

CHAPTER 5

REPLICATING SYSTEMS CONCEPTS: SELF-REPLICATING LUNAR FACTORY AND DEMONSTRATION

5.1 Introduction

As the cost of fossil-fuel energy continues to escalate and supplies of readily accessible high-grade ores and minerals gradually become depleted, the utilization of non-terrestrial sources of energy and materials and the development of a nonterrestrial industrial capacity become increasingly desirable. The Moon offers plentiful supplies of important minerals and has a number of advantages for manufacturing which make it an attractive candidate factory site compared to Earth. Given the expense and danger associated with the use of human workers in such a remote location, the production environment of a lunar manufacturing facility should be automated to the highest degree feasible. The facility ought also to be flexible, so that its product stream is easily modified by remote control and requires a minimum of human tending. However, sooner or later the factory must exhaust local mineral resources and fall into disrepair or become obsolete or unsuitable for changing human requirements. This will necessitate either replacement or overhaul, again requiring the presence of human construction workers with the associated high costs and physical hazards of such work.

The Replicating Systems Concepts Team proposes that this cycle of repeated construction may possibly be largely eliminated by designing the factory as an automated, multi-product, remotely controlled, reprogrammable Lunar Manufacturing Facility (LMF) capable of constructing

duplicates of itself which would themselves be capable of further replication. Successive new systems need not be exact copies of the original, but could, by remote design and control, be improved, reorganized, or enlarged so as to reflect changing human requirements. A few of the benefits of a replicative growing lunar manufacturing facility (discussed at greater length in secs. 5.4 and 5.5) include:

- (1) The process of LMF design will lead to the development of highly sophisticated automated processing and assembly technologies. These could be used on Earth to further enhance human productivity and could lead to the emergence of novel forms of large-scale industrial organization and control.
- (2) The self-replicating LMF can augment global industrial production without adding to the burden on Earth's limited energy and natural resources.
- (3) An autonomous, growing LMF could, unaided, construct additional production machinery, thus increasing its own output capacity. By replicating, it enlarges these capabilities at an increasing rate since new production machinery as well as machines to make new machines can be constructed.
- (4) The initial LMF may be viewed as the first step in a demonstration-development scenario leading to an indefinite process of automated exploration and utilization of nonterrestrial resources. (See fig. 5.1.) Replicating factories should be able to achieve a very general manufacturing

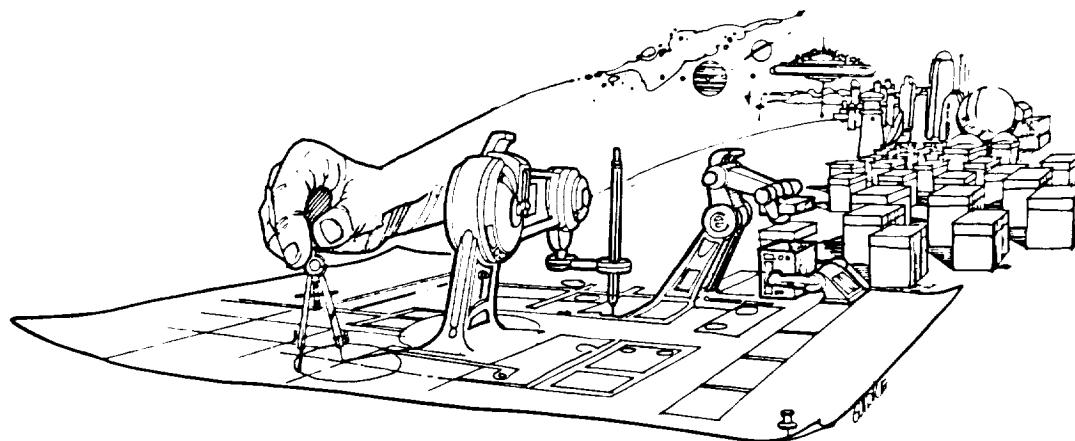


Figure 5.1. – Automated space exploration and industrialization using self-replicating systems.

capability including such products as space probes, planetary landers, and transportable “seed” factories for siting on the surfaces of other worlds. A major benefit of replicating systems is that they will permit extensive exploration and utilization of space without straining Earth’s resources.

5.1.1 Summary of Chapter Contents

The history of the concept of machine replication is reviewed in section 5.2. This theoretical background is largely a consideration of the work of John von Neumann – in particular, his kinematic and cellular models of automata self-reproduction. Post-von Neumann research is reviewed next, noting particularly the established theoretical capabilities of machines in the realm of general construction, inspection, and repair strategies. Such strategies may prove useful, even vital, to the successful design, realization, and operation of actual replicating systems.

Section 5.3 deals with the engineering feasibility of the concept of self-replicating systems (SRS). An attempt is made to confront two important general problems in creating a lunar replicating factory:

- Given that in theory, machines can construct duplicates of themselves, how might systems designers and engineers identify all functions which must be carried out to achieve machine replication and also develop the technological means by which to implement these functions?
- Given the constraints obtaining in the lunar environment, particularly in terms of the inventory of known kinds and quantities of naturally occurring raw materials and the existing repertoire of materials processing technologies, can all machine functions required both for production and for replication and growth be implemented?

