SCALING ISSUES IN LARGE NETWORKS OF SMALL SENSORS: ENERGY AND COMMUNICATION MANAGEMENT

Mel Siegel

The Robotics Institute – School of Computer Science

Carnegie Mellon University, Pittsburgh PA 15213 USA

Emails: mws@cmu.edu

Abstract- We discuss five scaling-related issues in robotics and robotic systems. They are unified by an understanding of the fundamental engineering constraints imposed by the essentially fixed fundamental strength of materials and the essentially fixed practical density at which energy can be stored. These constraints have profound effects on when big bodies are in danger of collapsing under their own weight and when small bodies are in danger of running out of fuel. The section on strength provides a review of the lesson that small is strong, big is weak. The section on speed relates design and scale to inherent speed capability, including a discussion of why humanoid robots walk in unnatural-looking ways. The section on energy reinforces the realization that stored energy scales as the cube of a body's characteristic linear dimension, whereas its baseline power requirement usually scales as a lower power, e.g., the square of that dimension, so running time and range usually decrease cripplingly rapidly with size. The section on power shows how baseline power – and its dependence on speed – in combination with stored energy determines a body's running time and range. The section on communication discusses scale-related communication quality and duration, especially in systems that are simultaneously large and small, e.g., in networks composed of a very large number of very small robot-like nodes.

**Index terms*: scaling, robotics, strength, speed, energy, power, communication.*

I.   INTRODUCTION

laws of physics make no reference to any absolute time or absolute spatial coordinates: all other things being equal an experiment done tomorrow in Palmerston North can be reliably expected to give the same result as that experiment done yesterday in Pittsburgh. The number of natural invariances like these are so numerous and so emphasized in the teaching of the subject that students – and the public – often come to believe that "everything is relative". But in reality not

everything is relative. One feature of reality that is not relative is size or *scale*. Two structures or bodies in which every part is in the same linear proportion – *scale models* – are not equally strong, do not stay warm equally long when exposed to the cold cold night, and do not have the same lifetime or travelling range when running on stored energy. This paper is about the consequences in robotics of nature not being *scale invariant*. Much of our understanding of these matters comes from the long history in science of observing and modeling families of *geometrically similar* animals – like house cats, cougars, and lions. Thus in my presentation I will move around freely, without explicitly drawing the analogies, among examples from biomechanics [1], engineering, robotics, and everyday experience.

In 1959 Feynman presented a now famous lecture that was later published as *There's Plenty of Room at the Bottom: An Invitation to Enter a New Field of Physics* [2]. The optimism of this lecture is often regarded as having launched the current micro- and nano-scale revolutions.

In 1965 Feynman presented a now famous series of lectures that were later published as *The Character of Physical Law*. In the lecture titled *Symmetry in Physical Laws* [3] – after having in the previous lecture enumerated a host of nature's invariances – he tells us:

> of physics are symmetrical under any kind of change whatsoever, so now I will give a few that do not work. The first one is change of scale. It is not true that if you build an apparatus, then build another one, with every part made exactly the same, of the same kind of stuff, but twice as big, that it will work in exactly the same way .... You have probably seen ... that somebody has made a cathedral with matchsticks -- several floors, everything more Gothic than any Gothic cathedral has ever been, and more delicate. Why do we never build big cathedrals like that, with great logs, with the same degree of 'ginger cake', the same enormous degree of detail? The answer is that if we did build one it would be so high and so heavy that it would collapse ....This fact that the laws of physics were not unchanged under change of scale was first discovered by Galileo. In discussing the strength of rods and bones, he argued that if you need a bone for a bigger animal -- say an animal twice as high, wide, and thick -- you will have eight times the weight, so you need

a bone that can hold the strength eight times. But what a bone can hold depends on its cross-section, and if you made the bone twice as big it would only have four times the cross-section and would be only be able to support four times the weight ....

As described by Feynman and others, Galileo apparently considered his discovery that nature is not scale-invariant – and his development of a quantitative understanding of scaling laws – an achievement on a par with his invention of experimental physics. In fact, in Galileo's *Dialogs on Two New Sciences* [4], his discussion of scaling came first. Here is part of the dialog:

(Sagredo, the straight-man in this conversation) "... if a large machine be constructed in such a way that its parts bear to one another the same ratio as in a smaller one, and if the smaller is sufficiently strong ... the larger also should be able to withstand any severe and destructive tests to which it may be subjected ..." (Salviati, who straightens him out) "... the mere fact that it is matter makes the larger machine built of the same material not so strong ... the larger the machine the greater its weakness ... who does not know that a horse falling from a height of three or four cubits will break his bones, while a dog falling from the same height or a cat from a height of eight or ten cubits will suffer no injury? ..."

