant consideration; for others, the undesirability of consuming advanced glycation end-products (AGEs, resulting from the chemical reaction between sugars and animal proteins or fats, typically taking place during frying or pasteurization) outweighs that risk; for yet others taste is the important consideration.

It is perhaps appropriate at this point to raise the question of regulation. *Caveat emptor* is a universal injunction, which should actually render regulation superfluous, since it is generally called for to protect the unsuspecting consumer from unscrupulous purveyors of goods or services, but if the consumer took the trouble to properly investigate what he was letting himself in for, presumably the good or service would not be bought, and the unscrupulous purveyor would be less likely to continue to attempt to sell whatever it was, profit presumably being the sole motive. This is an example of market forces at work.

Yet, despite repeatedly hearing that developed countries are "knowledge-based economies", it seems that the knowledge necessary to properly apply the principle of *caveat emptor* is lacking among the general public. Moreover, in many countries this knowledge is lacking among the legislative bodies. Therefore, one of the most urgent political needs is to ensure that parliamentarians and the like raise their level of knowledge and understanding of our technologically advanced society to ensure that they can effectively fill the knowledge gap between the technologists and the still largely ignorant

[15] This, incidentally, highlights the importance of the members of society being sufficiently knowledgeable to be able to make an informed choice (see also Chapter 14).

[16] For further examples, see J.J. Ramsden, Complex technology: a promoter of security and insecurity. In: J.J. Ramsden and P.J. Kervalishvili (eds), *Complexity and Security*, pp. 249–264. Amsterdam: IOS Press (2008).

consumer. A discussion of how this might be done goes beyond the scope of this already lengthy section, however.[17]

Less controversial than nanoadditives, but just as nano, are developments to achieve not a chemical modification of a foodstuff, but a modification to its *structure*. This particularly affects not taste, which is perhaps above all dependent on the actual chemicals sensed by our tastebuds, but mouthfeel, which is a very difficult characteristic of a natural foodstuff to imitate. Hence the food processing industry is devoting a great deal of ingenuity to find (with the application of nanometrology, since it does seem that physical structure at the nanoscale is responsible) ways of mimicking desirable mouthfeel, in products such as "ice cream" (which, if industrially manufactured, may contain very little cream).

The Social Context "Man isst, was man isst", as Martin Luther famously remarked in one of his *Tischgespräche*. Given the centrality of food for human existence, it can hardly be discussed in isolation as a purely physiological matter. Indeed, there is even evidence that the intake of folic acid by a pregnant mother can influence the methylation of the baby's proteins.[18] Furthermore, it is too much to expect that we always eat "sensible" foods, or carefully examine the list of ingredients on a packet (which anyway is usually woefully inadequate, particularly regarding the actual quantities of the substances mentioned), or acquire a personal biosensor for verifying on the spot the absence of hormone-active substances in vegetables sold on the market. It may be nowadays trite to repeat John Donne's dictum "No man is an island", but if anything it is even more true today, in a world wide web-connected age, than it was at the end of the Middle Ages. We are all affected by the foods around us, whether we partake of them or not.

The dominant social aspect of nutrition is malnutrition coupled with obesity. It is estimated that current world production of food is adequate for the current world population, but much of that food is in the wrong place at the wrong time. Technologies, such as cold storage and nanoparticle-containing gas-resistant wrappers, should therefore contribute to alleviating unevennesses of supply. The technologies come at a price, however; for example, many modern farming practices, including intensive agriculture of all kinds

[17] Nevertheless, it is worth remarking that we seem to be no closer to the dilemma, repeatedly pointed out by dispassionate observers during the last 50 years, posed by two equally unsatisfactory possible solutions, namely giving scientists control of our society, and allowing governments to be only very feebly influenced by rationality. New thinking to solve this problem is very much needed!

[18] The methylation pattern of the repertoire of genes is a key controlling factor in development.

and fish farming, tend to yield produce that is less wholesome than their nonintensive counterparts.

The solution to eating the wrong foods—such as those that leave one undernourished, or overweight or both—is surely more knowledge. This is the perennial problem of a society based on high technology—it can only be truly successful if all members are sufficiently knowledgeable to properly partake in its development. Hence we also need to inquire how nanotechnology can contribute to the education of the population.

Nanotechnology and the Food Crisis In June 2008 the Food and Agriculture Organization (FAO, part of the United Nations) held a 3-day summit conference in Rome in order to explore ways of overcoming problems caused by steeply rising food prices around the world, which have caused especially grave problems in poor countries. Although part of the problem lies in the commercial sphere, and may be dealt with by considering the effects of export restrictions and price controls, a sustainable solution clearly lies in the technological realm. In the short term, charitable deliveries of seeds and fertilizers may alleviate the problem; in the medium term such measures are likely to make things worse. Therefore, a thorough appraisal of the state of agronomy is needed. In fact, a number of recent reports have pointed to the research deficit in the field accumulated during the last few decades.[19] Yet in its call for increasing public support (in the developing countries) for agronomy research, the FAO is essentially still thinking of traditional approaches. In view of the generally revolutionary nature of nanotechnology, it must be expected that here too it can make a decisive contribution.

There seem to be two timescales involved. One covers the next few years, and is based on the intense nanoscale scrutiny of the processes of comestible biomass production in its entirety, followed by inspired intervention at that scale. An example is biological nitrogen fixation. A wealth of detail is already known about the process: at the molecular level (the nitrogenase enzymes responsible for actually fixing the nitrogen); at the microbiological level (the symbiotic rhizobia); and at the ecological level (soil and inoculation, although here there are still inexplicable mysteries). Intervention, e.g. with functionalized nanoparticles, especially to improve fixation in difficult (e.g., dry or saline) conditions, seems to be feasible. The actual need is for laboratory and field research to establish the possibilities and limitations of such an approach.

---

[19] *World Development Report 2008: Agriculture for Development*. Washington, DC: World Bank (2007).

The more distant timescale involves the introduction of molecular manufacturing. Should this ever come about, anything, including any kind of foodstuff, could be made from acetylene, ammonia, oxygen, phosphorus (along with some metals essential for our enzymes) and electrical power. This, more than anything else, will revolutionize the world order; whether hunger is abolished will depend on politics (and demography).

Looking back over the past millennia of human civilization, improvements in technology have enormously increased agricultural output, but this has also led to a concomitant increase in world population, hence global nutritional difficulties remain. Geographical mismatch of supply and demand is also frequently mentioned as a contributor to malnutrition—somewhat ironically, in an age of unprecedentedly large global trade volumes. A serious current problem is that it is becoming increasingly clear that product volume increases imply product quality decreases. This goes well beyond mere unpalatability.[20] The output of the agro-industrial complex, unfortunately including residues of pesticides and the presence of hormone-active substances, may solve the basic malnutrition problem, but may introduce new problems of ill-health that may be deeply unsustainable, although the manufacturers of pharmaceuticals may see it as a source of new opportunities.

6.10 METROLOGY

The primary products included in this category are the scanning probe microscopes that are indispensable for observing processes at the nanoscale, and which may even be used to assemble prototypes. The market is, however, relatively small in value—estimated at around $250 million per annum for the whole world. This represents about a quarter of the total microscope market. Optical microscopes have a similar share (but are presently declining, whereas scanning probe microscopes are increasing), and electron microscopes have half the global market.

At the other end of the spectrum are telescopes, looking at very large and very distant objects. New generations of space telescopes and terrestrial telescopes used for astronomy require optical components (especially mirrors) finished to nanoscale precision. The current concept for very large terrestrial telescopes is to segment the mirrors into a large number of unique pieces (of

[20] "Mere" perhaps belies the significant contribution of the enjoyment of food to social harmony, creativity, etc.

the order of one square meter in area), the surface of each of which must be finished to nanoscale precision.[21]

6.11 PAPER

This commodity is made in vast quantities (globally, about 100 million tonnes per annum) in most countries in the world. The primary constituent is cellulose fiber, but as much as half of the annual production contains nanoparticles (0.02–0.2% of the total mass). The use of such "fillers" in paper-making has a long history.[22] The purpose is to better control attributes such as porosity, reflectance and ink absorption. A new application for nanoparticles is to tag sheets of paper with distinguishable nanoparticles (for example, made up from different metals) for security and identification purposes. Individual cellulose fibers are being coated with nanoscale polyelectrolyte films in order to enhance strength and other attributes of paper such as electrical conductivity.[23] The coating is a self-assembly process whereby the fiber is merely dipped in a solution of the polyelectrolyte. Ease of manufacture makes the treatment quite cost-effective (e.g., by doubling the tensile strength, single-ply sacks can be used instead of double-ply, but the cost per unit area of the paper is less than double that of the untreated material).

6.12 SECURITY

Although military organizations such as the Department of Defense in the USA are spending a great deal on nanotechnology, most of the applications are generic. In other words, most military applications of nanotechnology currently under investigation are adaptations of civilian products.[24] Homeland security is heavily focused on the detection of explosives. This calls for chemical sensors of trace volatile components, using the same kind of

---

[21] P. Shore, Ultra precision surfaces. *Proc. ASPE*, pp. 75–80. Portland, OR (2008).

[22] M.A. Hubbe, *Emerging Technologies in Wet End Chemistry*. Leatherhead: Pira International (2005).

[23] Z. Zheng, J. McDonald, T. Shutava, G. Grozdits and Yu. Lvov, Layer-by-layer nanocoating of lignocellulose fibers for enhanced paper properties. *J. Nanosci. Nanotechnol.* 6 (2006) 324–332.

[24] J. Altmann, *Military Nanotechnology*. London: Routledge (2006).

technology as is used for medical applications (see Chapter 8). Nanotechnology also enters into the video surveillance technology rapidly becoming ubiquitous in the civilian world, notably through the great processing power required for automated pattern recognition.

6.13 TEXTILES

A natural textile fiber such as cotton has intricate nanostructure; the comfortable properties of many traditional textiles result from a favorable combination of chemistry and morphology. Understanding these factors allows the properties of natural textiles to be equaled or even surpassed by synthetic materials.

Furthermore, nanoadditives can enhance textile fibers with properties unknown in the natural world, such as ultrastrength, ultradurability, flame resistance, self-cleaning capability, modifiable color, antiseptic action and so forth. Textiles releasing useful chemicals, either passively or actively, are also conceivable (of which the antiseptic textile, in which silver nanoparticles are incorporated, is a simple example). Such functionally enhanced textiles are typically used in specialty applications, such as serving as a living cell scaffold assisting tissue regeneration, and as wound dressings assisting healing.

CHAPTER 7

Information Technologies

CHAPTER CONTENTS

It has already been pointed out that in information processing (including data storage) applications, nanotechnology offers many advantages because the intrinsic lower limit of the representation of one bit of information is around the atomic (nano) scale. The process of nanification of information processing technology is well represented by Moore's law—which in its original form states that the number of components (i.e., resistors, capacitors and resistors) per chip doubles each year.[1] When Moore revisited this prediction 10 years later,[2] he somewhat refined this statement, pointing out that the result was the consequence of three technological developments: increasing area per chip, decreasing feature size (defined as the average of linewidth

[1]G.E. Moore, Cramming more components onto integrated circuits, *Electronics* 38 (19 April 1965) 114–117.

[2]G.E. Moore, Progress in digital integrated electronics. In: *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 11–13. Washington, D.C. (1975).

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

and spacewidth), and improved design of both the individual devices and the circuit. Only the second of these three is a nanification process.

The direct economic consequence of these technological developments is a roughly constant cost per area of processed silicon, while the processing power delivered by the chip becomes steadily greater. Furthermore, nanification makes the transistors not only smaller, but also lighter in weight, faster (because the electrons have less distance to travel), less power-hungry and more reliable. These are all strong selling points. Therefore, although technology push is undoubtedly important in maintaining Moore's law, the ultimate driver is economic.

As a result of these developments the microprocessor, which nowadays contains nanoscale components,[3] has become ubiquitous throughout the world. For example, even a small company employing fewer than 50 people probably uses a computer to administer salaries etc. (even though it would almost certainly be cheaper and more effective to do it manually).

7.1 SILICON MICROELECTRONICS

The starting point of chip production is the so-called wafer, the circular disc cut from a single crystal of silicon with a diameter of at least 300 mm and a thickness typically between 500 and 800 μm. Using lithography and etching technology the structures of integrated circuits are built up layer by layer on the surface of a chip.[4] Transistor construction has been based on complementary metal oxide semiconductor (CMOS) technology for decades. The size of the smallest structures, currently standing at 65 nm, has been steadily diminishing and should already have fallen to 45 nm in 2009, to 32 nm in 2012, and 22 nm in 2015 (the transistor "roadmap")—this last value is close to the operational limit for metal oxide semiconductor field effect transistor technology. These developments represent tremendous technological challenges, not only in the fabrication process itself, but also in testing the finished circuits and in heat management—a modern high-performance chip may well dissipate heat at a density of 100 W/cm^2, greater than that of a domestic cooking plate.

---

[3]Nowadays, the sizes of apparatus such as a cellphone or a laptop computer are limited by peripherals such as screen, keyboard and power supply, not by the size of the information processing unit.

[4]A.G. Mamalis, A. Markopoulos and D.E. Manolakos, Micro and nanoprocessing techniques and applications. *Nanotechnol. Perceptions* 1 (2005) 63–73.

Silicon itself is still foreseen as the primary semiconducting material (although germanium, gallium arsenide, etc. continue to be investigated), but in order to fabricate ever-smaller structures, new photoresists will have to be developed. Furthermore, the silicon oxide thin film, which insulates the gate from the channel in the field effect transistor, becomes less and less effective as it becomes thinner and thinner (of the order of 1 nm). Other metal oxides (e.g., hafnium oxide) are being investigated as alternative candidates.

Some design issues arising from this relentless miniaturization are discussed in Chapter 11.

7.2 DATA STORAGE TECHNOLOGIES

Electrons have spin as well as charge. This is of course the origin of ferromagnetism, and hence magnetic memories, but their miniaturization has been limited not by the ultimate size of a ferromagnetic domain but by the sensitivity of magnetic sensors. In other words, the main limitation has not been the ability to make very small storage cells, but the ability to detect very small magnetic fields.

The influence of spin on electron conductivity was invoked by Nevill Mott in 1936, but remained practically uninvestigated and unexploited until the discovery of giant magnetoresistance (GMR) in 1988. The main present application of spintronics (loosely defined as the technology of devices in which electron spin plays a rôle) is the development of ultrasensitive magnetic sensors for reading magnetic memories. Spin transistors, in which the barrier height is determined by controlling the nature of the electron spins moving across it, and devices in which logical states are represented by spin belong to the future (Chapter 12).

Giant magnetoresistance (GMR) is observed in thin (a few nanometers) alternating layers (superlattices) of ferromagnetic and nonmagnetic metals (e.g., iron and chromium).[5] Depending on the width of the nonmagnetic spacer layer, there can be a ferromagnetic or antiferromagnetic interaction between the magnetic layers, and the antiferromagnetic state of the magnetic layers can be transformed into the ferromagnetic state by an external magnetic field. The spin-dependent scattering of the conduction electrons in the nonmagnetic layer is minimal, causing a small resistance of the material, when the magnetic moments of the neighboring layers are aligned in parallel,

[5] M.N. Baibach, J.M. Broto, A. Fert, F. Nguyen Van Dau, F. Petroff, P. Etienne, G. Creuzet, A. Friederich and J. Chazelas. Giant magnetoresistance of (001)Fe/(001)Cr magnetic superlattices. *Phys. Rev. Lett.* 61 (1988) 2472–2475.

whereas for the antiparallel alignment the resistance is high. The technology is nowadays used for the read–write heads in computer hard drives. The discovery of GMR depended on the development of methods for making high-quality ultrathin films (such as molecular beam epitaxy).

A second type of magnetic sensor is based on the magnetic tunnel junction (MTJ), in which a very thin dielectric layer separates ferromagnetic (electrode) layers, and electrons tunnel through this nonconducting barrier under the influence of an applied voltage. The tunnel conductivity depends on the relative orientation of the electrode magnetizations and the tunnel magnetoresistance (TMR): it is low for parallel alignment of electrode magnetization and high in the opposite case. The magnetic field sensitivity is even greater than for GMR. MTJ devices also have high impedance, enabling large signal outputs. In contrast with GMR devices, the electrodes are magnetically independent and can have different critical fields for changing the magnetic moment orientation. The first laboratory samples of MTJ structures (NiFe–Al_2O_3–Co) were demonstrated in 1995.

7.3 DISPLAY TECHNOLOGIES

The results of a computation must, usually, ultimately be displayed to the human user. Traditional cathode ray tubes have been largely displaced by the much more compact liquid crystal displays (despite their disadvantages of slow refresh rate, restricted viewing angle and the need for back lighting). The main current rival of liquid crystal displays are organic light-emitting diodes (OLEDs). They are constituted from an emissive (electroluminescent), conducting organic polymer layer placed between an anode and a cathode (see also Section 9.7).

Any light-emitting diode requires one of the two electrodes to be transparent. Traditionally indium-doped tin oxide (ITO) has been used, but the world supply of indium is severely limited, and at current rates of consumption may be completely exhausted within 2 or 3 years. Meanwhile, relentless onward miniaturization and integration make it more and more difficult to effectively recover indium from discarded components. Hence there is great interest in transparent polymers doped with a small volume percent of carbon nanotubes to make them electrically conducting (see Section 6.2).