To attack the first of these problems – identification of necessary functions for practical machine replication – the team proposes a specific phased demonstration-development scenario, described in section 5.3. For the second problem – establishing that machine replication can feasibly take place in the actual lunar environment – a strawman mission concept was employed. In this scenario, a 100-ton initial “seed” factory is planted on the Moon with access only to local resources and established materials processing techniques. The initial system should be able to successfully develop into an expanded machine system capable of conducting all functions necessary for autonomous replication, growth, and automated production and manufacturing.

The problem of “closure” is also considered at length in section 5.3. The issue of closure is whether autonomous manufacturing and construction systems can make available to themselves all of the materials, parts, and assembly techniques required for all internal operations. An iterative strategy is presented for detecting and eliminating closure gaps, and for optimizing the resulting augmented system.

Section 5.4 deals with possible applications of the SRS concept. Applications of replication technology include enormous gains in terrestrial industrial productivity (automation and computer-aided design and manufacturing), utilization of Solar System resources, orbital and planetary opportunities, and the possibility of interstellar exploration on a grand scale. Indefinitely large masses can be organized in extraterrestrial environments using self-replicating systems.

Section 5.5 deals with just a few of the many implications of SRS. The advantages of space-based replicative manufacturing are considered, together with possible political, social, economic, cultural, and psychological consequences of the proposed SRS development program.

Section 5.6 sets forth in some detail how NASA can take action at once toward the achievement of the ultimate goal of establishing a replicating manufacturing facility. Suggested statements of work (SOWs) and a listing of institutions that might undertake the tasks outlined in the work statements are included. A series of specific conclusions and recommendations generated by the Replicating Systems Concepts Team are presented in section 5.7.

5.2 Theoretical Background

The notion of a machine reproducing itself has great intrinsic interest and invariably elicits a considerable range of responses – some directed toward proving the impossibility of the process, others claiming that it can be carried out, but almost all of them indicating an unwillingness to subject the question to a thorough examination. In discussing self-replication by automata it is essential to establish early rather important ground rules for the discussion. According to Kemeny (1955), “If [by ‘reproduction’] we mean the creation of an object like the original out of nothing, then no machine can reproduce – but neither can a human being....The characteristic feature of the reproduction of life is that the living organism can create a new organism like itself out of inert matter surrounding it.”

Often it is asserted that only biological organisms can reproduce themselves. Thus, by definition, machines cannot carry out the process. On the other hand, others argue that all living organisms are machines and thus the proof of machine reproduction is the biosphere of Earth. Also, sometimes it is claimed that although machines can produce other machines, they can only produce machines less complex than themselves. This “necessary degeneracy” of the machine construction process implies that a machine can never make a machine as good as itself. An automated assembly line can make an automobile, it is said, but no number of automobiles will ever be able to construct an assembly line.

Another common argument is that for a machine to make a duplicate copy it must employ a description of itself. This description, being a part of the original machine, must itself be described and contained within the original machine, and so on, until it is apparent we are forced into

an infinite regress. A variant of this is the contention that a machine not possessing such a description of itself would have to use itself for a description, thus must have the means to perceive itself to obtain the description. But then what about the part of the machine that does the perceiving? It cannot perceive itself, hence could never complete the inspection needed to acquire a complete description. (A simple counter is that the original machine might possess multiple perceiving organs, so that the perceiving could be shared.) Yet another related objection is that for the process to be carried out, the machine must come to “comprehend” itself – at which point it is said to be well known that “the part cannot possibly comprehend the whole.” These disputation suggest that there is a very deep-seated resistance to the notion of machines reproducing themselves, as well as an admittedly strong fascination with the concept.

The Hungarian-American mathematician John von Neumann (1966), who first seriously came to grips with the problem of machine reproduction, once noted that it would be easy to make the whole problem go away. One could, for example, make the elementary parts of which the offspring machine was to be composed so complex as to render the problem of replication trivial. In one example of this considered by the team, a robot required only to insert a fuse in another similar robot to make a duplicate of itself would find “reproduction” very simple (see sec. 5.2.3). As von Neumann also pointed out, it is equally useless to go to the other extreme and try to account for the placement of every atomic particle in the system – one would quickly become mired in incomprehensible detail. Even most lifeforms do not have DNA-encoded instructions for reproduction to this fantastic level of detail – their descriptions are largely at the molecular level.

As will be demonstrated presently, although reproduction may be transparently trivialized or intractably complexified, there appear to be no fundamental inconsistencies or insoluble paradoxes associated with the concept of self-replicating machines.

5.2.1 Von Neumann's Contributions and Subsequent Research

John von Neumann began studying automata replication because he was interested in very complex machines and their behaviors. The early history of the theory of reproducing machines is basically the history of von Neumann's thinking on the matter, and this is reviewed below.