I added the italics to emphasize that the issue is that when you change the scale of a structure or a body you may imagine that you also correspondingly change the strength of the materials of which it is constructed, but in reality nature only give you only   one kind of stuff with which to build.

I.a STRENGTH

Although there is considerable – and still growing – interest in big robots – from large earth-moving machines to enormous dynamic precision structures like radio telescopes and space stations – they are never built without benefit of the expertise of mechanical engineers who know how to design structures that support their own weight and more. This is not to say these large structures never collapse anyway. Occasionally a building or a bridge falls without obvious environmental assault, and at least one 100-meter-diameter radio telescope collapsed spontaneously and spectacularly [5]. But these are unusual exceptions to the rule that big machines are expertly designed and their designs are expertly executed. Small robots, on the

other hand, are often designed by enthusiastic researchers who have innovative agendas but no experience in taking a design concept from tabletop model scale to the nanoscale implementation that is their goal. It is easy to make small things strong, so these designs are usually static successes. Some current milli-, micro-, and nano-scale devices are illustrated in Figure 1.

They are static successes but most to date are dynamic failures: they don't run as long as expected because on the microscale it is impossible to carry along enough of your own energy to do much of anything useful, and they don't move as expected because on the nanoscale viscosity is so powerful as to make gravity invisible. These difficulties, approaches to understanding them quantitatively, and approaches to living with the limits that nature imposes are my main topic.
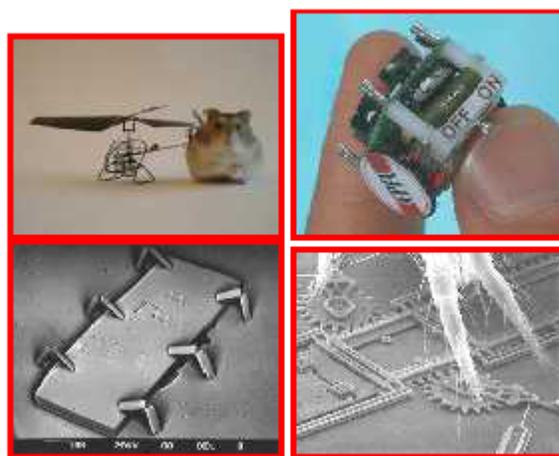

Figure 1: Mouse-Sized Helicopter; Watch-Motor Based Wheeled Robot; Six-Legged Silicon Bug; Silicon Rack-And-Pinion Compared With The Hairs On The Hairs Of A Spider's Foot.

I.b SPEED

Our expectations about how fast bodies should move depend strongly on context. In the lab we tend to believe that speed should be scaled to size: two mobile robots whose speed is the same in body lengths per unit time are perceived to have the same speed. On the highway we expect all vehicles to have the same top speed and to cruise at the speed limit. There are important intermediate cases: we will see that humanoid walking robots that use energy efficiently have cruising speeds that increase as the square root of height h: the smaller they get the faster

their legs move but the more slowly they go. This is true even when speed is measured relative to height: the pendulous step frequency of energy-efficient walking or running scales as $h^{-1/2}$, step length scales as h, so absolute speed scales as $h^{1/2}$, and speed relative to height scales as $h^{-1/2}$.

I.c SPEED

How much energy a body can carry is proportional to its volume, i.e., the cube of a characteristic linear dimension like its height h. How much energy it needs – either per unit time or to complete a particular task – depends on its scale, its metabolism, the speed at which it moves, etc. It almost always turns out that smaller bodies need more energy per unit mass than larger ones.

I.d POWER

The almost trivial realization that running time equals stored energy divided by power demand and range isrunning time times speed lead to deep understandings of the inherent practical limitations of small mechanical systems – and how to think about alternative approaches that will allow us to work around these limitations so as to fulfill the promises of nanotechnology.

I.e COMMUNICATION

We are particularly interested in communication issues in networks composed of a very large number of very small robot-like nodes. Smaller nodes consume less power, but as noted, they have even less capacity to store energy, so they have correspondingly shorter lifetimes. Smaller size also means lower efficiency as an antenna for the longer radio wavelengths that may be preferred for communication on the grounds that they correspond to lower energy per photon.