7.4 SENSING TECHNOLOGIES

Information technology has traditionally focused on arithmetical operations, but information transduction belongs equally well to the field. Information represented as the irradiance of a certain wavelength of light, or the bulk concentration of a certain chemical, can be converted (transduced) into an electrical signal. From careful consideration of the construction of sensors consisting of arrays of discrete sensing elements, it can be clearly deduced that atomically precise engineering will enable particle detection efficiency to approach its theoretical limit.[6] Since a major application of such sensors is to clinical testing, they are considered again in the next chapter.

[6] S. Manghani and J.J. Ramsden, The efficiency of chemical detectors. *J. Biol. Phys. Chem.* 3 (2003) 11–17.

CHAPTER 8

Applications to Health

CHAPTER CONTENTS

Nanomedicine is defined as the application of nanotechnology to human health. The dictionary definition of medicine is "the science and art concerned with the cure, alleviation and prevention of disease, and with the restoration and preservation of health". As one of the oldest of human activities accessory to survival, it has indeed made enormous strides during the millennia of human civilization. Its foremost concern is well captured by a phrase in the spirit of Hippocrates, "Primum non nocere". During the past few hundred years, and especially during the past few decades, medicine has been characterized by enormous technization, with a concomitantly enormous expansion of its possibilities for curing disease. The application of nanotechnology, the latest scientific–technical revolution, is a natural continuation of this trend.

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

Medicine is, of course, closely allied to biology, and molecular biology might well be called an example of conceptual nanotechnology—scrutinizing a system with regard to its structure at the nanoscale. Furthermore, much of the actual work of the molecular biologist increasingly involves nanometrology, such as the use of scanning probe microscopies.

Rather closer to nanotechnology is the mimicry, by artificial means, of natural materials, devices and systems structured at the nanoscale. Ever since Drexler presented biology as a "living proof of principle" of nanotechnology,[1] there has been a close relationship between biology and nanotechnology.

It is customary nowadays to take a global view of things, and in assessing the likely impact of nanotechnology on medicine this is very necessary. Nanotechnology is often viewed to be the key to far-reaching social changes (this theme will be taken up again in Chapter 14), and once we admit this link then we really have to consider the gamut of major current challenges to human civilization, such as demographic trends (overpopulation, aging), climate change, pollution, exhaustion of natural resources (including fuels), and so forth. Nanotechnology is likely to influence many of these, and all of them have some implications for human health. Turning again to the dictionary, medicine is also defined as "the art of restoring and preserving health by means of remedial substances *and the regulation of diet, habits, etc.*" It would be woefully inadequate if the impact of nanotechnology on medicine were restricted to consideration of the development of more sophisticated ways of packaging and delivering drugs (important as that is).

8.1 PRINCIPAL APPLICATIONS

The three main developments currently envisaged by leading pharmaceutical companies are: sensorization,[2] automated diagnosis, and customized pharmaceuticals. Sensorization belongs predominantly to direct nanotechnology. With their ever-diminishing footprint, nanoscale sensors are not only able to penetrate inside the body via minimally invasive procedures such as endoscopy, but are moving towards the ability to be permanently implanted. The latter are potentially capable of yielding continuous outputs

[1] K.E. Drexler, Molecular engineering: an approach to the development of general capabilities for molecular manipulation. *Proc. Natl Acad. Sci. USA* 78 (1981) 5275–5278.

[2] This word is defined as meaning "embedding large numbers of sensors in a structure".

of physiologically relevant physicochemical parameters such as temperature and the concentrations of selected biomarkers. The downside of these developments is the enormous quantity of data that needs to be handled, but here indirect nanotechnology comes to the rescue, with ever-increasing information processing power becoming available. One of the greatest current challenges is the automatic diagnosis of disease. If it is generally true that "about 85% of [medical examination] questions require only recall of isolated bits of factual information",[3] this looks to be achievable even by currently available computing systems. The third development, (affordable) customized pharmaceuticals, is supposed to be enabled by miniaturized (microfluidic) mixers and reactors, but this is micro rather than nano and, hence, outside the scope of this book.

8.2 IMPLANTED DEVICES

Prostheses and biomedical devices must be biocompatible.[4] This attribute can take either of two forms: (i) implants fulfilling a structural rôle (such as bone replacements) must become assimilated with the host; failure of assimilation typically means that the implant becomes coated with a layer of fibrous material within which it can move, causing irritation and weakening the structural rôle; (ii) for implants in the bloodstream (such as stents, and possibly implanted sensors in the future) the opposite property is required: blood proteins must not adsorb on them. Adsorption has two deleterious effects: layers of protein build up and may clog the blood vessel; or the proteins that adsorb may become denatured, hence foreign to the host organism and triggering inflammatory immune responses.

In order to promote assimilation, a favorable nanotexture of the surface seems to be necessary, to which the cell responds by excreting extracellular matrix molecules, humanizing the implant surface. Years of empirical studies have enabled this to be achieved in some cases, e.g. the surfaces of cell culture flasks. Intensive research is meanwhile under way to provide the basis for a more rational design of the surfaces with a pattern specified at the atomic level; it is still not known whether the pattern only needs to fulfill certain statistical features.

---

[3]According to a University of Illinois study by G. Miller and C. McGuire (quoted by W.E. Fabb, Conceptual leaps in family medicine: are there more to come? *Asia Pacific Family Med.* 1 (2002) 67–73).

[4]J.J. Ramsden, *Biomedical Surfaces*. Norwood, MA: Artech House (2008).

In order to prevent adsorption, its free energy (ΔG_{123}) is analyzed according to:[5]

$$\Delta G_{123} = \Delta G_{22} + \Delta G_{13} - \Delta G_{12} - \Delta G_{23} \quad (8.1)$$

where subscripts 1, 2 and 3 denote the implant (surface), the biofluid (blood) and the protein respectively. ΔG_{22} is thus the cohesive energy of water, which is so large that this term alone will ensure that adsorption occurs unless it is countered by strong hydration. The biomedical engineer cannot influence ΔG_{23} and must therefore design ΔG_{12} appropriately. Coating material 1 with an extremely hydrophilic material such as poly(ethylene oxide) (PEO) is one way of achieving this.

For medical devices that are not implanted, such as scalpels or needles, attention is paid to finishing them in such a way that they cut the skin very cleanly, minimizing pain, and have low coefficients of friction to allow penetration with minimal force.[6] This may be achieved by ultraprecision machining, finishing the surfaces to nanometer-scale roughness.

Long-term implants must be designed in such a way as not to host adventitious infection by bacteria. Once they colonize an implant, their phenotype usually changes and they may be resistant to the attentions of the body's immune system (thus causing persistent inflammation without being destroyed) and to antibiotics.

Implants with rubbing surfaces, such as joint replacements, typically generate particles as a result of wear. Traditional tribopairs such as metal–polyethylene generate relatively large microparticles (causing inflammation); novel nanomaterials with otherwise improved properties may generate nanoparticles, with unknown consequences.

8.3 NANOPARTICLE APPLICATIONS

The oldest well-documented example of the use of nanoparticles in medicine is perhaps Paracelsus's deliberate synthesis of gold nanoparticles (called "soluble gold") as a pharmaceutical preparation.[7] The use of nanoparticles in medicine has recently become a burgeoning field of activity. Applications

[5] M.G. Cacace, E.M. Landau and J.J. Ramsden, The Hofmeister series: salt and solvent effects on interfacial phenomena. *Q. Rev. Biophys.* 30 (1997) 241–278.

[6] J.J. Ramsden et al., The design and manufacture of biomedical surfaces. *Annals CIRP* 56/2 (2007) 687–711.

[7] See R. Zsigmondy and P.A. Thiessen, *Das kolloide Gold*. Leipzig: Akademische Verlagsgesellschaft (1925).

include: magnetic nanoparticles steered by external fields to the site of a tumor, and then energized by an external electromagnetic field in order to destroy cells with which the particle is in contact; nanoparticles as carriers for drugs; and nanoparticles as "sensors" (perhaps "markers" would be a more accurate word) for diagnosis (and as a tool for investigating biochemical mechanisms).

Nanoparticles for drug delivery are being intensively researched and developed, and many products are undergoing clinical trials. A major hindrance to successfully developing new drugs is the fact that many of the candidate molecules showing good therapeutic interaction with a target (e.g., an enzyme) are very poorly soluble in water. Such compounds can be encapsulated in nanoparticles with a hydrophilic outer surface. Such a surface is also important for preventing the adsorption and adverse immune response-triggering denaturation of proteins during the passage of the particle through the bloodstream (cf. Section 8.2).

An example of a "smart" (probably it is more accurate to describe it as merely "responsive") drug delivery particle is a hollow shell of calcium carbonate destined for the stomach: the strongly acidic environment there will dissolve the mineral shell away, releasing the contents.

A niche market for nanoparticles is in molecular biology and clinical research. They can be useful as biomarkers: by coating them with chemicals having a specific affinity for certain targets (e.g., antibodies), the locations of those targets can be much more easily mapped using microscopy. The particles might simply be heavy metals, which are easy to see in the electron microscope, or they may fluoresce, in which case the nanoparticles substitute for organic fluorescent dyes.

8.4 TISSUE SCAFFOLDS

It is now known that the extracellular matrix (ECM), which acts as a scaffold in the body on which cells grow, has a complex structure made up of several different kinds of large protein molecules (e.g., laminin, tenascin). The responses of cells in contact with the ECM has revealed dramatic changes correlated with subtle differences in the molecules. The main research question at present is to determine what features of artificially nanostructured substrata are required to induce a cell to differentiate in a certain way. An enormous quantity of results has already been accumulated, but overall it seems not to have been sufficiently critically reviewed, and therefore it is difficult at present to discern guiding principles, other than rather trivial ones.

8.5 PARAMEDICINE

The use of toxic materials for cosmetic purposes (e.g., applying them to the skin of the face) has a long history—antimony salts were popular among the Romans, for example. Advances in our knowledge of toxicity have since then ushered in far more benign materials, although the recent use of extremely fine particles (for example, zinc oxide nanoparticles in sunscreens) has raised new concerns about the possibility of their penetration through the outer layers of the skin, or penetration into cell interiors, with unknown effects. The data given in Table 5.3 testify to the popularity of nanotechnology in this area. Many modern cosmetic products are amazingly sophisticated in their nanostructure. An important goal is to devise delivery structures for poorly water-soluble ingredients such as vitamin A and related compounds, vitamin E, lycopene, and so forth. The liposome (a lipid bilayer enclosing an aqueous core; i.e., a vesicle) is one of the most important structures; the first liposome-based cosmetic product was launched by Dior in 1986. Variants such as "transferosomes" (liposomes with enhanced elasticity), "niosomes" (using non-ionic surfactants instead of the lipid bilayer), "nanostructured lipid carriers", "lipid nanoparticles" and "cubosomes" (fragments of the bicontinuous cubic phase of certain lipids) point to the intense development activity in the field.

8.6 NANOBOTS

Microscopic or nanoscopic robots are an extension of existing ingestible devices that slowly move through the gastrointestinal tract and gather information (mainly images). As pointed out by Hogg,[8] minimal capabilities required of future devices are: (chemical) sensing; communication (receiving information from, and transmitting information to, outside the body, and communication with other nanobots); locomotion—operating at very low Reynolds numbers, estimated at about 1/1000 (i.e., viscosity dominates inertia); computation (e.g., recognizing a biomarker would typically involve comparing sensor output to some preset threshold value; due to the tiny volumes available, highly miniaturized molecular electronics would be very attractive for constructing on-board logic circuits); and of course power—it is estimated that picowatts would be necessary for propelling a nanobot at a speed of around 1 mm/s. It is very likely that to be effective, these nanobots

[8] T. Hogg, Evaluating microscopic robots for medical diagnosis and treatment. *Nanotechnol. Perceptions* 3 (2007) 63–73.

would have to operate in swarms, putting an added premium on their ability to communicate.

8.7 TOXICOLOGY ASPECTS

It may strike the reader as somewhat incongruous that, on the one hand, we have seen elsewhere (Section 5.5) that there are concerns about the toxicity of nanoparticles and strong recommendations to undertake more extensive and systematic investigations of the matter; on the other hand, they are already being widely incorporated into or proposed for pharmaceutical products. At least any medicinal preparation is subjected to a strict régime of testing before it is released for general use, but cosmetics are subject to much lighter requirements. Section 14.2 further discusses issues of regulation.

FURTHER READING

J.J. Ramsden, The rôle of biology, physics and chemistry in human health. *J. Biol. Phys. Chem.* 7 (2007) 153–158.

CHAPTER 9

The Business Environment

CHAPTER CONTENTS

In Chapter 3, innovation was examined as an essential part of the development of nanotechnology. This theme is now taken up again with particular reference to the financing of nanotechnology enterprises.

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

9.1 THE UNIVERSALITY OF NANOTECHNOLOGY

Reference is often made to the diversity of nanotechnology. Indeed, some writers insist on referring to it in the plural as nanotechnologies (perhaps an unnecessarily refined nuance). Inevitably, a technology concerned with building matter up atom by atom is a universal technology with enormous breadth.[1] Nanostructured materials are incorporated into nanoscale devices, which in turn are incorporated into many products, as documented in Part 2. An artefact is considered to be part of nanotechnology if it contains nanostructured materials or nanoscale devices even if the artefact itself is of microscopic size; this is the domain of indirect nanotechnology.[2] The fact that the feature sizes of components on semiconductor microprocessor chips are now smaller than 100 nm, and hence within the nanoscale, means that practically the entire realm of information technology has become absorbed by nanotechnology. Nanotechnology is, therefore, already pervasive.[3] The best current example of such a universal technology is probably information technology, which is used in countless products.

Any universal technology—and especially one that deals with individual atoms directly—is almost inevitably going to be highly upstream in the supply chain. This is certainly the case with nanotechnology at present. Only in the case of developments whose details are still too nebulous to allow one to be anything but vague regarding the timescale of their realization, such as quantum computers and general-purpose atom-by-atom assemblers, would we have pervasive direct nanotechnology.

Universal technologies form the basis of new value creation for a broad range of industries; that is, they have "breadth". Such technologies have some special difficulties associated with their commercialization because of their upstream position far from the ultimate application (see Figure 9.1).

The most important difficulty is that the original equipment manufacturer (OEM) needs to be persuaded of the advantage of incorporating the nanoscale component or nanomaterial into the equipment. The most convincing way of doing this is to construct a prototype. But if the technology is

[1] Synonyms for "universal" are "generic", "general purpose" and "platform".

[2] J.J. Ramsden, What is nanotechnology? *Nanotechnol. Perceptions* 1 (2005) 3–17.

[3] There is a certain ambiguity here, since the nanoscale processors (which, I suppose, should now be called nanoprocessors) have only been introduced very recently. Hence, the majority of extant information processors strictly speaking belong to microtechnology rather than nanotechnology.

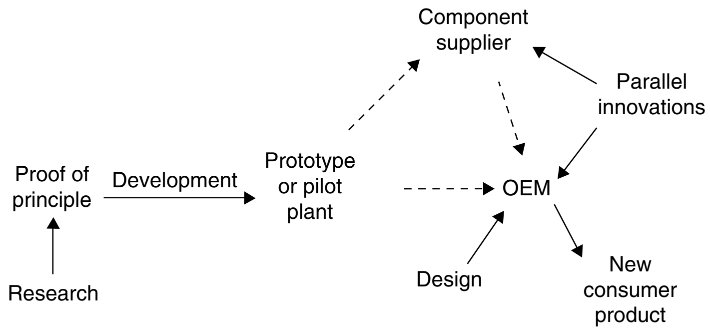


FIGURE 9.1 *Diagram of immediate effects showing the supply chain from research to consumer product. The dashed lines indicate optional pathways: the route to the original equipment manufacturer (OEM) is very likely to run via one or more component suppliers. Parallel innovations may be required for realization of the equipment. These include legally binding regulatory requirements (particularly important in some fields—e.g., gas sensors). Note that most of the elements of Porter's value chain are included in the last arrow from OEM to consumer product.*

several steps upstream from the equipment, constructing such a prototype is likely to be hugely expensive (presumably it will anyway be outside the domain of expertise of the nanotechnology supplier, so will have to be outsourced). The difficulty is compounded by the fact that many OEMs, especially in the important automotive branch, as well as "Tier 1" suppliers, rarely pay for prototype development. The difficulty is even greater if a decision to proceed is taken at the ultimate downstream position, that of the consumer product itself. The nanotechnology supplier, which as a start-up company is typically in possession of only proof of principle, often obtained from the university laboratory whence the company sprang, is likely to be required to make its most expensive investments (e.g., for a prototype device or an operational pilot plant) before it has had any customer feedback.

This distance between the technology and its ultimate application will continue to make life difficult for the technologist even if the product containing his technology is introduced commercially, because the point at which the most valuable feedback is available—from the consumer—is so far away. There is perhaps an analogy with Darwin's theory of evolution here, in its modern interpretation that incorporates knowledge of the molecular nature of the gene—variety is introduced at the level of the genome (e.g., via mutations), but selection operates a long way downstream, at the level of the organism. The disparity between loci is especially acute when the exigencies of survival include responses to potentially fatal sudden threats.