Von Neumann had a tremendous range of interests – he contributed to the logical foundations of quantum theory, was the co-inventor of the theory of games, and he worked on the Manhattan Project (contributing to the design of the implosion mechanism for the plutonium bomb). It is believed that his participation in the Manhattan Project and the tremendous volume of calculations necessary for bomb

design led him into automatic computing. Hearing of the ENIAC computer project at the Moore School of Electrical Engineering at the University of Pennsylvania, von Neumann was fascinated by the potential of a computer very much faster than any of the devices that had previously been produced. In the early 1940s there existed only simple relay machines and analog devices such as the differential analyzer. But the new electronic machines that interested von Neumann promised to be perhaps millions of times faster than relay machines.

So von Neumann immersed himself in the ENIAC project, the first electronic computer program where some actual useful computing was produced. Late in 1945 and early 1946, the first problems that were put on ENIAC are believed to have been calculations involving the feasibility of a hydrogen bomb. Von Neumann, although he remained very much interested in nuclear energy and was appointed a member of the Atomic Energy Commission, became fascinated with the idea of large and complex computing machines. He devised the organization employed today in almost all general purpose computational machines – the so-called von Neumann concept of serial processing stored-program or the “von Neumann machine.” After that work was completed he began thinking seriously about the problems of extremely large machines – their reliability, programming, design, how to understand what they do – and he became involved with the many possible analogies to the complex behavior of living systems.

Von Neumann set for himself the goal of showing what the logical organization of a self-reproducing machine might be. He had in mind a full range of self-replicating machine models which he intended to explore, including the (a) kinematic machine, (b) cellular machine, (c) neuron-type machine, (d) continuous machine, and (e) probabilistic machine. As it turned out, he ultimately was only able to produce a very informal description of the kinematic machine. Although he wrote a great deal on the cellular machine, his *magnum opus* on the subject was left in the form of unfinished notes at the time of his death. Almost no work was done on the other three kinds of self-reproducing machines. For this reason, only the postulated workings of the kinematic and cellular machines are presented below, with brief comments on the other three types. For an additional review of these two models of reproduction, see Burks (1970).

In dealing with machines that could reproduce, von Neumann concluded that the following characteristics and capabilities should be demonstrable for each:

(1) Logical universality – the ability to function as a general-purpose computing machine able to simulate a universal Turing machine (Turing, 1936). This was necessary because SRS must be able to read instructions to carry out complex computations.

(2) Construction capability – to self-replicate, a machine must be capable of manipulating information,

energy, and materials of the same sort of which it itself is composed.

(3) Constructional universality – in parallel to logical universality, constructional universality implies the ability to manufacture any of the finitely sized machines which can be formed from specific kinds of parts, given a finite number of different kinds of parts but an indefinitely large supply of parts of each kind.

(4) Self-reproduction – follows immediately from the above, since the universal constructor must be constructable from the set of manufacturable parts. If the original machine is made of these parts, and it is a constructable machine, and the universal constructor is given a description of itself, it ought to be able to make more copies of itself.

Von Neumann formally demonstrated that his cellular model of reproduction possessed these four properties.

Not much was done on a fifth property also believed to be important – evolution – though there have been some more recent results in this area. If one has a machine, and it makes a machine, which then itself makes a machine, is there any proof that the line of machines can become successively “better” in some fashion – for instance more efficient, or able to do more things? Could they evolve to higher and higher forms? This problem raises issues in learning, adaptation, and so forth, and was left largely untouched by von Neumann.

The kinematic machine. The kinetic machine is the one people hear about most often in connection with von Neumann’s work on self-reproducing machines, probably because it received the earliest attention and publicity. John Kemeny (1955) produced a paper for the popular publication *Scientific American* detailing this model, and a further description appeared in a paper by von Neumann (1951).

The notion of kinematic machine self-reproduction was dealt with by von Neumann only informally. The mathematician envisioned a machine residing in a “sea” of spare parts. The machine has a memory tape which instructs it to go through certain mechanical procedures. Using a manipulative appendage and the ability to move around in its environment, the device can assimilate and connect parts. The tape-program first instructs the machine to reach out and pick up a part, then to go through an identification routine to determine whether the part selected is or is not the specific one called for by the instruction tape. If not, the component is thrown back into the “sea” and another is withdrawn for similar testing, and so on, until the correct one is found. Having identified a required part the device searches in like manner for the next, then joins the two together in accordance with instructions.

The machine continues following the instructions to make something, without really understanding what it is doing. When it finishes it has produced a physical duplicate

of itself. Still, the second machine does not yet have any instructions so the parent machine copies its own memory tape onto the blank of its offspring. The last instruction on the parent machine’s tape is to activate the tape of its progeny.

Von Neumann’s logical organization for a kinematic machine is not the only one possible, but probably is the simplest way to achieve machine self-replication. In its logic it is very close to the way living organisms seem to reproduce themselves (Dyson, 1979). One conceptual problem with the model is that the parts involved are supplied free to the machine, and those parts are of a relatively high order. The machine dwells in a universe which supplies precisely the sorts of things it needs as a kinematic device to make a duplicate of itself. This raises the issue of closure, a problem which is discussed and conceptually resolved in section 5.3.