I.f ORGANIZATION OF THE PAPER

The paper exists in two versions, the complete (16 page) version as a link off the author's home page http://www-2.cs.cmu.edu/~mws and this short (8 page) version for the proceedings of the International Conference on Sensor Technology (ICST) 2005 held at Massey University,

Palmerston North, New Zealand, 2005 November 21-23. Sections that expand in detail the above subsections 1.a (in section 2, *strength*), 1.b (in section 3, *speed*), 1.c (in section 4, *energy*), and 1.d (in section 5, *power*) are included in the long version, but except for section 5.e, describing a fundamental problem for nano-robotics, these sections are omitted from the short version; the section that expands on section 1.e, (in section 6, *communication*), is presented in its entirety in the short version. The omitted material has been presented previously, but not in as much detail or with as comprehensive a set of application scenarios as are described and analyzed in the long version. The material in the section on communication is about half new, and the remaining half has been presented previously only in very abbreviated form.

5.e A PROBLEM FOR NANO-ROBOTICS

It is hard to imagine scenarios in which the required power per unit mass decreases when characteristic linear dimension decreases. On the contrary, most plausible models conclude that power demand per unit mass is inversely proportional to characteristic linear dimension h, making running time proportional to h In most scenarios the required power is also proportional to a power of the speed v^n. For rolling wheel resistance range is independent of speed, but for most other scenarios range drops rapidly with speed – even if we mentally scale our expectations so we are satisfied with speed proportional to body size. It is thus apparent that at some small scale, no matter how clever energy and power engineering become, nano-robots will need to receive or extract energy from their environments, as it will become impossible for them to carry enough for any mission of useful duration or distance. Since how small a body has to be before it qualifies as a nano-robot is ill defined, I propose that a nano-robot be defined as a robot that is so small that it cannot carry enough energy to do a useful job.

Plausible alternatives to carrying all the energy needed to compete a task include:
- A physical or a virtual umbilicus, e.g., collect natural or beamed in light, RF energy, etc.
- Extract chemical energy from the environment, e.g.,as microbes do.
- Extract wind energy from gusts and thermal updrafts, e.g., as birds do.
- Build a Maxwell's Demon, e.g., extract energy from macroscopic thermal gradients (which does not violate the second law of thermodynamics) or from microscopic thermal

fluctuations (which probably does violate the second law, though there is some speculation that it could be legal given sufficiently intelligent sensing and actuation).

Some research groups, e.g., the builders of the EcoBot I and II at University of West England, have demonstrated robots with a rudimentary ability to forage for chemical energy, e.g., by attracting, catching, digesting, and extracting electrical energy from flies [9].

VI COMMUNICATION

In this section we discuss scaling issues in communication, particularly in systems that are large in the sense that they consist of a very large number of network nodes, yet simultaneously small in the sense that each node is a micro- or nano-scale device.

VI.a GLOBAL ENVIRONMENTAL MONITORING

As a concrete example of this scenario we describe a Global Environmental Monitoring System (GEMS). GEMS is envisioned as a large network of small sensor nodes that would be sprinkled throughout the earth's atmosphere to gather location, temperature, pressure, humidity, and ancillary data whose primary use would be to initialize global-scale weather prediction models. Its goal is to sample the atmosphere at a resolution of one node per km3 over the surface of the earth to an altitude of 20 km. The number of nodes needed for this resolution is about 10^{10}. To the extent that nodes are effectively carried by the wind, wind velocity would be inferred from consecutive time and position reports. If the ultimate node design results in aerodynamics that do not assure node motion to be essentially the same as wind motion it would be necessary also to sense and explicitly to report wind speed and direction.

VI.a.i MOTIVATION

The ability to predict the weather at any location and any future time is of enormous social, economic, political and military value. Agriculture, industry, and transportation are all substantially influenced by the weather; a small but statistically significant connection between

stock prices and the weather at the location of the stock market was demonstrated recently. This effect probably relates to investors' feelings of physical and emotional wellness, which are well know to be affected by weather. So are the outcome of elections and military operations.

The cost of ignorance is high. Prior to three enormous Caribbean hurricanes recently making landfall in the southeast US (Katrina, Rita, and, as this is being written, Wilma) it was estimated that an error of 1 hr or 1 km in predicting a hurricane's landfall might cost $10 million in unnecessary preparation in one place and inadequate preparation in another. Based on the two recent catastrophes and the apparently impending one, this figure now appears to be lower than reality by a factor of ten or more.