The further upstream one is, the more difficult it is to "capture value" (i.e., generate profit) from one's technology.[4] Hence cash tends to be limited, hence the possibilities for financing the construction of demonstration prototypes.

The difficulty of the position of the upstream technologist is probably as low as it can be if the product or process is one of substitution. As will be seen later (§9.7) this is likely to be a successful path for a small nanotechnology company to follow. In this case, demonstration of the benefits of the nanotechnology is likely to be relatively straightforward, and might even be undertaken by the downstream client.

On the other hand, the nanotechnology revolution is unlikely to be realized merely by substitutions. Much contemporary nanotechnology is concerned with a greater innovative step, that of miniaturization (or nanification, as miniaturization down to the nanoscale is called)—see Figure 1.1. As with the case of direct substitution, the advantages should be easy to describe and the consequences easy to predict, even if an actual demonstration is likely to be slightly more difficult to achieve.

A curious, but apparently quite common difficulty encountered by highly upstream nanotechnology suppliers is related to the paradox (attributed to Jean Buridan) illustrated by an ass placed equidistantly between two equally attractive piles of food and unable to decide which one to eat first, and which starved to death through inaction. Potential buyers of nanoparticles have complained that manufacturers tell them "we can make any kind of nanoparticle". This is unhelpful for many downstream clients, because their knowledge of nanotechnology might be very rudimentary, and they actually need advice on the specification of which nanoparticles will enhance their product range. Start-up companies that offer a very broad product range typically are far less successful than those that have focused on one narrow application, despite the allure of universality (see Section 9.7), not least because—ostensibly—it widens the potential market.

On the other hand, for a larger company universal technologies are attractive commercial propositions. They allow flexibility to pursue alternative

---

[4] This can be considered as quasi-axiomatic. It seems to apply to a very broad range of situations. For example, in agriculture the primary grower usually obtains the smallest profit. The explanation might be quite simple: it is customary and acceptable for each purveyor to retain a certain percentage of the selling price as profit; hence, since value is cumulatively added as one moves down the amount, the absolute value of the profit will inevitably increase. In many cases the percentage actually increases as well, on the grounds that demand from the fickle consumer fluctuates, and a high percentage profit compensates for the high risk of being left with unsold stock. As one moves upstream, these fluctuations are dampened and hence the percentage diminishes.

market applications, risks can be diversified, and research and development costs can be amortized across separate applications. The varied markets are likely to a have a corresponding variety of stages of maturity, hence providing revenue opportunities in the short-, medium- and long-term. As commercialization develops, progress in the different applications can be compared, allowing more objective assessments of performance than in the case of a single application; and the breadth and scope of opportunity might attract more investment than otherwise.[5]

9.2 THE RADICAL NATURE OF NANOTECHNOLOGY

But nanotechnology is above all a radical, disruptive technology whose adoption implies discontinuity with the past. In other words, we anticipate a qualitative difference between it and preceding technologies. In some cases, this implies a wholly new product; and at the other extreme an initially quantitative difference (progressive miniaturization) may ultimately become qualitative. While a generic technology has breadth, a radical technology has depth, since changes, notably redesign, might be needed all the way down the supply chain to the consumer; they affect the whole of the supply chain, whereas an incremental technology typically only affects its immediate surroundings. Insofar as the very definition of nanotechnology includes words such as "novel" and "unique" (see Section 1.6), "true" nanotechnology can scarcely be called anything but radical, otherwise it would not be nanotechnology.

The costs of commercialization are correspondingly very high. Redesign at a downstream position is expensive enough, but if it is required all the way, the costs of the introduction might be prohibitive. Furthermore, the more radical the technology, the greater the uncertainty in predicting the market for the product. High uncertainty is equivalent to high financial risk, and the cost of procuring the finance is correspondingly high. "Cost" might mean simply that a high rate of interest is payable on borrowings, or it might mean that capital is difficult to come by at all. This is in stark contrast to an incremental technology, for which the (much smaller) amount of capital required should be straightforward to procure, because the return on the investment should be highly predictable.

In addition, the more radical the innovation, the more likely it is that other innovations will have had to be developed in parallel to enable the one under consideration to be exploited. If these others are also radical, then

[5] E. Shane, *Academic Entrepreneurship*. Cheltenham: Edward Elgar (2004).

maybe there will be some synergies since comprehensive redesign is anyway required even for one. There may also be regulatory issues, but at present nanotechnology occupies a rather favorable situation, because there is a general consensus among the state bureaucracies which manage regulation that nanoparticulate X, where X is a well-known commercially available chemical, is covered by existing regulations governing the use of X in general. This situation stands in sharp contrast to the bodies (such as the FDA in the USA) entrusted with granting the *nihil obstat* to new medicinal drugs, which following the thalidomide and other scandals have become extremely conservative. Things are, however, likely to change, because one of the few clearly articulated recommendations of the influential Royal Society of London–Royal Academy of Engineering report on nanotechnology was that the biological effects of nanoparticles required more careful study before allowing their widespread introduction into the supply chain.[6]

The implications go even further, because an existing firm's competences may be wholly inadequate to deal with the novelty. Hence the infrastructure required to handle it includes the availability of new staff qualified for the technology, or the possibility of new training for existing staff.[7]

Nanotechnology is clearly both radical and universal. This combination is in itself unusual, and justifies the need to treat nanotechnology separately from other technically-based sectors of the economy.

9.3 FINANCING NANOTECHNOLOGY

Figure 9.2 summarizes the overall path of value creation by a nanotechnology company. We need only consider the two most typical types of nanotechnology company: (1) the very large company that is well able to undertake the developments using internal resources; and (2) the very small university spin-out company that in its own special field may have better intellectual resources than the large company, but which is cash-strapped. Examples of

[6] *Nanoscience and Nanotechnologies: Opportunities and Uncertainties*. London (2004). This conclusion created a considerable stir and triggered a flurry of government-sponsored research projects. Nevertheless, given the considerable literature that already existed on the harmful effects of small particles (e.g., P.A. Revell, The biological effects of nanoparticles. *Nanotechnol. Perceptions* 2 (2006) 283–298 and the many references therein), and the already widespread knowledge of the extreme toxicity of long asbestos fibers, the sudden impact of that report is somewhat mystifying.

[7] A particularly attractive mode of training is the various courses, typically ranging from a few intensive days dealing with a particular facet through a 1-year full-time M.Sc. to a collaborative 3-year Ph.D., offered by postgraduate institutes of technology such as Cranfield University in the UK.

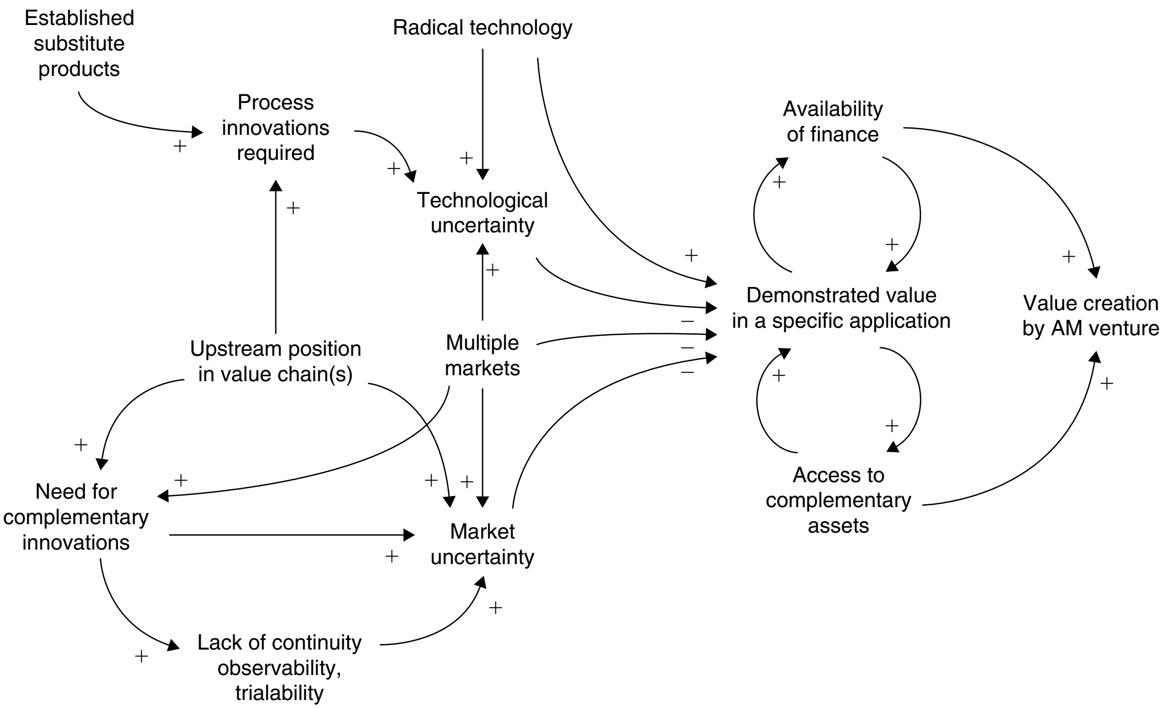


FIGURE 9.2 *Diagram of immediate effects for a nanotechnology company. "+" indicates that the factor causes an increase and "−" that it causes a decrease. Reproduced from S. Lubik and E. Garnsey, Commercializing nanotechnology innovations from university spin-out companies.* Nanotechnol. Perceptions *4 (2008) 225–238 with permission of Collegium Basilea.*

(1) include IBM (e.g., the "Millipede" mass data storage technology)[8] and Hewlett-Packard ("Atomic Resolution Storage" (ARS) and medical nanobots). Examples of (2) are given in Section 9.7.

We have already mentioned Thomas Alva Edison's "1% inspiration, 99% perspiration" dictum. If the research work needed to establish proof of principle costs one monetary unit, then the development costs to make a working prototype are typically 10 units, and the costs of innovation—introducing a commercial product—are 100 units. The last figure is conservative. An actual example is DuPont's introduction of Kevlar fiber: laboratory research cost $6 million, pilot plant development cost $32 million, commercial plant construction cost approximately $300 million, and marketing, sales and

[8]See S. de Haan, NEMS—emerging products and applications of nanoelectromechanical systems. *Nanotechnol. Perceptions* 2 (2006) 267–275.

distribution cost $150 million.[9] Moreover, commercial development is typically lengthy. It took about 17 years for Kevlar to reach 50% peak annual sales volume, which was in fact rather fast in comparison with other similar products (31 years for Teflon, 34 years for carbon fibers, and 37 years for polypropylene).[10] Hence immense sources of capital are necessary; even a large firm may balk at the cost.

Three main sources of capital are available: (i) internal funds of the company; (ii) private investors (typically venture capitalists and angel investors); and (iii) government funds. Generally speaking, (i) is only an option for very large firms, and even they seem to prefer to reserve their cash for acquiring small companies with desirable know-how, rather than developing it themselves. For various reasons connected with problems of internal organization and its evolution, large-company research is often (but not, of course, always) inefficient; the problem is that all firms, as they grow, inevitably also proceed along the road to injelititis.[11] Option (iii) is fraught with difficulties. The establishment of extensive state programs to support nanotechnology research and development is presumably based on the premise that nanotechnology is something emerging from fundamental science, implying that there is insufficient interest from existing industry willing to lavish funds upon its development, and still less because of its potentially disruptive nature. However, this government largesse might actually hinder development. It has long been a criticism of the European Union "Framework" research and technical development programs that they actually hinder innovation in European industry.[12] Generic weaknesses of government funding programs are: excessive bureaucracy, which not only saps a significant proportion of the available funds, but also involves much unpaid work (peer review) by working scientists, inevitably taking time away from their own research; excessive interference in the thematic direction of the work supported, which almost inevitably leads in the wrong direction, since by definition the officials administering the funds have left the world of active research, hence are removed from the cutting edge, nor are they embedded in the world of industrial exigencies; an excessively leisurely timetable of deciding which work to support—12 months is probably a good estimate of the average time that elapses *after* submitting a proposal before the final decision is made by the research

[9]E. Maine and E. Garnsey, Commercializing generic technology. *Res. Policy* 35 (2006) 375–393.

[10]Maine and Garnsey, loc. cit.

[11]C.N. Parkinson, *Parkinson's Law*, pp. 86 ff. Harmondsworth: Penguin Books (1965).

[12]House of Lords Select Committee on the European Communities, Session 1993–94, 12th Report, Financial Control and Fraud in the Community (HL paper 75). London: HMSO (1994).

council, and to this should be added the time taken to prepare the proposal (9 months would be a reasonable estimate), in which an extraordinary level of detail about the proposed work must typically be supplied (to the extent that, in reality, some of the work must already be done in advance in order to be able to provide the requested detail), and further months elapse after approval before the work can actually begin, occupied in recruiting staff and ordering equipment (6 months would be typical). Operating therefore on a timescale of two or more years between having the idea and actually beginning practical work on testing it, it is little wonder that research council projects tend to be repositories for incremental, even pedestrian work, the main benefit of which are the accompanying so-called overhead payments that help to maintain the central facilities of the proposer's university. We are therefore left with (ii) as the main or most desirable source of funding for truly innovative spin-out companies. Even the most angelic of investors seeks an eventual return on his capital, however. Very important elements of Figure 9.2 are the two small loops on the right hand side of the diagram. They are expanded in Figure 9.3.

An attractive route, with several successful examples, is for the nanotechnology company to enter into a close partnership with a company

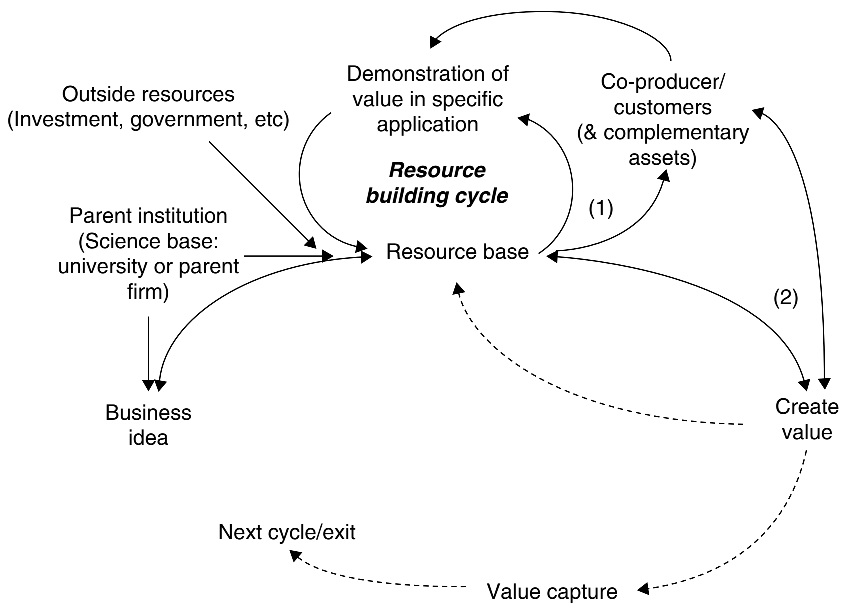


FIGURE 9.3 *Diagram of immediate effects for a nanotechnology company, from a slightly different viewpoint compared with Figure 9.2, here focusing on the resource-building cycle. Reproduced from S. Lubik and E. Garnsey, Commercializing nanotechnology innovations from university spin-out companies.* Nanotechnol. Perceptions *4 (2008) 225–238 with permission of Collegium Basilea.*

established in the application area for which the new technology is appropriate. The pooling of complementary interests seems to create a powerful motivation to succeed in the market (see also Section 9.7).

9.4 GOVERNMENT FUNDING

The biggest current investment in nanotechnology comes from the public domain. It is interesting to compare government funding for nanotechnology (Table 9.1). A bald comparison of the absolute values is less revealing than key ratios: funding per capita (F/N) is indicative of the general level of public interest in pursuing the new technology, and the fraction of GDP spent on nanotechnology (F/G) indicates the seriousness of the intention. Despite the weakness of not knowing exactly how F has been determined, several general conclusions are interesting. Japan is clearly the leader, both in interest and intention. The USA follows in interest, and then France and Germany. France's strong interest is in accord with its current image as a powerhouse of high technology (the output of which includes the Ariane spacecraft, the Airbus and high-speed trains). Switzerland's interest is surprisingly low, given its past lead as a high-technology country (but see Table 9.2). But when it comes to "putting one's money where one's mouth is", only Japan does creditably well. Switzerland in particular could easily afford to double or triple its expenditure.[13] And, impressive as the USA's contribution looks relative to that of Brazil, for example, it is barely half of the sum allocated to "funds for communities to buy and rehabilitate foreclosed and vacant properties" (taken as a somewhat random example of part of the US federal stimulus plan promulgated in February 2009).

The lower half of the table presents a less encouraging picture. The very low level of activity in Argentina, which formerly had a relatively strong science sector, is indicative of the success of the International Monetary Fund (IMF) in insisting on a substantial downscaling of that sector as part of its economic recovery prescription. In our modern high-technology era, this is simply not how to create the basis for a strong future economy, and reflects the archaic views that still dominate the IMF. Brazil's performance is also disappointing, given its aspirations to become one of the new forces in the world economy. Among these countries, only Malaysia reveals itself as a true Asian "tiger" able to take its place in the world nanotechnology community.