The cellular model. Von Neumann evidently was dissatisfied with his original kinematic model because of its seemingly mathematical inelegance. This model of machine self-reproduction, while qualitatively sound, appeared not easily susceptible to mathematically rigorous treatment and so might not serve to convince a determined skeptic.

Stan Ulam, a Polish-American mathematician who had also worked on the Manhattan Project, suggested to von Neumann that the notion of a self-reproducing machine would be amenable to rigorous treatment if it could be described in a “cell space” format – a geometrical grid or tessellation, regular in all dimensions. Within each cell of this system resides a finite state automaton. These cell automata can only be affected by certain of their neighbors, and only in very specific ways. In the model von Neumann finally conceived, a checkerboard system is employed with an identical finite state automaton in each square (fig. 5.2). In this system, as it evolved with subsequent research, the cell-automata can be in one of 29 possible different states (fig. 5.3). Each automaton can communicate with its four cardinal direction neighbors. The state of a cell-automaton is determined by its own state and by the states of its cardinal direction neighbors.

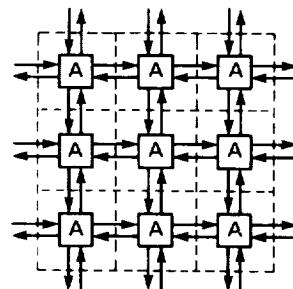


Figure 5.2.— Finite state automation cellular space.

	U			
UNEXCITABLE	→	↑	←	↓
ORDINARY TRANSMISSION	•→	•↑	•←	•↓
SPECIAL TRANSMISSION	⇒	↑↑	⇐	↓↓
CONFLUENT	⇒•	•↑↑	⇐•	•↓↓
SENSITIZED	C ₀₀	C ₀₁	C ₁₀	C ₁₁
	S ₀	S ₀	S ₁	S ₀₀
	S ₀₁	S ₁₀	S ₁₁	S ₀₀₀

Figure 5.3.— Twenty-nine states of von Neumann's cellular automata.

At the beginning of operation, all but a finite number of the cell automata are in a "U" or "unexcitable" state. If a given cell is in the "U" state, and all its neighbors also are in the "U" state, then at the next moment of time, the given cell remains in the "U" state. Thus the "U" states can be viewed as representing undifferentiated, passive underlying substrate. Their passivity implies that they may in some cases serve as "insulation" surrounding more active cells in the system.

Then there are "ordinary transmission" cell states. These are states which direct their activity in each of the four cardinal directions. Each of these may be in an excited or quiescent mode, so there is a total of eight different kinds of ordinary transmission states. In addition, there are eight "special transmission states," similar to the ordinary states in that they also point in each of the cardinal directions and can be in excited or quiescent modes. The two basic kinds of transmission states — ordinary and special — differ in that the primary intended role of ordinary transmission states is the routing of informational signals, whereas the primary role of special states is to inject transforming signals into cell locations and thereby convert "U" cells into active elements (or, if need be, convert active elements back into "U" cells).

The system also has four "confluent" states. They are activated if they receive signals from all cells in their neighborhood which are directed toward them. If activation occurs, then after two moments of time they emit signals outward toward any cell in their neighborhood which does not have a transmission directed toward it. Thus, confluent cells can serve as "and" gates, and as wire branching elements. Since they do not emit their output until two moments of time have elapsed, the confluent cells can also be employed to create time delays in the transmission of signals. The eight remaining cell states of the 29 originally employed by von Neumann are of less importance. These are temporary cell states which arise only as the operational states are being created from "U" cells.

Von Neumann first showed how to design a general purpose computing machine in his cell space system. He did this by showing the design of various basic organs — "pulsers" to emit any desired finite train of pulses upon activation, "periodic pulsers" to emit repeated trains of desired pulses after activation until signaled to stop, "decoders" to detect the presence of certain patterns of pulses, and the like. Using these organs, von Neumann developed a design for the control portion of a computing machine in one region of the cell space. He then showed how to organize an adjacent but indefinitely extendable portion of the cell space into a memory or information storage unit, which could be accessed by the control unit.

For the process of construction, von Neumann designed a construction unit, which, taking instructions from the memory unit, could send out a constructing arm (by creating an active pathway of transmission cells into a region of "U" cells) and at the end of the arm, convert "U" cells to the cell types specified in memory (see fig. 5.4). He showed that this constructor could create any pattern of passive cells whatsoever. Thus, he had designed with mathematical rigor a universal constructor, relative to all possible passive configurations of cells in the cell space.

Since the parent machine itself can be created in passive form, it can make a duplicate of itself by the following process. The parent machine is supplied initially with instructions to make a duplicate of its control, construction and memory units (the memory unit initially is empty). After it completes this major construction phase, the instructions call for the parent machine to make a copy of the instructions in its memory and to feed into the memory unit of the newly constructed machine. Then the parent machine activates the heretofore passive offspring machine, and withdraws the constructing arm. At that moment the offspring is a duplicate, in all respects, of the parent at the time the original machine commenced its reproductive activities.