The state-of-the-art in weather prediction is impressive, and the increasing availability of relatively low cost supercomputing resources promises to make it rapidly more impressive. Dynamical models for the evolution of large-scale weather patterns are now comprehensive and accurate. Within a few years it is expected that these codes will run usefully faster-than-real-time on supercomputers that will be able to handle 10^{10} grid points. However data to initialize the grid will then needed at the same resolution. Those initialization data are the object of the GEMS.

VI.a.ii FEATURES

GEMS is envisioned as a self-organizing network of sensing and communication nodes that are scattered throughout the atmosphere, e.g., by high-flying aircraft. The nodes fall to earth in a few hours to a few weeks, depending on their aerodynamics, primarily their size, and the extent to which aerodynamic control features might be incorporated in their design. As is typical in networks of this sort, data are passed from node-to-node, and eventually to one or more base stations, perhaps via one out of every thousand or so nodes having super-node status. For the present discussion we will consider in detail only the peer-to-peer communication portion of the system. Some variants of the model also incorporate the possibility that the density of nodes will be nonuniform so as to accommodate higher resolution sampling, e.g., in active storm areas, and lower resolution sampling, e.g., over vast ocean areas with negligible place-to-place variation.

VI.a.iii GEMS NODES ARE ROBOTS

What exactly constitutes a robot is subject to debate, but almost everyone accepts the sense-think-act paradigm – to which I think it is essential that we now add *communicate*. GEMS nodes easily meet criteria based on this augmented paradigm:

- Sense: environmental sensing, proprioception, and location sensing.
- Think: processing for data collection and reduction, power and communication management, and possibly mobility control.
- Act: transmission of data and reception of commands via the physical communication layer, and possibly position and orientation control, and aerodynamic re-configuration.
- Communicate: receive commands, send data, aggregate data from nearby nodes.

VI.a.iii TECHNOLOGICAL CHALLENGES

Table 1 enumerates the key issues and enabling technologies affecting the design and development of GEMS. These represent individually and collectively an enormous challenge to technology and engineering, but none of the fundamental scientific questions are in doubt.

These are primarily system-level issues. At the node level the challenges are expanded:

- Size: if each node uses only 1 mm3 of finished silicon, 1010 nodes require 10 m3. While this may not seem like much, current annual world production of 300 mm wafers is estimated at only 200 m3, so the GEMS project would constitute 4% of a year's total production capacity – enough to affect the world price.
- Residence time: longer residence time, i.e., slower fall, is generally to be preferred; terminal velocity scales as surface area, i.e., h2, so small nodes are substantially better.
- Efficient communication: for energy efficiency, atmospheric transmission and (possibly) directionality, radio frequency transmission is preferred; but efficient antennas are of the same scale as the wavelength they transmit/receive, so bigger is better.

Given these conflicting issues, a practical GEMS node obviously will have to be a design compromise in a multi-dimensional optimization space.

Table 1: Key issues and enabling technologies affecting the design and development of GEMS

VI.b POWER REQUIRED TO COMMUNICATE

Now we address how much power a node must transmit for how long for its message to be received uncorrupted a nearby node. We will examine the issues from two perspectives. In the first perspective, rather than starting from first principles, we will bootstrap an engineering estimate onto the known current adequacy of the GPS system. The data rate that a GEMS node will need to transmit is comparable to the data rate that a GPS receiver needs to receive, so the comparison is essentially geometrical, based on the different distances and solid angles involved. In the second perspective, we will start from first principles and form best-case lifetime estimates based on the energy actually transmitted. The key engineering assumption in the case of the first perspective is that by the time a GEMS network can be deployed the capability of today's hand-held GPS will be available on a microchip– the antenna excepted, of course. The key scientific assumption in the case of the second perspective is that it is reasonable to base this sort of estimate on a photon shot-noise limited model.