[13] Incidentally, EU member states and countries associated with their research and development program receive an additional 40% of the stated F.

Table 9.1 Government Funding F (2004) for Nanotechnology Research and Development, Together with Population N and GDP G.[a]

Country	$F^b/10^6$€	$N^c/10^6$	$G^c/10^{12}$€	F/N	F/G (%)
France	223.9	61	1.43	3.67	0.016
Germany	293.1	82	1.86	3.57	0.016
Italy	60.0	59	1.18	1.02	0.0051
Japan	750	128	3.03	5.86	0.025
Switzerland	18.5	7.4	0.25	2.50	0.0074
UK	133	60	1.49	2.22	0.0089
USA	1243.3	298	8.27	4.17	0.015
Argentina	0.4	38	0.12	0.010	0.00033
Brazil	5.8	188	0.59	0.031	0.0010
Malaysia	3.8	26	0.09	0.15	0.0042
Mexico	10	106	0.51	0.094	0.0020
South Africa	1.9	49	0.16	0.039	0.0012
Thailand	4.2	62	0.12	0.068	0.0035

[a] *The upper portion contains selected Category I countries (Section 5.6).*
[b] *Source: Unit G4 (Nanosciences and Nanotechnologies), Research Directorate General, European Commission.*
[c] *Source:* Global Market Information Database. *Euromonitor International (2008).*

Table 9.2 Number of papers (P, 2005) and P/F (Table 9.1).

Country	P^a	$(P/F)/10^{-6}$€$^{-1}$
France	3994	18
Germany	5665	19
Italy	2297	38
Japan	7971	11
Switzerland	1009	55
UK	3335	25
USA	14750	12

[a] *R.N. Kostoff et al., The growth of nanotechnology literature.* Nanotechnol. Perceptions *2 (2006) 229–247.*

What of the effectiveness of the expenditure? If the main outcome of this kind of funding is papers published in academic journals, the ratio (P/F) of the number of papers P to funding is a measure of effectiveness (Table 9.2). By this measure the big spenders (Japan and the USA) appear to be less effective, and Switzerland's spending appears to be highly effective. This simple calculation of course takes no account of the existing infrastructure (the integral of past expenditure), nor to the extent to which funds result in products rather than papers. One nanotechnology paper costs about 5600 Euro in France, which seems remarkably cheap, suggesting that expenditure on nanotechnology is underestimated (even if European Union funds are taken into account, it still amounts to less than 8000 €).

In most countries, this public support for research and development covers the entire range of the technology, with little regard for ultimate utility. What is lacking is a proper assessment of which sectors might best benefit from nanotechnology at its current level. Such an assessment would make it possible to appraise the utility of current research, indicating into which sectors investment should be directed towards research, development and innovation, and hence provide a better basis for public investment decisions, as well as being useful for private investors interested in backing nanotechnology-based industry.

It may be presumed that private (industrial) funding for nanotechnology is more directed. Incidentally, this is about double the level of state funding in Japan, about equal to it in the USA, and about half of the European level (cf. Section 5.6.1). The contrast between Europe and Japan is therefore especially marked.

9.5 INTELLECTUAL NEEDS

As well as material capital, the innovating company also has significant intellectual needs. It is perhaps important to emphasize the depth of those needs. Although the scientific literature today is comprehensive and almost universally accessible, simply buying and reading all the journals would not unlock the key to new technology: one needs to be an active player in the field just to understand the literature, and one needs to be an active contributor to establish credibility and allow one to participate in meaningful discussions with the protagonists.

Science, technology and innovation all require curiosity, imagination, creativity, an adventurous spirit and openness to new things. Progress in advanced science and technology requires years of prior study in order to reach the open frontier, and to perceive unexplored zones beyond which the frontier has already passed. Governments mindful that innovation is the wellspring of future wealth do their best to foster an environment conducive to the advance of knowledge. Hence it is not surprising that the state typically plays a leading rôle in the establishment of research institutes and universities.

Nevertheless, in this "soft" area of human endeavor it is easy for things to go awry. The linear Baconian model has recently recaptured the interest of governments, who wish to expand the controlled legal framework supposedly fostering commercially successful innovations (such as the system of granting patents) by extending their control upstream to the work of scientists. Even the Soviet Union under Stalin, a world steeped in state control, realized that extending it this far was inimical to the success of enterprises (such as the development of atomic weapons) that were considered to be vital to the survival of the state.

This lesson seems to have been forgotten in recent decades. The system of allocating blocks of funds to universities every 5 years or so and letting them decide on their research priorities has been replaced by an apparatus of research councils to which scientists must propose projects, for which funds will be allocated if they are approved. Hence, the ultimate decision on what is important to investigate is taken away from the scientists themselves and put in the hands of bureaucrats (some of whom, indeed, are themselves former scientists, but obviously cannot maintain an acute knowledge of the cutting edge of knowledge). To any bureaucrat, especially one acting in the public interest, the file becomes the ultimate object of importance (for, as C.N. Parkinson points out,[14] there may subsequently be an inquiry about a decision, and the bureaucrat will be called upon to justify it). Therefore great weight is placed on clearly measurable outcomes ("deliverables") of the research, which should be described in great detail in the proposal, so that even an accountant would have no difficulty at the end of the project in ascertaining whether they had indeed been delivered. The most common criticism of proposals by reviewers seems to be that they lack sufficient detail, a criticism that is frequently fatal to the chances of the work being funded. Naturally, such an attitude does nothing to encourage adventurous, speculative thinking. Even de Gaulle's Centre National de la Recherche Scientifique (CNRS), modeled on the Soviet system of Academy institutes, and offering a place where scientists can work relatively free of constraints, is now in danger of receiving a final, mortal blow (in fact, for years the spirit of the endeavor had not been respected; the resources available to scientists not associated with any particular project had become so minimal that they were only suitable for theoretical work requiring neither assistants nor apparatus).

[14]C.N. Parkinson, *In-Laws and Outlaws*, pp. 134–135. London: John Murray (1964).

One can hardly imagine that such a system could have been introduced, despite these generally recognized weaknesses, were there not failings in the alternative system. Indeed we must recognize that the system of allocating a block grant to an institute only works under conditions of "benign dictatorship". Outstanding directors of institutes (such as the late A.M. Prokhorov, former director of the General Physics Institute of the USSR Academy of Sciences)[15] impartially allocated the available funds to good science—"good" implying both intellectually challenging and strategically significant. Unfortunately, the temptations to partiality are all too frequently succumbed to, and the results from that system are then usually disastrous. A possible alternative is democracy: the faculty of science receives a block grant, and the members of the faculty must agree how to divide it among themselves. It is perhaps an inevitable reflection of human nature that this process almost invariably degenerates into squabbling. Besides, the democratic rule of simple majority would ensure that the largest blocs appropriated all the funds. Hence in order for democracy to yield satisfactory results, it has to be accompanied by so many checks and balances it ends up being unworkably cumbersome.

Is there a practical solution? Benign dictatorship would appear to yield the best results, but depends on having an inerrant procedure for choosing the dictator; in the absence of such a procedure (and there appear to be none that are socially acceptable today) this way has to be abandoned. The opposite extreme is to give individual scientists a grant according to their academic rank and track record (measured, for example, by publications). This system has a great deal to commend it (and, encouragingly, appears to be what the Research Directorate of the European Commission is aiming at with its recently introduced European Research Council awarding research grants to individual scientists[16]). The only weakness is that, almost inevitably, scientists work in institutes, with all that implies in terms of possibilities for partiality in the allocation of rooms and other institutional resources by those in charge of the administration, who are not necessarily involved in the actual research work.

---

[15]The author spent some weeks in his institute in 1991. For a published account, see I.A. Shcherbakov, 25 Years of A.M. Prokhorov General Physics Institute, RAS. *Quantum Electronics* 37 (2007) 895–896.

[16]Unfortunately the procedure for applying for these grants is unacceptably bureaucratic and thus vitiates what would otherwise be the benefit of the scheme; furthermore the success rate in the first round was only a few percent, implying an unacceptable level of wasted effort in applying for the grants and evaluating them. The main mistake seems to have been that the eligibility criteria were set too leniently. This would also account for the low success rate. Ideally the criteria should be such that every applicant fulfilling them is successful.

9.5.1 Company–University Collaboration

The greatest need seems to be to better align companies with university researchers. Many universities now have technology transfer offices, which seem to think that great efforts are needed to get scientists interested in industrial problems. In reality, however, rarely are such efforts required—a majority of devoted scientists would agree with A.M. Prokhorov about the impossibility of separating basic research from applied (indeed, these very expressions are really superfluous). On the contrary, university scientists are usually highly interested in working with industrial colleagues; it is usually the institutional environment that hinders them from doing so more effectively. Somehow an intermediate path needs to be found between the consultancy (which typically is too detached and far less effective for the company than access to the available expertise would suggest should be the case) and the leave of absence of a company scientist spent in a university department (which seems to rapidly detach the researcher from "real-life" problems), and the full-time company researcher, who in a small company may be too preoccupied by daily problems that need urgent attention, or who in a larger company might be caught up in a ponderous bureaucracy. Furthermore, companies are typically so reticent about their real problems that it is hard for the university scientist to make any useful contribution to solving them. One seemingly successful model, now being tried in a few places, is to appoint company "researchers in residence" in university departments—they become effectively members of the department, but would be expected to divide their time roughly equally between company and university. Such schemes might be more effective if there were a reciprocal number of residencies of university researchers in the company. Any expenses associated with these exchanges should be borne by the company, since it is they who will be able to gain material profit from them; misunderstanding over this matter is sometimes a stumbling block. It is a matter of profound regret that the current obsession with gaining revenue from the intellectual capital of universities has poisoned relationships between them and the rest of the world. The free exchange of ideas is thereby rendered impossible. In effect, the university becomes simply another company. If the university is publicly funded, then it seems right to expect that its intellectual capital should be freely available to the nation funding it. In practice, "nation" cannot be interpreted too literally; it would be contrary to the global spirit of our age to distinguish between nationals and foreigners—whether they be students or staff—and if they are roughly in balance, there should be no need to do so.

9.5.2 Clusters

Evidence for the importance of personal intellectual exchanges comes from the popularity, and success, of clusters of high-technology companies that have nucleated and grown, typically around important intellectual centers such as the original Silicon Valley in California, the Cambridges of England and Massachusetts, and the Rhône-Alpes region of south-eastern France. The additional feature of importance is the availability of centralized fabrication and metrology facilities, the use of which by any individual member of the cluster would scarcely be at a level sufficient to justify the expense of installing and maintaining them.

9.6 THE COST OF NANOTECHNOLOGY

Many downstream manufacturers are attracted by what they hear about nanotechnology and became interested in incorporating upstream nanomaterials or devices into their products. In order to make a business decision, they need to know the cost of such incorporation. This will depend on the choice of material or device, how much further research and development will be necessary—there may be little or no experience with similar applications to draw upon—the degree of redesign, the manufacturing process, and any special requirements regarding end-of-life disposal. For substitution and incremental improvement, existing methodologies of cost engineering may be adequate.

9.7 COMPANIES

Wilkinson has identified four generic business models (Figure 9.4), all beginning at the most upstream end of the supply chain, but extending progressively downstream.

The following subsections consist of case studies of four small or medium nanotechnology companies.[17]

9.7.1 Hyperion

Hyperion was founded in 1981 to develop carbon filament-based advanced materials for a variety of applications. Located in Cambridge, MA, they developed their own process for the fabrication of multiwalled carbon nanotubes

[17] Information about Hyperion and CDT are from Maine and Garnsey (loc. cit.) and about Q-Flo and Owlstone from Lubik and Garnsey (loc. cit.).

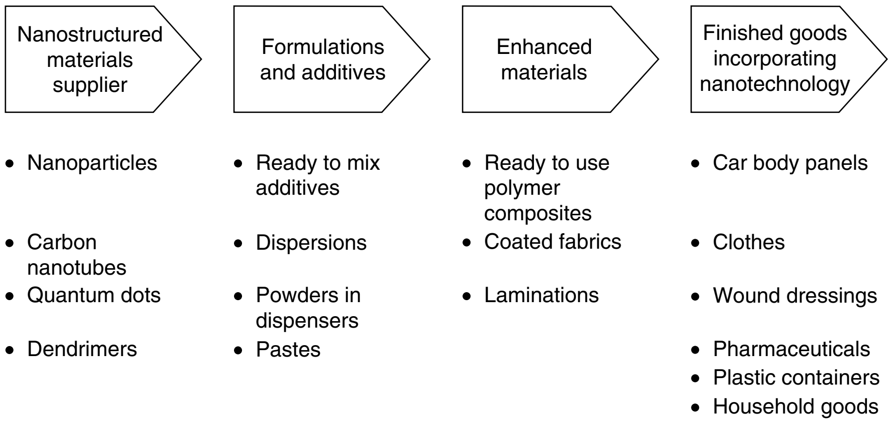


FIGURE 9.4 *Generic business models for nanomaterial suppliers. Model A (e.g., Thomas Swan) produces only nanostructured materials. Model B (e.g., Zyvex) produces nanostructured materials and formulates additives. Model C (e.g., Nucryst) produces nanostructured materials, formulates additives and supplies enhanced materials. Model D (e.g., Uniqema Paint) produces nanostructured materials, formulates additives, makes enhanced materials and finished goods incorporating those materials. Reproduced from J.M. Wilkinson, Nanotechnology: new technology but old business models?* Nanotechnol. Perceptions *2 (2006) 277–281 with permission of Collegium Basilea.*

(MWCTs), their key intermediate product. By 1989 they could make them in-house on a fairly large scale and to a high level of purity. The problem then was to choose a downstream application. In the absence of prototypes, they widely advertised their upstream product with the aim of attracting a partner. Their first was a company that had developed a competitive polymer automotive fuel line as a substitute for the existing steel technology, but still needed to make the polymer electrically conducting (to minimize the risk of static electricity accumulating and sparking, possibly triggering fuel ignition). Dispersing MWCTs in the polymer looked capable of achieving this, and by 1992 Hyperion had developed a process to disperse their material into the polymer resin, meanwhile also further upscaling their process to reach the tonne level. Related applications followed from the mid-1990s onward—conductive polymer automotive mirror casings and bumpers, which could be electrostatically painted along with the steel parts of the bodywork and hence fully integrated into existing assembly lines. The company moved slightly downstream by starting to compound its MWCTs with resin in-house. Efforts to diversify into structural aerospace parts did not succeed in demonstrating adequate enhanced value to enter the market, but the company did successfully break into internal components of consumer electronics devices. Research into supercapacitors and catalysts was pursued with the help of government funding. To date, Hyperion have filed over 100 patents. The product line remains based on carbon nanotubes dispersed in resin to make it conductive. They have 35 employees and annual revenues are $20–50 million.

9.7.2 CDT

Cambridge Display Technology (CDT) was founded in 1992 as a spin-out from Cambridge University (UK), where during the preceding decade polymer transistors and light-emitting polymers had been invented (the key polymer electroluminescence patent was filed in 1989). CDT's objective was to manufacture products for flat-panel displays, including back lighting for liquid crystal displays. It soon became apparent that a small company could have little impact on its own, hence it abandoned in-house manufacturing and sought licensing arrangements with big players such as Philips and Hoechst (finalized by 1997), and in 1998 embarked on a joint venture with Seiko-Epson Corp. (SEC) to develop a video display. Other strategic allies included Bayer, Sumitomo, Hewlett-Packard and Samsung. CDT continued patenting (end-products developed with allies were included in the portfolio), but R&D costs remained huge, far exceeding license revenues. In 2000 the company was acquired by two New York-based private equity funds. This caused some turbulence: the departure of the energetic CEO (since 1996) and the decision of the founder, Richard Friend, to form a new company on which he focused his continuing research efforts. CDT then decided to recommence manufacturing and released an organic light-emitting diode (OLED) shaver display in 2002, but an attempt to extend this to the far more significant cellular telephone market came to nought and the commercial-scale production line was closed in 2003, retaining only the ability to make prototypes. The company thus reverted to the licensing mode. By 2003 it held 140 patents, generating $13–14 million per annum, compared with annual running costs of *ca* $10 million (the company had 150 employees at that time). The strategy of getting the technology into small mobile displays in the short term, and aiming at the huge flat-panel market (estimated as $30 milliard annually) in the medium term has remained attractive to investors despite the ups and downs, and the company went public on the NASDAQ in 2004.

9.7.3 Q-Flo

Q-Flo was founded in 2004, also as a spin-out from Cambridge University (UK), in order to commercialize a novel process for making carbon nanotube (CNT) fiber (at a cost potentially one-fifth that of current industrial CNT fiber) as a very strong material in the form of a textile or a film. Favorable electrical properties are reflected in envisaged applications in supercapacitors and batteries. Other opportunities include bulletproof body armor, shatterproof concrete, ultrastrong rope, tires and antennae. However, the company is too small to be able to afford to make prototypes for value-demonstration purposes, but in their absence cannot attract the investment needed to be able to afford to make them. The key resource-building cycle (see Figure 9.3) cannot therefore start turning. Because the company is so small, none of its current seven employees work full-time for Q-Flo, which also limits the intrinsic dynamism of the available human resources.