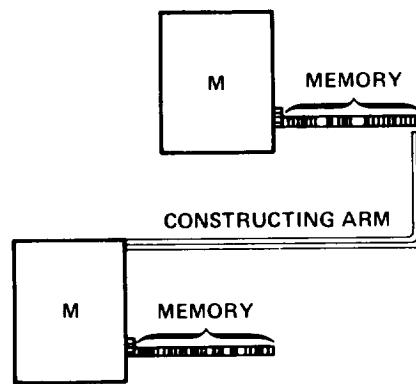


Figure 5.4.— Universal construction in the cellular model of machine self-reproduction.

Critique of the cellular model. Although the 29-state von Neumann cellular array system permits a more elegant mathematical approach to the problem of machine construction and self-reproduction, it is more difficult to envision an actual useful physical implementation of the process (compared, say, to the kinematic model of replication). The entire cell space enterprise proceeds in a highly constrained artificial environment, one which is very special despite some features relating in a general way to the natural world. For example, the movement of objects in space, a ubiquitous and familiar phenomenon in the real world, becomes a complex process of deletion of cell states at one location and re-creation of these states at some other location.

There is also an assumption of synchronous behavior throughout the system. All cells, no matter how distant, are subject to change of state at the same instant, a property which would be difficult to implement in any practical large cell space. Indeed, the requirement of a source of clocking pulses violates the array symmetry which makes the cell space notion an attractive object for mathematical treatment.

It is also very difficult to design machines of interest which can be embedded in the cell array format. To make design and embedding easier, a higher-level machine design language would have to be created. It is likely that, rather than undertake that task, one would first redesign the underlying cell space properties to rid the system of the deficiencies already noted.

For instance, one might wish to introduce a new primitive cell state in the system to permit signals to cross without interference. A "wire-crossing" organ can be devised using only the original von Neumann primitive cell types, but this introduces an unnecessary complexity into the machine design process since the organ contains initially active cell states whose creation involves considerable extra care to avoid the propagation of spurious signals. This extra care is especially critical because the cell system, as von Neumann originally constituted it, is highly susceptible to signal errors. (He undoubtedly intended his probabilistic machine model to mitigate this sensitivity and fragility.)

The cell space system has very limited capacity to detect the states of cells. It has some capacity to detect states, for this is required in the operation of the memory unit. But a machine cannot analyze an arbitrary encountered cell to determine what state it is in, thus cannot "read" the states of an encountered machine. This inability severely restricts the capacity of cell-space machines to repair other machines or to attempt self-repair. Such limitations also are evident in the construction process, where the constructing machine must assume that the region in which a new machine is to be created consists entirely of elementary quiescent cells. Should this not be the case, there is no systematic and complete way to detect it. A machine can send destruction signals into cells to reduce them to the quies-

cent form. Unfortunately, in some cases one must know the state of the cell ahead of time in order to determine what destructive signal must be sent to destroy it.

Finally, all machines that can be produced in von Neumann's cell space system are essentially information transactional devices. Even construction is, in this context, a form of information processing. Physical construction and material transformations can possibly be viewed as informational processes but, in a practical sense, the cell-space notion is far from providing a readily useful paradigm of actual manipulation and transformation of physical materials.

Von Neumann's other self-reproducing machine concepts. In addition to his kinematic and cellular models, von Neumann planned to examine three other models of self-reproducing machines. These were to be a neuronal or "excitation-threshold-fatigue" model, a continuous model, and a probabilistic model. Von Neumann is not known to have left any completed work whatsoever on these models at the time of his death, so his intentions are almost entirely a matter of conjecture.

Following Burks' speculations on this matter (von Neumann, 1966), we can guess that von Neumann's neuronal system might have been a version of the cell-space model in which the individual cell automata in the space were to be constructed of neuron-like elements. This would have been a rather straightforward process, as it is well known that idealized neurons of the McCulloch-Pitts (1943) variety can be employed to implement the kinds of logical gatings and delays called for in the 29-state cell automaton system. The reason for employing neuron-like elements seems mainly an attempt to recast the model in a more "biological" vocabulary.

Von Neumann's postulated continuous model might have been an attempt to comprehend machine reproduction in an even more biological format. The usual mathematical tools for handling actual neuron activity are differential equations expressing the electrochemical flows through and along neuron soma and axons. Thus the actions of cell automata (implemented with neurons) could be expressed by sets of differential equations. In this way the more highly developed tools of mathematical analysis might be employed in representing the behavior of the machine system, in contrast to the use of combinatorics which von Neumann himself characterized as one of the most intractable of mathematical specialties.

Finally, in his proposed probabilistic model von Neumann perhaps intended to consider using whole congeries of neuron-like elements in implementing the behaviors of what in the neuronal model could be carried out by single neurons. By employing redundancy techniques similar to those described in his classic paper on reliability, von Neumann (1956) may finally have hoped to

design a reliable, biologically oriented, self-reproducing machine characterizable by differential equations. We can only guess.