VI.b.i GPS BOOTSTRAP

The altitude of a GPS satellite R_{GPS} is approximately 19,000 km. The transmitted power per channel P_{GPS} is approximately 20 W. We will assume the transmitter footprint covers the earth uniformly, so the transmitter power is concentrated into a solid angle Ω_{GPS} given approximately by the square of the ratio of the earth diameter to the sum of R_{GPS} and the earth radius. We assume the GEMS network nodes are deployed in the atmosphere at a density of $N\ km^{-3}$, so the average node separation is

$$R_N \approx \sqrt{3} N^{-1/3} km \qquad (1)$$

We will consider node transmission solid angles ΩN both very small and very large compared to Ω_{GPS}. Very small will require sensing and actuation to achieve active pointing of the node's transmitting antenna. Very large essentially means omnidirectional, which requires much more power but no knowledge of where to point and demands no ability to achieve pointing. We can safely assume the future GEMS node receiver sensitivity is approximately the same as today's GPS receiver sensitivity, so we equate the received power density at the two receivers:

$$\frac{P_{GPS}}{\Omega_{GPS} R_{GPS}^2} \approx \frac{P_N}{\Omega_N R_N^2} \qquad (2)$$

Solving for P_N and substituting R_N from Eq. 20 gives

$$P_N \approx 3 P_{GPS} \frac{\Omega_N}{\Omega_{GPS} R_{GPS}^2 N^{2/3}} \qquad (3)$$

Now we ask how large a node is needed to support power P_N for a useful time. We assume an energy density W_N of about 0.5 J/mm^3, consistent with the highest energy density available in the batteries illustrated in **Error! Reference source not found.**. If the power budget is dominated by the transmitter then the node's lifetime as a function of its volume V_N is:

$$t_N \approx \frac{W_N V_N}{P_N} \approx W_N V_N N^{2/3} R_{GPS}^2 \frac{\Omega_{GPS}}{\Omega_N} \frac{1}{3 P_{GPS}} \qquad (4)$$

Consider first a moderately large "micro-node" with $V_N \approx 1 mm^3$, a moderately large density of nodes N $\approx$ 1000 km^{-3}, and a tightly collimated transmitted beam of angular size 0.01 *radian*. Using these values and other stated previously Equation 23 evaluates to a node operating time of $t_N \approx 5 \times 10^{16}$ *s*. So in this scenario the on-board power lasts forever.

Now consider a more ambitious scenario in which the sizes are compatible with the GEMS scenario. This time consider a "nano-node" with $V_N \approx 1 \mu m^3$ and a node density $N \approx 1$ km^{-3}. We now obtain $t_N \approx 5 \times 10^5$ s, about 5 days. That is encouraging, but remember it assumes a tightly collimated beam and thus demands precise sensing and actuation capability. If we instead assume an omnidirectional transmitter, $\Omega_N = 4\Pi$, then $t_N \approx 400$ *s*, or about 7 minutes.

So we see that the feasibility of carrying enough energy to support the transmitted power depends

strongly on the assumptions we make about size and antenna pointing ability. We remind the reader that this back-of-the-envelope estimate completely ignores the power needed to support sensing, processing, receiving, and whatever actuation might be incorporated in the design, e.g., to achieve the pointing capability that seems to be essential to operating a feasible transmitter power for a useful time.

VI.b.ii FIRST-PRINCIPLES APPROACH

Now we briefly outline an alternative first-principles approach to optimization of communication between network nodes when total network mass – how much silicon we are allowed to use – is the essential constraint. This is actually a realistic constraint: as noted above, a GEMS network of 10^{10} nodes would consume a significant fraction of the world's current annual production of finished silicon even if each node required only *1 mm^3*.

We begin along the same lines followed in the previous section. If we deploy N nodes in a volume V_{atm} of the atmosphere then the node spacing is approximately $RN \approx N^{-1/3}$. If the total mass M of all nodes is fixed then $h^3 \approx M/N \Rightarrow h \infty N^{1/3}$. So in this capacity-limited model $h \infty R_N$.

For communication to be efficient the antenna size h must be comparable to the communication wavelength λ, i.e., $h \approx \lambda$. If this requirement is satisfied then

$$P_{receive} \approx P_{transmit} \frac{h^2}{R^2} \qquad (5)$$

Transmitted power $P_{transmit}$ is the product of the number of transmitted photons per second *dn/dt* and the energy per photon. The energy per photon in inversely proportional to wavelength. So

$$P_{transmit} \propto \frac{1}{\lambda} \frac{dn}{dt} \qquad (6)$$

Calling the transmitter on-time τ, integrating Eq. 25 gives

$$n \propto \lambda P_{transmit} \tau \qquad (7)$$

$P_{transmit} \tau$ is the energy stored on board, hence proportional to h^3, and λ is proportional to h, so the number of photons we can send from a node of characteristic size h is proportional to h^4:

$$n \propto h^4 \qquad\qquad\qquad (8)$$

The integral signal-to-noise ratio is thus $n/\Delta n = n^{1/2} \propto h2 \propto N\text{-}2/3$.