9.7.4 Owlstone

Owlstone was also founded in 2004 as a spin-out from Cambridge University (UK). Its technology is nanoscale manufacturing to produce a microelectromechanical system (MEMS) gas sensor, based on field-asymmetric ion mobility spectrometry (FAIMS). This generates a "fingerprint" for any gas or vapor entering the sensor, which is matched against a collection of standard fingerprints. The device is several orders of magnitude smaller than existing competitors and detection takes less than 1 second. The company's first investor was Advanced Nanotech, which acquired a majority interest, but after other companies owned by Advanced Nanotech failed to reach expectations, Owlstone took over its erstwhile owner.

The original aim was to make the FAIMS chip and sell it to sensor suppliers, leaving it to them to incorporate it into their products. However, the uniqueness of the device meant that outsourcing production of the chip alone would incur high development overheads with general foundries in any case, hence it was decided to aim instead to produce the finished downstream sensor. With the help of SBIR funds (fortunately for this purpose Advanced Nanotech was registered in both the UK and the USA) the first production model sensor was launched in 2006. Further products were subsequently launched with partners already in the market. Reveunue in 2008 is expected to exceed $2 million.

9.7.5 Analysis

The above case studies, of one indubitably very successful and one perhaps haltingly successful company in each category of medium and small company, shows that key ingredients of success are:

- Focusing on a single application
- Launching as downstream a product as possible

- Making a prototype to demonstrate value
- Having dedicated staff.

Spin-out companies are often tempted to economize by continuing to use university facilities and part-time staff, but this seems to ensure that the necessary pressures to succeed never surmount what might be a critical threshold. Doubtless location in a thriving center of high technology is important (but even this might be becoming less so in the age of the internet). Given the novelty of the upstream product, persuading downstream companies to incorporate it into their final product, with all the attendant expense of redesign (even if the upstream product is merely substitutional) may be even more expensive than pursuing the downstream product in-house. Here, a rational basis for estimating the costs is important (cf. Section 9.6). And even if the downstream client is a partner, it may still be difficult to obtain accurate information about key attributes. Finally, as already mentioned it is known that the further upstream one is positioned, the harder it is to capture value from any specific application, which diminishes the attraction for investors.

There is of course an element of luck in finding investors. A social setting (which might be as unpretentious as a college bar) in which would-be investors and technologists mix informally is probably a crucial ingredient. The fiscal environment is also crucial. Despite globalization, this is still a distinctively national characteristic. Given the outstanding success record of the SBIR grant scheme in the USA, it is astonishing that other countries have not sought to adopt it (Japan has its own very successful mechanisms, but unlike the situation in the USA and Europe as a whole, they are geared towards a far more socially homogeneous environment, as foreigners working in Japan cannot fail to notice). The situation within the European Union is especially depressing, marked as it is by ponderous, highly bureaucratic mechanisms and an overall level of funding running at about one-third of the equivalent in the USA or Japan. Switzerland manages to do better, but could actually easily exceed the (per-capita) effort of the USA and Japan (given that it has the highest per-capita income in the world). It is particularly regrettable that it has failed to maintain its erstwhile lead as a high-technology exporting nation, choosing instead to squander hundreds of milliards of francs on dubious international investments that have now (i.e., in 2008 and 2009) been revealed as worthless. One can only wonder what might have been achieved had these same monies been spent instead on building up world-leading nanotechnology research and development facilities.

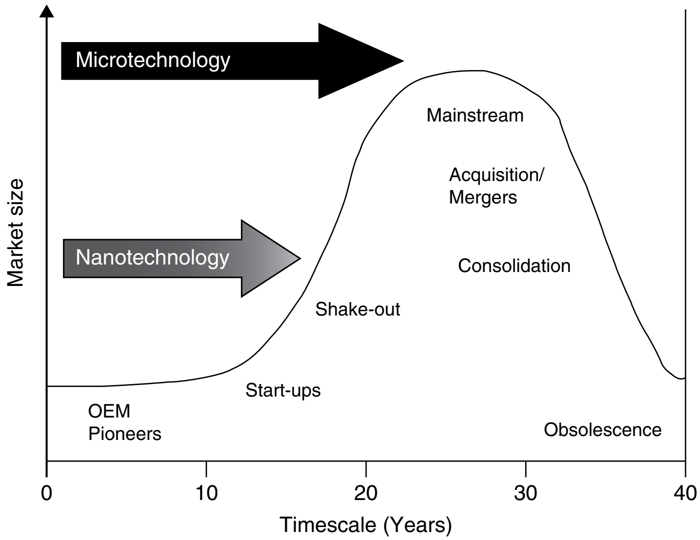


FIGURE 9.5 *Generic model proposed for the temporal evolution of nanotechnology companies (originally developed by Prismark Associates, New York for the printed circuit board industry; it also seems to fit the evolution of the microtechnology industry). Reproduced from J.M. Wilkinson, Nanotechnology: new technology but old business models?* Nanotechnol. Perceptions *2 (2006) 277–281 with permission of Collegium Basilea.*

9.8 TEMPORAL EVOLUTION

Figure 9.5 is based on a model describing the printed circuit board industry, and so far has fitted the observed course of events in microsystems. It can certainly be considered as a model for nanotechnology as far as its substitutional and incremental aspects are concerned. Insofar as it is universal and radical, however, prediction becomes very difficult.

If one examines in more detail the early stages, it appears that there might be a gap between early adopters of an innovation and development of the mainstream market.[18] The very early market peaks and then declines, followed by the mainstream development as shown in Figure 9.5. This feature underlines the importance of patient investors in new high-technology companies.

[18]G.A. Moore, *Crossing the Chasm*. New York: Harper Business (1991).

9.9 PATENTS AND STANDARDS

The system of patents—monopoly privileges accorded to inventors enshrined in law—seems to have begun in the 15th century. They were known within the glassmaking community of Venice, but the oldest continuous patenting tradition in the world began in England with the patent granted in 1449 to John of Utynam for making stained glass. The circumstances were perhaps rather special—John had come to England from Flanders to make the windows for Eton College, whose patron was King Henry VI and whose seal also validated the patent. Hence, in this particular case the monopoly might be viewed as a royal reward to a distinguished craftsman. Although patents offer an obvious advantage to their holder, their benefit to the nation as a whole is doubtful. Indeed, it seems anomalous that England, a traditional champion of free trade and competition, should effectively have pioneered the modern patent system.[19] The contemporary argument is that patents provide an incentive to the inventor. This is clearly specious; evidence overwhelmingly shows that inventors invent regardless of "incentives". It would be more accurate to state that they provide an incentive to the innovator. A guaranteed monopoly of supply does not, of course, guarantee that there will be a demand for the new product, but given the difficulty of predicting such demand, one could justify such a legal guarantee. It must be weighed against the effects patents might have in stifling innovation. A large company may buy up patents for rival products held by smaller companies in order to further entrench its monopoly (on the other hand, additional invention might be stimulated by companies seeking to evade a patent). Given the essentially irremediable absence of controlled "experiments" in the field of political economy, it is very difficult to ascertain whether patents have a net positive or a net negative benefit on innovation. It is, however, noteworthy that the patent laws were anathema to Isambard K. Brunel, one of the most brilliant engineers of the Victorian era. A principal argument of his against them was that they could be, and were, exploited by taking out patents of principle, thereby stifling actual innovation.[20]

Technical Committee (TC) 229 of the International Standards Organization (ISO) is currently developing definitions and terminology appropriate to this early stage in the field of nanotechnology (Section 1.6).

---

[19]The 1624 Statute of Monopolies placed clear limitations on the extent of monopoly, however, not only temporal (a maximum of 14 years), but also stipulated that the public interest must be respected.

[20]L.T.C. Rolt, *Isambard Kingdom Brunel*, p. 217. London: Longmans, Green & Co. (1957).

CHAPTER 10

Assessing Demand for Nanotechnology

CHAPTER CONTENTS

When a decision has to be made regarding the viability of an investment in a nanotechnology venture, it might not be easy to predict costs. In more traditional industries, these costs are generally well determined. Given nanotechnology's closeness to the fundamental science, however, it is quite likely that unforeseen difficulties may arise during the development of a product for which proof of principle has been demonstrated. By the same token, difficulties may have been anticipated on the basis of present knowledge, but subsequent discoveries could enable a significant shortcut to be taken. On balance, these positive and negative factors might balance out; it seems, however, to be part of human nature to minimize the costs of

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

undertaking a future venture even when the desire to undertake it is high.[1] There is a strong element of human psychology here.

The development, innovation and marketing costs determine the amount of investment required. The return on investment arises through sales of the product (the market); that is, it depends on demand, and the farther downstream the product, the more fickle and unpredictable the consumer.

A starting point for assessing these costs would appear to be the elasticities of supply and demand. Extensive compilations have been made in the past,[2] and (updated) might be useful for products of substitution and innovation. Of course, this would represent only a very rudimentary assessment, because all the cross-elasticities would also have to be taken into account. Furthermore, the concept has not been adequately developed to take quality into account (which is sometimes difficult to quantify).

A perpetual difficulty is that only very rarely can the impact of the introduction of a new product be compared with its non-introduction. Change may have occurred in any case and even the most carefully constructed models will usually fail to take into account the intrinsic nonlinearities of the system.

10.1 PRODUCTS OF SUBSTITUTION

These represent the lowest level of innovation. The consumer may not even be aware of any change; the main advantage is to the producer (lower manufacturing costs through a simplified process or design), and possibly to the environment (a smaller burden, due to the use of a smaller quantity of raw materials, hence less weight to transport and less waste to ultimately be disposed of). In this case the anticipated market is the same as the present market; if there is an increasing or decreasing trend it may be considered to continue (e.g., exponential, linear or logarithmic or a combination of all three, i.e. logistic) in the same fashion.

If the innovation reduces production costs, the enhanced profitability may attract other manufacturers (assuming that the innovation is not protected by patent or secrecy), which would tend to depress the price in the long term.

---

[1] This state of affairs has led to the failure of many (geographical) exploratory expeditions. It is understandable, given the prudence (some would say meanness) of those from whom resources are being solicited, but is paradoxical because the success of the venture is thereby jeopardized by being undertaken with inadequate means. Failure might also decrease the chances of gathering support for the future expeditions of a similar nature.

[2] E.g., H.S. Houthakker and L.D. Taylor, *Consumer Demand in the United States: Analyses and Projections*. Cambridge, MA: Harvard University Press (1970).

10.2 INCREMENTALLY IMPROVED PRODUCTS

Examples are tennis rackets reinforced with carbon nanotubes, making them stronger for the same weight. Very often this will make the product more expensive, so elasticity of demand is a significant factor. On the other hand, it is doubtful whether the laborious compilations of demand elasticity that have been made in the past are really useful. What degree of improvement ranks as incremental? It might not take very much for the product to be considered as essentially new. Furthermore, how is one to quantify quality? If a laptop computer originally weighing 2 kg can be made to weigh only 1.5 kg with the same performance, different users will value the difference in different ways.

10.3 RADICALLY NEW PRODUCTS

These are goods that, in their qualitative nature, did not exist before. Of course, it is perhaps impossible for something to be totally new. Polaroid "instant" film (that could be developed and made visible seconds after taking a snapshot) was certainly a radical concept, but on the other hand it was still based on a silver halide emulsion and the mode of actually snapping the shot was the same, essentially, as with a Kodak box camera.

The future is in this case very difficult to predict, and an ad hoc model (Section 10.4) is probably needed if any serious attempt at planning is to be made.

10.4 MODELING

A decision whether to invest in a new technology will typically be made on the basis of anticipated returns. While in the case of incremental technology these returns can generally be estimated by simple extrapolation from the present situation, by definition for any radical (disruptive) technology there is no comparable basis from which to start. Hence one must have recourse to a model, and the reliability will depend upon the reasonableness of the assumptions made. Naturally as results start to come in from the implementation of the technology, one can compare the predictions of the model with reality, and adjust and refine the model. An example of this sort of approach is provided by cellular telephony: the model was that the market consists of the entire population of the Earth.

One of the problems of estimating the impact of nanotechnology tends to be the over-optimism of many forecasters. The "dotcom" bubble of 2000 is a classic example. Market forecasts for mobile phones had previously assumed that almost every adult in the world would buy one and it therefore seemed not too daring a leap to assume that they would subsequently want to upgrade to the 3G technology. Although the take-up was significant it was not in line with the forecast growth of the industry—with all too obvious consequences. Nanotechnology market forecasting is still suffering from the same kind of problem; for example, will every young adult in the requisite socio-economic group buy an i-Pod capable of showing video on a postage stamp-sized screen? The next section offers a more sober way to assess market volume.

10.5 JUDGING INNOVATION VALUE

The life quality index Q, to be introduced in more detail in Section 14.3, is defined as

$$Q = G^q X_{\mathrm{d}} \tag{10.1}$$

where G is average work-derived annual earnings, q is optimized work–life balance (here defined as $q = w/(1 - w)$, where w is the optimized average fraction of time spent working, and considered as a stable constant with a value $q = 1/7$ for industrialized countries), and X_{d} is discounted life expectancy. From the manufacturer's viewpoint, any substitutional or incremental innovation that allows specifications to be maintained or surpassed without increasing cost is attractive. But how will a prospective purchaser respond to an enhanced specification available for a premium price? Many such consumer products are now available, especially in Japan.[3] Theoretically, if the innovation allows a chore to be done faster, then its purchase should be attractive if the increase of Q due to the increase of q is more than balanced by the decrease of Q due to the diversion of some income into the more expensive product.

10.6 ANTICIPATING BENEFIT

In which sectors can real benefit from nanotechnology be anticipated? What is probably the most detailed analysis existing of the economic consequences of molecular manufacturing assumes blanket adoption in all fields, even food production.[4] Classes of commodity particularly well suited for productive

[3]Unfortunately in Europe there is still a strong tendency to buy the cheapest, regardless of quality, which of course militates against technological advance.

[4]R.A. Freitas, Jr, Economic impact of the personal nanofactory. *Nanotechnol. Perceptions* 2 (2006) 111–126.

nanosystems (PN) include those that are intrinsically very small (e.g., integrated electronic circuits) and those in which a high degree of customization significantly enhances the product (e.g., medicinal drugs). In many other cases (and bear in mind that even the most enthusiastic protagonists do not anticipate PNs to emerge in less than 10 years), there are no clear criteria for deciding where "intermediate nanotechnology" could make a worthwhile contribution.[5]

Any manufacturing activity has a variety of valid reasons for the degree of centralization and concentration most appropriate for any particular type of product and production. The actual degrees exhibited by different sectors at any given epoch result from multilevel historical processes of initiation and acquisition, as well as the spatial structure of the relevant distributions of skills, power, finance and suppliers. The inertia inherent in a factory building and the web of feeder industries that surround a major center mean that the actual situation may considerably diverge from a rational optimum.

The emergence of a radical new technology such as nanoscale production will lead to new pressures, and opportunities, for spatial redistribution of manufacturing, but responses will differ in different market sectors. They will have different relative advantages and disadvantages as a result of industry-specific changes to economies of scale, together with any natural and historic advantages that underlie the existing pattern of economic activities. But we should be attentive to the possibility that the whole concept of economies of scale will presumably become irrelevant with the advent of productive nanosystems, and will have intermediate degrees of irrelevance for intermediate stages in the development of nanotechnology.

---

[5]As far as nanotechnology is concerned, the task of deciding whether agile manufacturing is appropriate is made more difficult by the fact that many nanotechnology products are available only from what are essentially research laboratories, and the price at which they are offered for sale is rather arbitrary; in other words, there is no properly functioning market.

CHAPTER 11

Design of Nanotechnology Products

CHAPTER CONTENTS

It has already been stressed in Chapter 9 that one of the difficulties faced by suppliers of any upstream technology is that they must ensure that its use is already envisaged in the design of the downstream products that will incorporate the technology. Apart from fulfilling technical specifications, aesthetic design is furthermore one of the crucial factors determining the allure of almost any product (perhaps those destined for outer space are an exception), but especially a consumer product. In this chapter we look at some peculiar features associated with design of nanodevices, here defined as devices incorporating nanomaterials.

11.1 THE CHALLENGE OF VASTIFICATION

There is little point in making something very small if only a few of those things are required.[1] The interest in making a very large-scale integrated

[1] Devices for which accessibility is the principal consideration might still be worth making very small even if only few are required; e.g., for a mission to outer space.

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

circuit with nanoscale components is rooted in the possibility of making vast numbers in parallel. Thus, the diameter of the silicon wafers has grown from 4″ to 8″ to 12″ in only a few years.

Hence, although the most obvious consequence of nanotechnology is the creation of very small objects, an immediate corollary is that there must be a great many of these objects. If r is the relative device size and R the number of devices, then usefulness may require that $rR \sim 1$, implying the need for making 10^9 nanodevices at a stroke.[2] This corresponds to the number of components (with a minimum feature size of 45–65 nm) on a very large-scale integrated electronic processor or storage chip, for example. At present, all these components are explicitly designed and fabricated. But will this still be practicable if the number of components increases by a further two and more orders of magnitude?

11.2 ENHANCING TRADITIONAL DESIGN ROUTES

Regarding processor chips, which are presently the most vastified objects in the nano world, aspects requiring special attention are: power management, especially to control leakage; process variability, which may require a new conception of architectural features; and a systems-oriented approach, integrating functions and constraints, rather than considering the performance of individual transistors. Nevertheless, the basic framework remains the same.