Alternative cell array systems. Work on cell-space automata systems in the period following von Neumann's contributions has taken several research directions. The underlying cell-space notion of a homogeneous medium with a local transition function that determines global properties has been employed in numerous modeling and simulation projects. For example, weather simulations use the idea of connected cells, the changes of each cell state described by a set of differential equations. Studies of the flow of excitation in heart tissue, the dispersal of medicinal drugs, and pattern recognition all have employed the cell-space concept. Cell spaces also have been investigated as abstract mathematical objects where, for instance, one tries to determine whether from every mathematical pattern all other patterns can be attained, and whether there are some patterns not attainable at all by means of the transition function, and various other specialized questions.

Some work in cellular automata has attempted to carry forth the von Neumann program of machine construction and self-reproduction. For instance, Codd (1968) recapitulated the von Neumann results in a simpler cell space requiring only 8 states rather than 29. This produced a machine design recognizably closer to that of present-day computing machines. Myhill (1970), trying to mitigate the artificiality of the indefinitely extended pre-existing cell space, designed a system in which componentry was drawn into a cell-grid system and was then employed as machine constituents somewhat as biological cell constituents might be drawn through a membrane to be used at an intracellular work site. Arbib (1966), attempting to make the movement of cell machines a less cumbersome matter, designed a cell-space system in which cells and blocks of cells might be joined together by a "welding" operation, thus becoming "co-moving" configurations.

Smith (1970) and Banks (1970) introduced additional simplifications to the cell-space notion, showing that the von Neumann program could be recapitulated in underlying cell spaces of an extremely elementary sort. Indeed, the so-called "Game of Life" designed by Conway (Gardner, 1971) is a cell-space system which, despite its very simple transition rules, has been claimed to be capable of expressing both universal computation and construction. (The game involves a checkerboard cell array with cells in one of two states, "0" or "1." A point whose state is "0" will change to state "1" if exactly three of its eight neighbors are in state "1." A point whose state is "1" will remain in that state if two or three of its neighbors are also in state "1." In all other cases, the state becomes or remains "0.")

Later research on self-reproducing automata. By the late 1960s, the original von Neumann program of machine construction and reproduction had been largely abandoned, although investigation of cell-space systems as abstract mathematical entities or as vehicles for "spatial" modeling and simulation has persisted. Indeed, research in the latter field has been especially vigorous and prolific — one recent author lists over 100 references for cell-space imaging applications (Preston et al., 1979).

Von Neumann's kinematic machine construction system appears to have had no intellectual progeny whatsoever. This is somewhat misleading, since practical application of computers to manufacturing and the persistent human interest in and investigation of robot mechanisms have, without explicit connection to von Neumann's earlier work, prepared the ground for a possible implementation of a hybrid computer/kinematic model of machine construction and reproduction.

The theoretical work of this later period, explicitly derived from von Neumann's research effort, has focused mainly on the molecular biological analogies that can be drawn. For example, in a series of papers Laing (1975, 1976, 1977, 1978, 1979) employs a hybrid cellular-kinematic model of machine construction and shows that neither existing natural nor artificial machines need be bound to follow the "classical" reproductive paradigm. In the classical paradigm, a program (DNA in living systems) is first interpreted to construct a machine (protein synthesis in lifeforms) and then is read a second time to make a copy of the program for insertion into the newly constructed duplicate machine (DNA replication in living cells). The principal contribution of Laing is to suggest reproductive strategies other than direct analogues to the known biological process. In this new conception, a machine is able to identify all of the components of which machine systems consist (not merely a subset as in the von Neumann cell system) and can access all of an existing machine structure without requiring dismantling of the system (as would be required in the von Neumann model).

Once this and other similar advanced concepts are brought to bear on the problems of machine reproduction, many alternative reproduction strategies become immediately apparent. A selected few of these are reviewed in the following section.

5.2.2 Alternative Replication Strategies

A number of alternative automata reproduction strategies have been suggested in the decades following the completion of von Neumann's work. Major strides have been made in the scientific understanding of the processes of biological reproduction at the molecular or biochemical level. Recent research has demonstrated the theoretical

possibility of inferring structure and achieving self-replication without first possessing a complete self-description. This suggests an enormous range of new machine capabilities which possibly may be technologically exploited in the future, according to specific rules and multiplication strategies for optimal deployment.

Biological reproduction. Biological reproduction is thought to obey the following underlying logical paradigm. The basic genetic program (encoded in the genetic DNA) is employed to make a copy of the same information in a slightly different medium (RNA). This modified form of the genetic program is transported to a work site within the cell where, with the aid of cellular enzymatic machinery, the RNA is interpreted as coding for amino acid strings (proteins). The protein produced plays two major roles: (1) it constitutes the basic structural material of living organisms, and (2) certain smaller and variably active proteins (enzymes) control the metabolic, interpretive, and constructive actions of the system.

When the genetic code embodied in the RNA has been read and acted upon, the machinery construction phase is complete. The cell must then undertake the copying of original genetic material (the DNA) to provide offspring organisms with the necessary instructions. This copying process is the well-known DNA replication phase, in which DNA (in most cases a twisted pair of complementary DNA molecules) untwists to permit new nucleotides to match with existing separated strands to form *two* twisted pairs of DNA. Reproduction is completed when the newly produced and original organism machineries are divided up, one DNA program remaining with each.