Smaller nodes mean we can deploy more nodes closer together, so the 1/r2 decrease of signal power density with distance is less severe than in the case of fewer but larger and further separated nodes. On the other hand, smaller nodes are less efficient antennas for long wavelengths, and employing a shorter wavelength costs more energy per photon.

From this simple analysis we can conclude that if constrained to use a fixed quantity of material, we can build more but smaller nodes and get more data at the price of lower signal-to-noise ratio. To a carefully optimized point doing this would probably be a good policy, since higher spatial sampling density is probably more useful than locally more precise but less densely sampled data.

VII. CONCLUSIONS AND FUTURE ISSUES

There are some simple conclusions:

- If you want to travel a long distance or for a long duration and you have to carry your own food, fuel, batteries, etc., then you need a big body. This is inconsistent with sending out a fleet or horde of small bodies, e.g., in robotic scenarios modelled on ant colonies.
- Making this qualitative point quantitative is usually simple:
- the energy you can carry scales as your linear dimension h cubed;
- your operating time at baseline power P thus scales as h^3/P;
- your operating range then scales as h^3v/P, where v is your speed;
- your challenge is to find the right functional form for P in your application.
- Below some size, untethered robots inevitably must forage for energy.
- Big networks composed of many small nodes present complex communication and control optimization challenges with many tradeoffs that must be balanced.

Scale can be a hindrance or a help. For vehicles and mobile robots bigger is usually better – at least until they get so big that the strength of materials becomes an issue. For sensors and instruments smaller is usually better, both because smaller means lower power and smaller weight carrying requirements, but often also for fundamental sensitivity and precision reasons.

The same factor may hinder or help depending on the specific application; for example, working in a liquid environment vs. an air or vacuum environment makes it easier to suspend network nodes, but harder for nodes to communicate with each other or with a base station.

Critical future needs can be identified as:

- energy sources for milli-, micro-, and nano-robots, both on board and via umbilicuses;

- investigation of advantageous shapes for atmospheric nodes, e.g., line-like vs. sphere-like shapes are advantageous for both mobility ("sailing") and communication (antennas);

- active mechanisms and closed-loop control strategies for antenna pointing;

- modelling the challenges of extreme environments, e.g., oceans, space, temperature;

- understanding the potential environmental impact of environmental monitoring networks.

VIII ACKNOWLEDGEMENT

REFERENCES

[1] C. J. Pennycuick, Newton Rules Biology: A physical approach to biological problems, Oxford University Press, 1992

[2] Richard P. Feynman, There's Plenty of Room at the Bottom: an Invitation to Enter a New Field of Physics, http://www.zyvex.com/nanotech/feynman.html

[3] Richard P. Feynman, *The Character of Physical Law*, chapter titled *Symmetry in Physical Laws*, pp. 95-96, (MIT Press, Fourth Paperback Printing, August 1971)

[4] Galileo*, Dialogues Concerning Two New Sciences*, translated by Henry Crew and Alfonso di Salvio, Prometheus Books, 1991. ISBN 0879757078.

[5] National Radio Astronomy Observatory, Green Bank, West Virginia, USA, 1988 November 15. See http://www.gb.nrao.edu/fgdocs/300ft/300ft.html.

[6] ASIMO's mechanical, etc., specifications are documented for several of the still-evolving versions of the robot on several of the Sony website pages. A compact summary is posted at http://www.answers.com/topic/asimo, which crossreferences a Wikipedia page.

[7] Many websites provide typical height and weight tables and graphs, and normal ranges, for children from birth to adulthood. A exemplary Growth Chart of Height and Weight is at http://www.teengrowth.com/index.cfm?action=development&drill=viewArt&art=11.

[8] Grant Imahara, personal communication, 2203 August 2003. See animatronic achievements at website http://www.mythbustersfanclub.com/html/grant_imahara.html.

[9] C. Blythsway and I. Gilhespy, links to several published articles on EcoBot I and Ecobot II. See http://www.ias.uwe.ac.uk/Energy-AutonomyNew/New Scientist - EcoBot II.htm.