Because it is not possible to give a clear affirmative answer to this previous question, alternative routes to the design and fabrication of such vast numbers are being explored. The human brain serves as an inspiration here. Its scale is far vaster than the integrated circuit: it has $\sim 10^{11}$ neurons, and each neuron has hundreds or thousands of connections to other neurons. So vast is this complexity there is insufficient information contained in our genes to specify all these interconnections. We may therefore infer that our genes specify an algorithm for generating them.[3]

In this spirit, evolutionary design principles may become essential for designing nanodevices. An example of an evolutionary design algorithm is shown in Figure 11.1. It might be initialized by a collection of existing designs, or guesses at possible new designs. Since new variety within the

---

[2]This is why vastification—the proliferation of numbers—almost always accompanies nanification.

[3]P. Érdi and Gy. Barna, Self-organizing mechanism for the formation of ordered neural mappings. *Biol. Cybernetics* 51 (1984) 93–101.

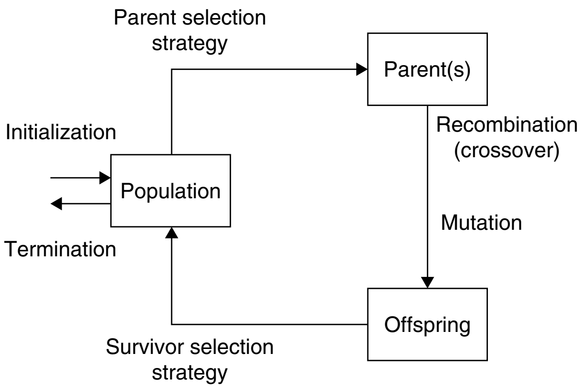


FIGURE 11.1 *An evolutionary design algorithm. All relevant design features are encoded in the genome (a very simple genome is for each gene to be a single digit binary value indicating absence (0) or presence (1) of a feature). The genomes are evaluated ("survivor selection strategy")—this stage could include human (interactive) as well as automated evaluation—and only genomes fulfilling the evaluation criteria are retained. The diminished population is then expanded in numbers and in variety—typically the successful genomes are used as the basis for generating new ones via biologically-inspired processes such as recombination and mutation.*

design population is generated randomly, the algorithm effectively expands the imagination of the human designer.

Although this strategy enables the design size (i.e., the number of individual features that must be explicitly specified) to be expanded practically without limit, one typically sacrifices knowledge of the exact internal workings of the device, introducing a level of unpredictability into device performance that may require a new engineering paradigm to be made acceptable.

Genetic algorithms[4] use bit strings to encode the target object. The genome is fixed in advance, only the combinations of presence and absence of individual features can be modified. In other words, the form of the solution is predetermined. For example, if the solution can be expressed as an equation, the coefficients evolve but not the form of the equation. More advanced algorithms relax these conditions; that is, the genome length can vary and additions and deletions are possible. These schemata are rather far from natural selection, and might best be described as artificial selection.

[4] J.H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press (1975).

Genetic programing[5] works at a higher level, in which the algorithm itself evolves. In other words, the form of the solution can evolve. Typically the solution is defined by trees of Lisp-like expressions, and changes can be made to any node of the tree. Genetic programing is closer to natural selection.

Human knowledge can be captured not only in the design of the algorithms, but also by incorporating an interactive stage in the fitness evaluation.[6]

11.3 MATERIALS SELECTION

Ashby has systematized materials selection through his property charts.[7] For example, Young's modulus E is plotted against density ρ for all known materials, ranging from weak light polymer foams to strong dense engineering alloys. The huge interest in nanomaterials is that it may be possible to populate empty regions on such charts, such as strong and light (currently natural woods are as close as we can get to this) or weak and dense (no known materials exist).

Material properties are only the first step. Shapability is also important, in ways that cannot be easily quantified. For example, rubber can readily be manufactured as a sealed tube, in which form it can serve as a pneumatic tire, but it is at risk from punctures, and a novel solid material may be useful, and more robust, for the same function. Finally availability (including necessary human expertise) and cost—linked by the laws of supply and demand—must be taken into consideration. Nanotechnology, by allowing rapid material prototyping, should greatly enhance the real availability of novelty. An assembler should in principle allow any combination of atoms to be put together to create new materials.

FURTHER READING

W. Banzhaf, G. Beslon, S. Christensen, J.A. Foster, F. Képès, V. Lefort, J.F. Miller, M. Radman and J.J. Ramsden, From artificial evolution to computational evolution: a research agenda. *Nature Rev. Genetics* 7 (2006) 729–735.

---

[5] J.H. Koza, *Genetic Programming*. Cambridge, MA: MIT Press (1992).

[6] E.g., A.M. Brintrup, H. Takagi, A. Tiwari and J.J. Ramsden, Evaluation of sequential, multi-objective, and parallel interactive genetic algorithms for multi-objective optimization problems. *J. Biol. Phys. Chem.* 6 (2006) 137–146.

[7] M.F. Ashby, *Materials Selection in Mechanical Design*. Oxford: Pergamon (1992).

CHAPTER 12

The Future of Nanotechnology

CHAPTER CONTENTS

Whereas Chapter 10 was mainly devoted to the prediction of substitutional and incremental nanotechnology, in this chapter we address the long-term future, for which the traditional methods of economic forecasting are of little use.

As pointed out by Toth-Fejel,[1] important ways to predict the future include:

Via prophets: who are individuals with charisma, a track record of successful predictions (ideally based on an intelligible chain of reasoning) and the courage to contradict popular opinion. The value of the prophet's work might be primarily derived from a cogent marshaling

[1] T. Toth-Fejel, Irresistible forces vs immovable objects: when China develops Productive Nanosystems. *Nanotechnol. Perceptions* 4 (2008) 113–132.

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

of relevant data; the prophecy is important when it is accompanied by a similar creative leap as when a theory emerges from a mass of experimental data.

Through history: one looks for patterns in the past to find analogies for, or extrapolations into, the future. Predictions tend to be necessarily rather vague—that is, made at a fairly high level. This method does not have a good track record, despite significant apparent successes (e.g., the First World War following from the Franco-Prussian War because the latter's terms of peace were too onerous for France, and the Second World War following from the First World War because the latter's terms of peace were too onerous for Germany; doubtless some participants of the Versailles peace conferences were aware of the dangers of what was being done, but the proceedings got bogged down in a morass of detail and were bedeviled by partisan considerations).

Trend ranking: one reasonably assumes that the significance of a trend depends on its rate of change and duration: Typically, highly significant trends (e.g., accelerating technology, increasing recognition of human rights) will enslave weaker ones (slow and short, e.g., business cycles and fashion).

Engineering vs science: scientific discoveries (e.g., X-rays, penicillin, Teflon) are impossible to predict (we exclude discoveries of facts (e.g., the planet Uranus) that were predicted by theories, in the formulation of which previously discovered facts played a rôle). On the other hand, engineering achievements (e.g., landing a man on the Moon) are predictable applications of existing knowledge that adequate money and manpower solved on schedule. According to the new model (Figure 2.3), new technologies (e.g., atomic energy and nanotechnology) are closely related to scientific discovery, making them concomitantly harder to predict.

The will to shape the future: the idea that the future lies in man's hands (i.e., he has the power to determine it).[2] This stands in direct opposition to predestination. Reality is, of course, a mixture of both: the future involves unpredictability but is subject to certain constraints.

[2]One of the more prominent philosophers associated with this idea was F. Nietzsche. However, he also believed in the idea of eternal return (the endless repetition of history).

Scenarios: not included in Toth-Fejel's list, but nevertheless of growing importance, e.g., in predicting climate change.[3]

12.1 PRODUCTIVE NANOSYSTEMS

The technological leap that is under consideration here is the introduction of desktop personal nanofactories.[4] These are general-purpose assemblers that represent the ultimate consummation of Richard Feynman's vision, capable of assembling things atom by atom using a simple molecular feedstock such as acetylene or propane, piped into private houses using the same kind of utility connection that today delivers natural gas. Such is the nature of this technology that once one personal nanofactory is introduced, the technology will rapidly spread, certainly throughout the developed world. It may be assumed that almost every household will purchase one.[5] What, then, are the implications of this?

Science of this era of productive nanosystems can be summed up as a quasi-universal system of "localized, individualized ultralow-cost production on demand using a carbon-based feedstock." Let us briefly take each of these attributes in turn.

Localized Production will practically eliminate the need for transport of goods. Transport of goods and people accounts for 28% of fossil fuel usage (compared with 32% used by industry),[6] at least half of which would no longer be necessary with the widespread introduction of the personal nanofactory. This would obviously have a hugely beneficial environmental impact.

We have become accustomed to the efficiency of vast central installations for electricity generation and sewage treatment, and even of healthcare, but future nanotechnology based on productive nanosystems will reverse that trend. Ultimately it will overturn the paradigm of the division of labor that was such a powerful concept in Adam Smith's conception of economics. In turn, globalization will become irrelevant and, by eliminating it, one of the gravest threats to the survival of humanity, due to the concomitant loss of diversity of thought and technique, will be neutralized.

---

[3] M. Anissimov et al., The Center for Responsible Nanotechnology Scenario Project. *Nanotechnol. Perceptions* 4 (2008) 51–64.

[4] K.E. Drexler, *Nanosystems: Molecular Machinery, Manufacturing, and Computation*. New York: Wiley (1992).

[5] Freitas (loc. cit.) assumes a 20-year interval for their introduction.

[6] G.C. Holt and J.J. Ramsden, Introduction to global warming. In: J.J. Ramsden and P.J. Kervalishvili (eds), *Complexity and Security*, pp. 147–184. Amsterdam: IOS Press (2008).

Individualized Production or "customized mass production" will be a powerful antidote to the products of the Industrial Revolution that are based on identical replication. In the past, to copy (e.g., a piece of music) meant writing it out by hand from an available version. This was in itself a powerful part of learning for past generations of music students. Nowadays it means making an identical photocopy using a machine. In Rome, although crockery was made on a large scale, each plate had an individual shape; almost two millennia later, Josiah Wedgwood rejoiced when he could make large numbers of identical copies of one design. The owner of a personal nanofactory (the concrete embodiment of a productive nanosystem) will be able to program it as he or she wishes (as well as having the choice of using someone else's design software).

Ultralow-Cost Production will usher in an era of economics of abundance. Traditional economics, rooted in the laws of supply and demand, are based on scarcity. The whole basis of value and business opportunities will need to be rethought.

Production on Demand also represents a new revolutionary paradigm for the bulk of the economy. Only in a few cases—the most prominent being Toyota's "just-in-time" organization of manufacture—has it been adopted in a significant way. A smaller-scale example is provided by the clothing company Benetton—garments are stored undyed centrally, and dyed and shipped in small quantities according to feedback regarding what is selling well from individual shops. Not only does this lead to a reduction of waste (unwanted prodution), but elimination of a significant demand for credit comes from production in anticipation of demand. Personal nanofactory-enabled production on demand represents the apotheosis of these trends.

Carbon-Based Feedstock The implications of carbon-based feedstock (acetylene or propane, for example) as a universal fabrication material are interesting. The production of cement, iron and steel, glass and silicon account for about 5% of global carbon emissions. Much of this would be eliminated. Furthermore, the supply of feedstock could, given an adequate supply of energy, be sequestered directly from the atmosphere.

12.2 SOCIAL IMPACTS

Although the anticipated course of nanotechnology-based technical development can be traced out, albeit with gaps, and on that basis a fairly detailed economic analysis carried out,[7] ideas regarding the social impacts of these revolutionary changes in manufacturing are far vaguer. An attempt was made a few years ago,[8] (typically) stating that "nanotechnology is being heralded as the new technological revolution ... its potential is clear and fundamental ... so profound that it will touch all aspects of the economy and society. Technological optimists look forward to a world transformed for the better by nanotechnology. For them it will cheapen the production of all goods and services, permit the development of new products and self-assembly modes of production, and allow the further miniaturization of control systems. They see these effects as an inherent part of its revolutionary characteristics. In this nano society, energy will be clean and abundant, the environment will have been repaired to a pristine state, and any kind of material artefact can be made for almost no cost. Space travel will be cheap and easy, disease will be a thing of the past, and we can all expect to live for a thousand years." Inevitably, such attempts are vaguest where, if not details, at least clues as to how the leaps will be made are given.[9] Furthermore, these writings remain silent about how people will think under this new régime; their focus is almost exclusively on material aspects. There is perhaps more recognition of nanotechnology's potential in China, where the Academy of Sciences notes that "nanodevices are of special strategic significance, as they are expected to play a critical role in socio-economic progress, national security and science and technology development."

Traditional technology (of the Industrial Revolution) has become something big and powerful, tending to suppress human individuality; men must serve the machine. Moreover, much traditional technology exacerbates conflict between subgroups of humanity. This is manifested in the devastation of vast territories by certain extractive industries, but also by the "scorched earth" bombing of cities such as Dresden and Hamburg in World War II.

In contrast, nanotechnology is small without being weak and is perhaps "beautiful". Since in its ultimate embodiment as productive nanosystems it becomes individually shapable, it does not have all the undesirable features of "big" technology; every individual can be empowered to the degree of his

[7] R.A. Freitas, Jr, Economic impact of the personal nanofactory. *Nanotechnol. Perceptions* 2 (2006) 111–126.

[8] S.J. Wood, R.A.L. Jones and A. Geldart, *The Social and Economic Challenges of Nanotechnology.* Swindon: Economic and Social Research Council (2003).

[9] For a critique, see J.J. Ramsden, The music of the nanospheres, *Nanotechnol. Perceptions* 1 (2005) 53–64.

or her personal interests and abilities. It is therefore important that in our present intermediate state nanotechnology is not used to disempower.[10]

12.3 TIMESCALES

"True" nanotechnologists assert that the goal of nanotechnology is productive nanosystems, and that the question is "when" not "if". Opponents implicitly accept the future reality of assemblers, and oppose the technology on the grounds of the dangers (especially that of "grey goo"—assemblers that run out of control and do nothing but replicate themselves, ultimately sequestering the entire resources of the Earth for that purpose). Finally there is a group that asserts that nanotechnology is little more than nanoparticles and scanning probe microscopes and that all the fuss, even the word "nanotechnology", will have evaporated in less than a decade from now.

This last attitude is rather like viewing the Stockton and Darlington Railway as the zenith of a trend in transportation that would soon succumb to competition from turbocharged horses. And yet, just as the company assembled on the occasion of the Rainhill engine trials could have had no conscious vision of the subsequent sophistication of locomotives such as the Caerphilly Castle, the Flying Scotsman or the Evening Star, and would have been nonplussed if asked to estimate the dates when machines fulfilling their specifications would be built, so it seems unreasonable to demand a strict timetable for the development of advanced nanotechnology. It should be emphasized that by the criterion of atomically precise manufacturing, today's nanotechnology—overwhelmingly nanoparticles—is extremely crude. But this is only the first stage, that of passive approximate nanostructures. Applications such as sunscreen do not require greater precision. Future envisaged phases are:

Active nanodevices: able to change state, transform and store information and energy, and respond predictably to stimuli. Integrated circuits with 65 nm features (made by "top-down" methods) belong here. Nanostorage devices (e.g., based on single electrons or molecules),

[10] An example of disempowerment is the recent development of "theranostics"—automated systems, possibly based on implanted nanodevices, able to autonomously diagnose disease and automatically take remedial action; for example by releasing drugs. In contrast to present medical practice, in which a practitioner diagnoses, perhaps imperfectly, and proposes a therapy, which the patient can accept or refuse, theranostics disempowers the patient, unless he was involved in writing the software controlling it.

biotransducers and the quantum dot laser are examples that have reached the stage of proof of principle. It is noteworthy that self-assembly ("bottom-up") nanofacture is being pursued for some of these.

Complex machines: able to implement *error correction codes*, which are expected to improve the reliability of molecular manufacturing by many orders of magnitude consider chemical syntheses with error rates around 1 in a 100 (a yield of 99% is considered outstanding); natural protein synthesis with error rates of 1 in 10^3–10^4, DNA replication with error rates of 1 in 10^6, and computers with error rates better than one in 10^{23} operations thanks to error detection and correction codes originally developed by Hamming and others, without which pervasive low-cost computing and all that depends on it, such as the internet, would not be possible. Algorithmic concepts are very significantly ahead of the physical realization (see Section 1.1). The main practical approaches currently being explored are tip-based nanofabrication (i.e., diamondoid mechanosynthesis; or patterned depassivation followed by atomic layer epitaxy) and biomimicry (DNA "origami" and bis-peptide synthesis).

Productive nanosystems: able to make atomically precise tools for making other (and better) productive nanosystems, as well as useful products. Current progress and parallels with Moore's law suggest that they might be available in 10–20 years.

12.4 SELF-ASSEMBLY

Although "passive" self-assembly creates objects of indeterminate size, except in the special case of competing interactions of different ranges,[11] and hence not useful for most technological applications (especially device nanofacture), biology shows that useful self-assembly is possible (e.g., the final stages of assembly of bacteriophage viruses[12]). It depends on initial interactions altering the conformations of the interacting partners, and hence the spectrum of their affinities (see Figure 12.1). Called programmable self-assembly (PSA), it

[11] J.J. Ramsden, The stability of superspheres. *Proc. R. Soc. Lond. A* 413 (1987) 407–414.