This highly simplified description of biological reproduction is offered only to illustrate the underlying logical strategies: (1) follow instructions to make machinery, (2) copy the instructions, (3) divide the machinery, providing a sufficient set in each half, (4) assign a set of instructions to each half, and (5) complete the physical separation.

Von Neumann's automata reproduction. Von Neumann's automata reproductive process closely mirrors the biological one. In the original model, instructions exist in two copies. One of the copies is read and acted upon to construct another machine, *sans* instructions. The second copy is then read and copied twice, and this double copy is inserted into the passive constructed offspring machine which is then turned on and released, thus completing the act of reproduction.

There is no logical necessity for having two sets of identical instructions. Von Neumann employed two copies of the instructions because it eliminated the criticism that the instructions might, in the first (construction) phase, become corrupted and so not be able to transmit a true version for the use of offspring machine. Also von Neumann feared that there might seem to be a paradox in the pro-

gram acting upon itself to make a copy of itself. There are, however, ways by which a program can successfully be made to make a copy of itself, and indeed many such programs, though exceedingly simple, have already been written (Burger, Brill, and Machi, 1980; Hay, 1980). Another solution is to provide the machine proper with an automatic "wired-in" copy routine which the program calls for at the proper time.

Simplified von Neumann automata reproduction. Consider a single instruction tape, and a constructor machine which reads the instructions once to build the offspring machine and again to make a copy of the instructions for the offspring machine. Notice that although the instructions available to the system yield a duplicate of the original system, this need not be the case. Thus, in the biological example, even though some DNA made available to a cell does not code the instructions for a duplicate cell, the cell-machine still may proceed to obey the instructions. This means that a cell can generate offspring not only different from itself and its normal constituents and products, but even inimical to it. This is precisely what happens when a virus possessing no metabolic machinery and no enzymatic protein machinery to read DNA or to manufacture anything parasitically insinuates itself into a host cell. The virus co-opts the host cell's interpreting and manufacturing capacity, causing it to make virus particles until the cell fills with them, bursts open, and is destroyed. The greatly multiplied viral agents are then free to parasitize other cells.

In artificial systems as well, machines may read and interpret instructions without knowing what they are being called upon to do. The instructions might call for some computational, constructional, or program-copying activities. The machine can make machines unlike itself, and can give these "unnatural" offspring copies of the instructions which were employed in their manufacture. If the offspring are also equipped to read and follow instructions, and if they have a constructional capability, their offspring in turn would be replicas of themselves — which might not resemble their "grandparent" machine at all. Thus, an original construction machine can follow instructions to make an indefinitely large number of diverse machines, that are like or unlike themselves, capable or not capable of constructing, reproducing, etc. And though a universal constructing machine might make large numbers of "sterile" machines, if it should once make a duplicate of itself which is also equipped with the instructional program for making duplicates of itself, the process can become "explosive." Such machines would tend to drive out all other "species" not possessing this reproductive "auto-catalytic" property.

Thatcher's variant: inferring structure. Thatcher (1970) showed that a machine need not have an explicit construction program made available to it initially in order to create

a duplicate of itself. First, it is sufficient that a machine can secure a *description* of itself (in place of instructions) if the machine is equipped with the capacity to read the description and convert this into the necessary constructive actions. Second, using a result obtained by Lee (1963) and himself (Thatcher, 1963), Thatcher showed that such a machine need not have its description loaded beforehand into its accessible memory organ. Instead, the machine has a partial self-description hard-wired into itself in the form of circuits which, when stimulated, make the description available to the machine in its accessible memory organ. These data describe all of the machine except the hardwired part which was stimulated to emit the description in the first place. The problem then, for the machine, is to obtain the description of this hidden part of itself. Lee and Thatcher showed that this section of the device can be constructed in such a simple fashion that the system can infer how this part must have been constructed merely by examining the consequences of its actions (e.g., the partial description it produced). After inferring the nature of this hidden part of itself, the machine possesses a complete self-description and can then follow von Neumann's paradigm for reproduction.

The principal practical significance of this form of automata replication is that it reminds the designer that the information required for machine construction (whether reproduction or not) need not be in the form of *instructions* for constructions but can be in the form of a *description*. Moreover, the description need not even reside in an accessible organ such as memory registers but may be embedded in "inaccessible" hardware. The hypothetical infinite regress likewise is shown to be baseless — it is possible for a machine to have within itself only a part of its own description, and from this to infer the rest.

Reproduction by component analysis. In von Neumann's cellular system, an embedded machine cannot send out an inspection arm to an encountered machine to identify all of its states. However, the cell-space system could be redesigned to permit this. In such a system an analyzing machine could examine an encountered passive machine and identify the type and location of all its cell-automata. (The analyzer might of course have to penetrate the machine, thus altering its automaton states, so the inspecting arm would have to send out appropriate restoration construction signals.)