[12] E. Kellenberger, Assembly in biological systems. In: *Polymerization in Biological Systems*, CIBA Foundation Symposium 7 (new series). Amsterdam: Elsevier (1972).

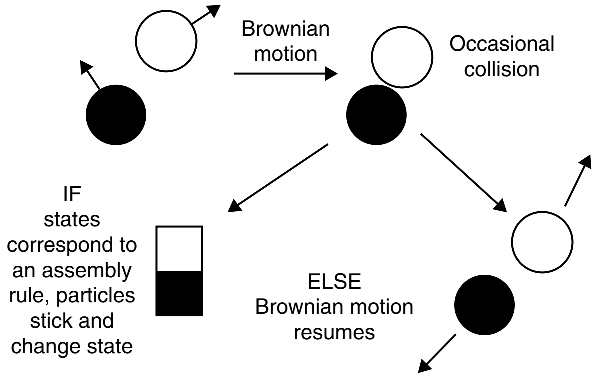


FIGURE 12.1 *Illustration of programmable self-assembly.*

can be formally modeled by graph grammar, which can be thought of as a set of rules encapsulating the outcomes of interactions between the particles.[13] While macroscopic realizations of PSA have been achieved with robots, we seem to be a long way off mimicking biological PSA in wholly artificial systems. Modifying biological systems is likely to be more achievable, and there is intensive research in the field.[14] The approach seems to be at least as promising as the assembler concept.

12.5 MOLECULAR ELECTRONICS

The industry view of the continuation of Moore's law is supposed to be guaranteed for several more years via further miniaturization and novel transistor architectures. Another approach to ultraminiaturize electronic components is to base them on single organic molecules uniting an electron donor (D^+, i.e. a cation) and an acceptor (A^-, i.e. an anion) separated by an electron-conducting bridge (i.e. a π-conjugated (alkene) chain). The molecule is placed between a pair of (usually dissimilar) metal electrodes $M^{(1)}$ and $M^{(2)}$,[15] chosen for having suitable work functions and mimicking a semiconductor p–n junction. Forward bias results in $M^{(1)}/D^+–\pi–A^-/M^{(2)} \rightarrow M^{(1)}/D^0–\pi–A^0/M^{(2)}$, followed by intramolecular tunneling to regenerate the

[13] E. Klavins, Universal self-replication using graph grammars. In: *Intl Conf. on MEMs, NANO and Smart Systems*, Banff, Canada (2004).

[14] J. Chen, N. Jonoska and G. Rozenberg (eds), *Nanotechnology: Science and Computation*. Berlin: Springer (2006).

[15] See, e.g., A.S. Martin et al., Molecular rectifier. *Phys. Rev. Lett.* 70 (1993) 218–221.

starting state. Reverse bias tries to create D^{2+}–π–A^{2-}, but this is energetically unfavorable and hence electron flow is blocked (rectification). This technology is still in the research phase, with intensive effort devoted to increasing the rectification ratio.

12.6 QUANTUM COMPUTING

Extrapolation of Moore's law to about the year 2020 indicates that component size will be sufficiently small for the behavior of electrons within them to be perturbed by quantum effects, implying the end of the semiconductor road map and conventional logic. Another problem with logic based on moving charge around is energy dissipation. Quantum logic (based on superposition and entanglement) enables computational devices to be created without these limitations and intensive academic research is presently devoted to its realization.

The physical embodiment of a bit of information—called a qubit in quantum computation—can be any absolutely small object capable of possessing the two logic states 0 and 1 in superposition, e.g., an electron, a photon or an atom. A single photon polarized horizontally (H) could encode the state $|0\rangle$ and polarized vertically (V) could encode the state $|1\rangle$ (using the Dirac notation). The photon can exist in an arbitrary superposition of these two states, represented as $a\,|\mathrm{H}\rangle + b\,|\mathrm{V}\rangle$, with $|a|^2 + |b|^2 = 1$. The states can be manipulated using birefringent waveplates, and polarizing beamsplitters are available for converting polarization to spatial location. With such common optical components, logic gates can be constructed.[16] Another possible embodiment of a qubit is electron spin (a "true" spintronics device encodes binary information as spin, in contrast to the so-called spin transistor, in which spin merely mediates switching).[17]

FURTHER READING

P.M. Allen, Complexity and identity: the evolution of collective self. In: J.J. Ramsden, S. Aida and A. Kakabadse (eds), *Spiritual Motivation: New Thinking for Business and Management*, pp. 50–73. Basingstoke: Palgrave Macmillan (2007).

[16] A. Politi and J.L. O'Brien, Quantum computation with photons. *Nanotechnol. Perceptions* 4 (2008) 289–294.

[17] S. Bandyopadhyay, Single spin devices—perpetuating Moore's law. *Nanotechnol. Perceptions* 3 (2007) 159–163.

CHAPTER 13

Grand Challenges

CHAPTER CONTENTS

Human society is widely considered to have entered a difficult period. It is confronted with immense challenges of a globally pervasive nature. Extrapolating present trends leads to a grim picture of the possibility of a miserable collapse of civilization. Because of globalization, the collapse is likely to be global—whereas in the past, different "experiments" (types of socio-economic organization) were tried in different places, and the collapse of one (e.g., the Aztec empire) did not greatly affect others. During the previous half-century, destruction by nuclear weapons was considered to be the greatest threat, but this was clearly something exclusively in human hands, whereas now, even though the origins of the threats are anthropogenic, mankind seems to be practically powerless to influence the course of events.

Can a new technology help? Several decades ago, nuclear fusion—using the vast quantities of deuterium in the oceans—was seen as a way to solve the challenge of rapidly depleting fossil fuel reserves. As it happens, that technology has not delivered the promised result, but now a similar latent

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

potential inheres in nanotechnology, which as a universal technology should in principle be able to solve all the crises.

13.1 MATERIAL CRISES

The crises, and the potential contributions of nanotechnology, are:

Climate change, especially global warming.[1] If the cause is anthropogenic release of carbon dioxide, then any technology that tends to diminish it will be beneficial, and nanotechnology in the long term may have that capability (see §12.1). If the cause is not anthropogenic and due to (for example) variations in the solar constant, then we anyway need to enhance our general technological capabilities to give us the power to combat the threat.

Demography. This comprises both population growth, considered to be excessive and becoming too large for Earth to support, and aging of the population.

The latter is partly a social matter. It is customary for elderly people to retire from active work, but their income as *rentiers* (i.e., old-age pensioners) is ensured by those still actively working, so if the ratio of the latter to the former decreases, the pension system collapses. There is also a matter of healthcare: elderly people tend to require more resources. Advances in medical care, to which nanotechnology is making a direct contribution (Chapter 8), will diminish the healthcare problem. Therefore, elderly people will be able to continue to work longer, diminishing the threat to the pension system. At the same time, advances in technology should further diminish the prevalence of unpleasant jobs from which one is glad to retire. In any case, unproductive work (e.g., involving an advisory or decision-making rôle without an impact on production—town planning is a good example, along with membership of similar councils or committees) could be preferentially assigned to elderly people.

If, in principle at least, the problem of aging populations can thus be solved, the same cannot be said for population growth. It is probably best to consider it as a medical problem (as it is already in many countries), in which case the technology of Chapter 8 is applicable.

[1] G.C. Holt and J.J. Ramsden, Introduction to global warming. In: J.J. Ramsden and P.J. Kervalishvili (eds), *Complexity and Security*, pp. 147–184. Amsterdam: IOS Press (2008).

Environmental degradation. This is mostly a gradual change, but by being slow it is pernicious, and suddenly we have a seemingly irreparable dust bowl in the Mid-West or an Aral Sea disaster. The latter was a fairly direct result of the massive diversion of the two main feeder rivers, the Amu Darya and the Syr Darya, into irrigation, mainly of cotton fields, although unpredictable nonlinearities appeared near the end. Whether the immediate restoration of full flow would regenerate the sea is a moot point, and anyway has not been seriously considered because of the immense social dislocation that would result from the collapse of the cotton agro-industry. Some of the world's great deserts such as the Sahara and the Gobi are also currently expanding, but this appears to be a cyclic phenomenon linked to long-term climate changes. In fact, the mechanisms of desertification are not well understood, and efforts to investigate it piecemeal. Presumably the United Nations Organization declared 2007 as the Year of Desertification in an effort to improve matters, but it was singularly ineffectual in inspiring a coordinated global effort to tackle the problem. In any case, it is not clear how nanotechnology can contribute.

Another major challenge is environmental pollution. In the long term, nanotechnology will certainly help to alleviate it via significantly increasing the overall efficiency of production (Chapter 12). Certain remediation technologies based on nanoparticles (cf. Section 6.8) may help to alleviate local problems.

Depletion of resources. Nanotechnology should have a generic beneficial effect, because if the same function can be achieved using a less material, obviously fewer resources will be used.

Furthermore, very light yet strong materials (probably based on carbon nanotubes) are likely to be of inestimable value for space travel—including the space elevator, which would enormously facilitate departures from the planet. Continuing increases in processing power will enhance the feasibility of unmanned space missions—for example, to neighboring planets in order to mine precious metals that may no longer be obtainable on Earth.

Financial chaos. The new economic order associated with nanotechnology—productive nanosystems (Chapter 12)—based on production on demand may solve this problem automatically, since the rôle of credit will diminish. Given that the realization of productive nanosystems is not anticipated before at least another decade, and quite possibly two, have elapsed, it must not be hoped for that nanotechnology will come to the rescue of the present troubles, but perhaps it will prevent a fresh occurrence of bubbles.

Terrorism. This is above all a social problem,[2] which might disappear if nanotechnology ushers in a new ethical era (see Chapter 14).

It is doubtful whether all the challenges can be met simultaneously, therefore priorities will have to be set. One notices that demography (especially population growth) is, in fact, the most fundamental challenge, in the sense that solving this one will automatically solve the others. It seems appalling that countries whose populations are falling (many European countries, Russia and Japan) are being encouraged to promote immigration—to stave off collapse of their social security systems! The prolongation of healthy human life is one of the more reliably extrapolatable trends, and a clear corollary is that world population should stabilize at a lower level than otherwise. In blunt ecological terms, "be fruitful and multiply" is an appropriate injunction in a relatively empty world in which r-selection (see Section 3.1) operates, but in our present crowded, technologically advanced era K-selection is appropriate, typified by a sparser, but longer-living population.

13.2 SOCIAL CRISES

Material problems are usually in the forefront of attention, but any technological revolution also brings psychological and social problems in its wake. One of the general problems of technology increasing leisure (see Figure 2.2) is that people might find it harder to lead meaningful lives. We may have to ask whether we really need a further increase in the abundance of labor-saving devices. More attention will need to be paid by everybody to continue to exercise body and mind: Hebb's rule essentially guarantees that the brain will atrophy in the absence of thought trajectories. This issue, and related ones, is taken up more fully in the next and final chapter.

13.3 IS SCIENCE ITSELF IN CRISIS?

The idea of scientific endeavor being harnessed to explicitly solve grave global problems is an attractive one. Bacon's stress on one of the purposes of scientific investigation being for the "relief of man's estate" would encourage

[2] S. Galam, The sociophysics of terrorism: a passive supporter percolation effect. In: J.J. Ramsden and P.J. Kervalishvili (eds), *Complexity and Security*, pp. 13–37. Amsterdam: IOS Press (2008).

that idea, and most scientists would probably agree that a basic humanitarian aim of science is to help promote human welfare. However, as Maxwell has pointed out, science seeks this by pursuing the purely intellectual aim of acquiring knowledge in a way (called standard empiricism (SE) by Maxwell) that is sharply dissociated from all consideration of human welfare and suffering.[3] Under the aegis of SE any desire to solve global challenges is likely to be little more than a velleity. Maxwell advocates replacing SE by aim-oriented empiricism (AOE) as a better philosophy of science. Not only is it more rigorous, but value (to humanity and civilization) becomes an intrinsic part of its pursuit. Even in a highly abstract discipline the adoption of AOE will be a step forward, because of its greater rigor, although the fruits (research output) are not likely to look very different. In areas other than the "hard" sciences, the difference is likely to be dramatic. The social sciences, including sociology and economics, have become largely useless to humanity. On the contrary, astonishingly and sadly, the attitudes prevailing among many academic sociologists and economists tend to drag humanity down whenever they are taken up by politicians or other activists. Social science should be replaced by something called social inquiry, social methodology or social philosophy, concerned to help humanity tackle its immense problems of living in more rational ways than at present, and seeking to build into social life progress-achieving methods arrived at by generalizing AOE, the progress-achieving methods of the natural sciences.

The natural sciences themselves need to acquire a tradition of criticism that has long been a part of literary and artistic work. Because, unlike those other areas of endeavor, the natural sciences contain their own internal validation—the predictions of a theory can always be tested via empirical observation—they have not felt the same need to develop a tradition of criticism that in the literary world is as esteemed as the creation of original works. In consequence, much science tends to move in sterile or even counterproductive directions.[4] One often hears reference to the "rapid progress" in many

---

[3] N. Maxwell, Do we need a scientific revolution? *J. Biol. Phys. Chem.* 8 (2008) 96–105.

[4] Scientometricians might argue that their work constitutes a kind of objective criticism. In principle perhaps, but in practice it degenerates into populism. For example, one of their best-known inventions is counting the number of times a published paper is cited, whence the infamous "impact factor" (the number of citations received by a journal divided by the number of papers published in the journal). Even the research director of the Institute of Scientific Information (ISI), which pioneered the extensive compilations of impact factors, recognized that they are only valid if authors scrupulously cite all papers that they should, and only those. This seems seldom to be the case. Now that the ISI has been taken over by a commercial organization (Thomson), there is even less reason to put any value on an impact factor.

fields, especially the biological sciences. To be sure, advances in techniques and instrumentation have yielded impressive results, but if the direction is wrong, it does not mean very much. These weaknesses will become more obvious when the challenges to which science as presently practiced might be able to respond, but does not or cannot, become more important.

13.4 NANOTECHNOLOGY-SPECIFIC CHALLENGES

Any revolution brings its own attendant new challenges (typically referred to as "birth pangs" and the like). There is already widespread implicit recognition of them. An example is presented by the report on nanoparticle risks commissioned by the British government.[5] This addresses the need to assess human exposure to engineered nanomaterials, evaluate their toxicity, and develop models to predict their long-term impacts. Similar investigations should be undertaken to establish effects on the overall ecosystem, including plant and microbial life. Given the extensive data that already exists, at least concerning human health impacts,[6] care should be taken to avoid the waste and futility of endless studies aimed at the same goal, and all deficient in some regard. Meanwhile, more attention should be paid to how to act efficaciously upon the findings.

Productive nanosystems (PNs, Section 12.1) raise the risk of "grey goo" (the uncontrolled subversion of all terrestrial matter to assembling assemblers). Given that PNs are expected to be at least 10–20 years in the future, should we already be concerned at this eventuality? Probably not, since it is still associated with so many imponderables.

There is strong military interest in nanotechnology,[7] raising the specter that it will "met en oeuvre des moyens de mort et de destruction incomparablement plus efficaces que par le passé".[8] Albert Schweitzer points out that history shows that victory by no means always belongs to the superior civilization; as often as not a more barbaric power has conquered. This problem is not specially associated with nanotechnology; but there is at least the hope

---

[5]C.L. Tran et al., *A Scoping Study to Identify Hazard Data Needs for Addressing the Risks Presented by Nanoparticles and Nanotubes*. London: Institute of Occupational Medicine (2005).

[6]E.g., P.A. Revell, The biological effects of nanoparticles. *Nanotechnol. Perceptions* 2 (2006) 283–298.

[7]J. Altmann, *Military Nanotechnology*. London: Routledge (2006).

[8]A. Schweitzer, *Le problème de la paix*. Nobel Lecture, 4 November 1954.

that a pre-singularity surge of human solidarity may neutralize it (see also Section 13.6 and Chapter 14).

13.5 GLOBALIZATION

Perhaps *the* greatest socio-economic-technical danger faced by humanity is that of globalization. Advances in transport and communication technology have made it seem inevitable, and it appears as the apotheosis of Adam Smith's economic system (based on the division of labor) that has been so successful in augmenting the wealth of mankind. Yet globalization carries within it the seeds of great danger: that of diminishing and fatally weakening the diversity that is so essential a part of our capability of responding to security threats.[9] The disappointing uniformity of products emanating from the far-flung reaches of the British Empire was already apparent to foreign (European) visitors to the British Empire (Wembley) Exhibition of 1924. Evidently productive nanosystems are *in principle* antiglobalizing, since products with the same function can be made anywhere. Only if subpopulations are too indolent to master the technology and produce their own designs will they lapse into a position of weakness and dependence.

13.6 AN INTEGRATED APPROACH

The message of this chapter is that there is little point in developing revolutionary nanotechnology without parallel developments in the organization of society. But the irrepressible curiosity and creativity of the scientist will inevitably drive the technology forwards—despite all the obstacles placed in the way by unsympathetic bureaucracy!—science is fundamentally a progressive activity, ever aiming at a distant goal without striving to be instantaneously comprehensive. But no similar tendency exists regarding society. The ebb and flow of social tendencies is ever-evolving and open-ended. At one level, the official vision of technology (as exemplified by policy declarations of government research councils, for example) is aimed at ever stricter scrutiny and surveillance, in which the bulk of the population is seen as a restless, unreliable mass in which criminal disturbances may break out at any instant. At another level, the spectral impalpability of so-called high finance is allowed to become ever more impenetrable and autonomous, which might

[9] J.J. Ramsden and P.J. Kervalishvili (eds), *Complexity and Security*. Amsterdam: IOS Press (2008).

be acceptable were it not for the real effects (in terms of ruined livelihoods and drastically adjusted currency exchange rates) that are now manifest.

One solid correlation that should give us undying optimism is the link between the growth of knowledge and the improvement of ethical behavior.[10] The best hope for the future—given the impracticability of anything other than piecemeal social engineering—is to constantly promote the growth of knowledge, and given that our knowledge about the universe is still very, very incomplete, keep in mind Donald Mackay's dictum, "When data is short, keep mind open and mouth shut."

[10] H.T. Buckle., *History of Civilization in England*, Vol. I, Chap. 4. London: Longmans, Green & Co. (1869).

CHAPTER 14

Ethics and Nanotechnology

CHAPTER CONTENTS

Buckle has pointed out that the foundations of ethics have essentially not advanced for at least the last two millennia. At the same time, there have been enormous advances in what we now call human rights. Since the philosophical foundations of ethics have not changed, we must look elsewhere for the cause. What has grown spectacularly is knowledge. Therefore, it can be concluded that the reason why we treat each other on the whole much better than formerly is because we know more about the universe:[1] one should recall that one of the functions of science is to enable man to better understand his place within the universe. This advance might actually be considered the most significant contribution of science to humanity, outweighing the many contributions ministering to our comfort and convenience. Thus, in a general

[1] As Buckle would have said, moral truths are stationary, and dependent on the state of intellectual knowledge for their interpretation.

Applied Nanotechnology: The Conversion of Research Results to Products, ISBN 9780815520238

way, the advance of knowledge, regardless of what that knowledge is, should be beneficial to humanity.

As already pointed out in Chapter 2, knowledge may be turned into technology. Now, when we look around at all that technology has brought us, we are confronted with a familiar paradox. Explosives allow the quarryman to more expeditiously extract stone with which we can build dwellings for ourselves, but fashioned as bombs have also wrought terrible destruction on dwellings—such as in Hamburg or Dresden in the Second World War (and, very recently, in Gaza in Palestine). Nuclear fission can provide electricity without emitting carbon dioxide, but also forms the core technology of the weapons dropped with such terrible effect upon Hiroshima and Nagasaki in the Second World War; further examples seem scarcely necessary. Hence, when it is (sometimes) stated that "technology is ethically neutral", the meaning must be that there is no *net* benefit or disadvantage from its application—with some kind of ergodic principle tacitly assumed to be valid; that is, neutral provided we observe a wide enough range of human activity, or over a sufficiently long interval.

But if so, then why is any technology introduced? On the contrary, technologists believe that they are embellishing life on Earth. There would be no sense in introducing any technology whose disbenefits outweigh the benefits. Hence technology is not neutral, but positive.[2]

A final question for this chapter is, are ethics associated with any particular technology? Is there an ethics of the steam engine, of motoring, of cement manufacture, of space travel—and of nanotechnology? We return to this question in Section 14.5.

14.1 RISK, HAZARD AND UNCERTAINTY

Technological progress typically means doing new things. It may be considered unethical to proceed with any scheme that exposes the population to risk. But how much risk is acceptable? Human progress would be impossible if every step taken had zero risk. In fact, the risk of doing something for the first time is formally unquantifiable, because the effects are unknown. In practice, the variety of past experience is used to extrapolate. Steam locomotives traveling on rails allowed faster travel than previously, but initially not that much faster than a man on a horse. Tunnels caused some problems, but

[2]In some cases, it transpires that what is a benefit to one subpopulation is a disbenefit to another (the latter usually having no say about the development). According to the principle of human solidarity, this is unethical.

natural caves were known and had been explored. Flying was a greater innovation, but birds and bats were familiar. Furthermore, only small numbers of people were initially involved.

New technology raises two aspects of risk: do we proceed with a new technology and do we need to regulate an already existing state of affairs? The former is typically the decision of a person or a company. The latter is typically a collective decision of society, through its institutions.

Both cases imply firstly the need to quantify risk, and then the need to decide on a limit. Let us tackle the quantification.

Risk can be defined as the hazard associated with an event multiplied by the probability of the event occurring. This decomposition is, indeed, the most common current basis of risk analysis and risk management. Two major difficulties immediately present themselves, however: how can hazard be quantified and how is the probability to be determined?

The answer to the first question is typically solved by using cost. Although in a particular case it might be difficult, practically speaking, because events are typically complex, nevertheless intelligible estimates can usually be made, even of the cost of events as complex as flooding. The insurance industry has even solved the problem of costing a human life. The numbers of people affected can be estimated. Hence, even if there are imponderables, a basis for estimation exists.

To answer the second question, one needs an appropriate probability model. If the event occurs reasonably frequently, the frequentist interpretation should be satistfactory. However, successive events may be correlated. Subevents aggregating to give the observed event may be additive or multiplicative, and so forth. As with the first question, heuristic approximation may be required.

The units of risk, as quantified in this manner, are cost per unit time.

Risk analysis comprises the identification of hazards (or "threats")—operational, procedural etc.—and evidently the better one understands the operation, procedure etc. the better one can identify the hazards. Risk management comprises attempts to diminish the hazard (its cost), or the probability of its occurrence, or both. The two are linked by comparing the costs of remedial action with the resulting change of the amount of risk.

The overall object is to decide whether to undertake some action to diminish the hazard, or the probability of its occurrence, or both. If the cost of the action is less than the value gained by diminishing the risk, then it is reasonable to undertake the action. A slight difficulty is that sometimes a single action is carried out and there are no recurrent costs. In this case the cost of the action should be divided by the anticipated duration of its effect. A more severe difficulty is that often hazard and probability are linked. For example, installing airbags in motor-cars diminishes the hazard of an accident, but the driver, knowing this, might tend to drive more recklessly, hence increasing the probability of an accident. This factor, often neglected, frequently makes the actual effects of remedial actions very significantly less than foreseen.

The application of this kind of approach to the introduction of new technology has already been tackled in Chapter 4. In the next section, we discuss its application to regulation.

14.2 REGULATION

Regulation represents a form of risk management; indeed, in the language of cybernetics it is precisely that. As evidence for the potential dangers of nanoparticles becomes more widely appreciated, calls to regulate their manufacture and use are heard. Regulators appointed as state officials have become widespread in recent years (Great Britain appears to be the leader of this trend); the motivation for their appointment is usually when a state monopoly, such as telephones or electricity, is privatized. There is a slight paradox here because the reason for privatization is typically the belief that the free market ensures a more cost-effective industry, but the perceived need for regulation implies that the market cannot be trusted. Moreover since there is at present no real theory of regulation the regulator acts by instinct, and his actions might well bring about a less cost-effective industry as often as the opposite. Furthermore regulation has itself a cost (apart from the regulator's salary): regulation should only be introduced if the benefits exceed the costs.

On the whole, regulators (and here we may include much of the "health and safety" apparatus that has become firmly entrenched in the industrial scene) have a deleterious effect upon activities. I.K. Brunel, opposing the appointment of Government Inspectors of Railways in 1841, aptly observed that "Railway engineers understand very well how to look after the public safety, and putting a person over them must shackle them. They have not only more ability to find out what is necessary than any inspecting officer could have, but they have a greater desire to do it." Seven years later (1848), he was obliged to express similar sentiments with respect to the Royal Commission on the Application of Iron to Railway Structures: "... it is to be presumed that they will lay down, or at least suggest, 'rules' and 'conditions' to be observed in the construction of bridges, or, in other words, embarrass and shackle the progress of improvement tomorrow by recording and registering as law the prejudices or errors of today. No man, however bold or however high he may stand in his profession, can resist the benumbing effect of rules laid down by authority. Devoted as I am to my profession, I see with fear and regret this tendency to legislate and rule."[3]

If, nevertheless, regulation is insisted upon, it should at least be done on a rational basis. This is covered in the next section.

14.3 A RATIONAL BASIS FOR SAFETY MEASURES

The rationale behind any measure designed to increase safety is the prolongation of life expectancy, but it must do so sufficiently to prevent life quality falling as a result of loss of income.

This provides the basis for a quantitative assessment of the value of safety measures, expressed as the judgment (J)-value,[4] defined as the quotient of the actual cost of the safety measure and the maximum amount that can be spent before the *life quality index* falls.

The life quality index Q (assuming that people value leisure more highly than work) is defined as (cf. 10.5)

$$Q = G^q X_\mathrm{d} \tag{14.1}$$

where G is average earnings (GDP per capita) from work, q is optimized work–life balance, defined as

$$q = w/(1-w) \tag{14.2}$$

where w is the optimized average fraction of time spent working ($q = 1/7$ seems to be typical for industrialized countries), and X_d is discounted life expectancy. Note that G^q has the form of a utility function: as pointed out by D. Bernoulli, initial earnings (spent on essentials) are valued more highly than later increments (spent on luxuries). Furthermore, money available now is valued more highly than that available tomorrow; a typical discount rate is 2.5% per annum.

An individual may choose to divert a portion of his income ΔG into a safety measure that will prolong his life by an amount ΔX. Assuming ΔG and ΔX are small, equation (14.1) yields

$$\Delta Q/Q = q\Delta G/G + \Delta X_\mathrm{d}/X_\mathrm{d}. \tag{14.3}$$

---

[3] L.T.C. Rolt, *Isambard Kingdom Brunel*, pp. 217–218. London: Longmans, Green & Co. (1957).

[4] P.J. Thomas, M.A. Stupples and M.A. Alghaffar, The extent of regulatory consensus on health and safety expenditure. Part 1: Development of the J-value technique and evaluation of regulators' recommendations. *Trans. IchemE, Part B* 84 (2006) 329–336.

Since it makes no sense to spend more on safety than the equivalent benefit in terms of life prolongation, the right-hand side of equation (14.3) should be equal to or greater than zero. The limiting case, equality, may be solved for ΔG and multiplied by the size of the population N benefiting from the measure to yield the maximum sensible safety spend

$$S_{\max} = -N\Delta G = (1/q)NG\Delta X_{\rm d}/X_{\rm d} \tag{14.4}$$

where the minus sign explicitly expresses the reduction in income. We can then write

$$J = S/S_{\max}. \tag{14.5}$$

Whether to proceed with a safety measure can therefore be decided on the basis of the J-value: if it is greater than 1, the expenditure S cannot be justified.[5]

14.4 SHOULD WE PROCEED?

The practical ethical question confronting the entrepreneur or the board of directors of a limited liability company is whether to proceed with some development of their activities. Let us travel back to the early Victorian era. One observes that "the engineers of the Industrial Revolution spent their whole energy on devising and superintending the removal of physical obstacles to society's welfare and development".[6] This, surely, was ethics to the highest degree. But a qualitative change subsequently occurred: "The elevation of society was lost sight of in a feverish desire to acquire money. Beneficial undertakings had been proved profitable; and it was now assumed that a business, so long as it was profitable, did not require to be proved beneficial."[7] And there we have remained to this day, it seems. Profit has become inextricably intertwined with benefit, but the former is no guide to the latter. The utilitarian principle (the greatest benefit to the greatest number) is only useful when two courses of action are being compared. The most important principle is *elevation of society*, which should be the primary criterion for deciding whether to proceed with any innovation, without even bothering about an

---

[5]Examples are given in P.J. Thomas, M.A. Stupples and M.A. Alghaffar, The extent of regulatory consensus on health and safety expenditure. Part 2: Applying the J-value technique to case studies across industries. *Trans. IchemE, Part B* 84 (2006) 337–343.

[6]A. Weir, *The Historical Basis of Modern Europe*, pp. 393–394. London: Swan, Sonnenschein, Lowrey (1886).

[7]A. Weir, loc. cit.

attempt to determine the degree of elevation—only the sign is important.[8] For Brunel, it was inconceivable that a railway engineer could have had anything other than the elevation of society in mind, hence the public was able to have total confidence in him, and he could be clear-minded in his opposition to regulators. Nowadays, we have to admit that this confidence is lacking (but not, one might hope, irremediably), and therefore society has built up an elaborate system of regulation, which seems, however, to have hampered the innovator while providing fresh opportunities for profit to individuals without benefit to society.

14.5 WHAT ABOUT NANOETHICS?

All that has been written so far in this chapter is generic, applicable to any human activity. How does nanotechnology fit in? Is there any difference between nanoethics and the ethics of the steam engine?

Possible reasons for according nanotechnology special attention are its pervasiveness, its invisibility (hence it can arrive without our knowledge), and the fact that it may be the technology that will usher in Kurzweil's singularity.

Pervasiveness scarcely needs special consideration. All sucessful technologies become pervasive—printing, electricity, the personal computer and now the internet.

The invisibility of modern technological achievement stands in sharp contrast to Victorian engineering, whose workings were generally open to see for all who cared to take an interest in such matters. Even in the first half of the 20th century, technical knowledge was widely disseminated, and a householder wishing to provide his family with a radio, or an electric bell, would be well able to make such things himself, as well as repair the engine of a motor-car. Nowadays, perhaps because of the impracticability of intervening with a soldering iron inside a malfunctioning laptop computer or cellphone, a far smaller fraction of users of such artefacts understand how even a single logic gate is constructed, or even the concept of representing information digitally, than, formerly, the fraction of telephone users (for example) who understood how the technology worked. And even if we do understand the principle, we are mostly powerless to intervene. Genetic engineering is, in

[8]It is a telling difference that it is quite typical nowadays for visionary and unexceptionably beneficial projects to be associated with the names of their business promoters, who may have nothing to do with the intellectual achievement of the innovation *per se*, the names of the engineers remaining unknown to the public, whereas in the Victorian age, the opposite was the case.

a sense, as familiar as the crossing of varieties known to any gardener, but few people in the developed world nowadays grow their own vegetables,[9] and there is mounting frustration at the unsatisfactory quality, in the culinary and gastronomic sense, of commercially available produce.

The answer to invisibility must, however, surely be a wider dissemination of technical knowledge. This should become as pervasive as basic literacy and numeracy. Hence it needs to be addressed in schools. It should be felt to be unacceptable that even basic concepts such as atoms or molecules are not held as widely as knowledge of words and numbers. Today they are not; to make "technical illiteracy" a thing of the past will require a revolution in eductional practice as far reaching as the Nano Revolution itself. To be sure, just as among the population there are different levels of literacy and numeracy, so we can expect that there will be different levels of technical literacy, but "blinding by science" should become far more difficult than it is at present.

Even without, or prior to the occurrence of, the singularity, productive nanosystems (PNs) imply an unprecedented level of technological assistance to human endeavor. All technologies tend to do this, and in consequence jobs and livelihoods are threatened—we have already mentioned Thimonnier's difficulties with tailors. The traditional response of engineers is that new jobs are created in other sectors. This may not be the case with PNs, in which case we can anticipate a dramatic shift of the "work–life balance" q in favor of leisure (see Section 14.3). Provided material challenges to human survival (Chapter 13) are overcome, finding worthwhile uses of this extra leisure remains as the principal personal and social challenge. Here it should be noted that as work–life balance shifts in favor of life (i.e., more leisure), implying decreasing q,[10] the marginal utility of money possessed saturates much more rapidly (equation 14.1). This may have profound social consequences (involving greed or its absence[11]), which have not hitherto been analyzed and which are difficult to predict.

The answer should, however, be available within the world of productive nanosystems (Section 12.1), with which everybody will be able to create his or her own personal environment. It represents the ultimate control over matter,

[9]Even if they did, they would find it difficult to procure seeds corresponding to their desires.

[10]Note that the simple equation (14.2) cannot be used directly here to compute q from a new w, because it deals with optimized quantities.

[11]J.J. Ramsden, Psychological, social, economic and political aspects of security. In: J.J. Ramsden and P.J. Kervalishvili (eds), *Complexity and Security*, pp. 351–368. Amsterdam: IOS Press (2008).

and does not depend on an élite corps active behind the scenes to maintain everything in working order. In this view, the age-old difference between the "sage" and the "common people" that is taken for granted in ancient writings such as the *Daodejing* should disappear. Is such a world, which each one of us can shape according to our interests and abilities, possible? That is at least as difficult a question to answer as whether productive nanosystems will be realized.

Even if they are, and even if we all become "shapers", will not our different ideas about shaping conflict with each other? Will there not be incompatible overlaps? That is presumably why we shall still need ethics. But human solidarity should be enhanced, not diminished, by more knowledge of the world around us, and it is that to which we should aspire. Anything less represents regression and loss. Let us proceed with nanotechnology in this spirit, not indeed knowing whither it will lead, but holding fast to the idea of elevation.

FURTHER READING

F. Allhoff, P. Lin, J. Moor and J. Weckert (eds), *Nanoethics*. Wiley (2008).

J.J. Ramsden, S. Aida and A. Kakabadse (eds), *Spiritual Motivation: New Thinking for Business and Management*. Basingstoke: Palgrave Macmillan (2007).

Index

S

T

U

V

W