In von Neumann's kinematic model a machine ostensibly could identify all parts of the system and thus determine the type and location of all components. This opens the possibility that a machine system might, for example, reproduce essentially two machines — one active, the other passive or able to assume passivity under a signal from the active machine. This possibility and others have been explored by Laing (1975, 1976, 1977, 1978, 1979) in a series of papers presenting alternative reproductive strategies which include the following:

- Beginning with two identical machines, one active and one passive, the active machine "reads" the passive machine twice, producing one active and one passive machine, thus completing reproduction.
- Beginning with two machines (not necessarily identical) one machine reads the second, and makes a duplicate of it. Then the second reads the first, and makes a duplicate of it, active and passive status being exchanged.
- By combining the capacity of machines to read machines with the Thatcher result, one can hardwire a machine to construct a second machine which is a duplicate of the original except for the hardwired part which produced the second machine. The original machine then "reads" the newly constructed partial duplicate, and infers what the missing hardwired part must be. The original machine then constructs the missing part, completing the reproductive process. This result explicitly confronts and overcomes the "necessary machine degeneracy" criticism of automata self-replication.

Machine reproduction without description. In the machine reproduction schemes explained thus far, some arbitrary part of the machine which cannot be inferred is always made explicitly available in memory initially, or is implicitly made available in memory or for inspection by means of an internal wired-in memory, also not directly accessible. Laing (1976) showed that even this wired-in description is not necessary. In effect, a machine can carry out a self-inspection which can yield a description which in turn can be made available to the machine in constructing a duplicate of itself.

The process begins with a wired-in construction routine which produces a semiautonomous analyzer machine. This analyzer moves over the original machine and identifies the type and location of its componentry. This is reported back to the original machine, which uses this information to make a duplicate of itself. Thus, though it may be that a part of a machine "may not comprehend the whole" in a single cognitive act, a part of a machine *can* examine in serial fashion the whole machine, and in time can make this information available to the machine for purposes of replication.

Exploitation of basic machine capabilities. The "simplified von Neumann" automata reproductive strategy — whereby a machine employs a stored program of instructions to make other machines (including duplicates of itself) and then also provides the program or parts of programs of instructions to newly constructed machines — should probably be the central strategy for any actual physical machine reproducing systems. The other strategies are, from most points of view, more complex than this and

thus perhaps are less preferable. The virtue of the alternative strategies is not as practical ways of implementing machine reproduction but rather in suggesting many basic capabilities, which, in a complex automated replicating LMF, may be usefully employed. The following are some of the behaviors of which, under *suitable* conditions and design, machines are actually and potentially capable:

(1) A machine can be "hard-wired" to carry out a computation.

(2) A machine can be programmed to carry out a computation.

(3) A machine can be a general-purpose computer, in that it can be given a set of instructions which will enable it to carry out the computation of any other computer. Alternatively, a general-purpose computing machine can be given the description of any other computing machine, and can carry out the computational actions of the machine described.

(4) A machine can be hard-wired to carry out a construction activity.

(5) A machine can be programmed to carry out a constructional activity.

(6) A sufficiently complex machine can be a general-purpose constructor, vis-a-vis a set of machines, in that it can be given a set of instructions which enables it to carry out the construction of any of the set of machines. Alternatively, a machine can be given the description of any machine of the set, and can, from this description, construct the machine described.

(7) A machine can construct a duplicate of itself, including the instructions or description used to guide the construction process.

(8) A machine, given a coded set of instructions for machine actions, or a coded description of a machine, can make a copy of the instructions or coded description.

(9) A machine, given a coded set of instructions for machine actions, can infer the structure of a machine which can carry out the actions described, and can construct such a machine.

(10) A machine, given a coded set of instructions for a machine, or a description of a machine, can carry out the actions of the machine whose instructions are given or whose description is supplied.

(11) A machine, given the instructions for or the description of an unknown machine, can examine the instructions or description and can (a) infer some of the properties of the machine, (b) simulate the actions of the machine, (c) construct the machine, and (d) observe the actions of the constructed machine.

(12) A machine can determine the component types of encountered machines.

(13) A machine can determine the structure (the component type and arrangement of components) of encountered machines.

(14) A machine can thus obtain a structural description of an encountered machine and simulate its actions, construct a duplicate, and then observe the duplicate in action.

(15) A machine can possess a copy of its own description, perhaps stored in a memory organ.

(16) A machine can obtain a copy of its own present structure. Note that the present structure of a machine may deviate from the original design, and also from its present stored description of itself (which may be out of date).

(17) A machine can compare the stored description of itself with the description obtained by inspection, and note the discrepancies.

(18) A machine can make a duplicate of itself on the basis of its stored "genetic" description or on the basis of its present (possibly altered) structure. This latter is an example of transmission of acquired characteristics.

(19) A machine can examine duplicates of itself constructed on the basis of an examination of itself, and note the discrepancies.

(20) The duplicates made from either of these two bases (genetic and observed) can be set in action and observed.

(21) For diagnostic purposes, the two kinds of descriptions can be compared, the two passive structures compared, the two kinds of structures in action observed and compared. The basis for machine self-diagnosis is thus available.

(22) A machine noting the discrepancies between two machine descriptions, or machine structures, or two machine behaviors, can in some cases act so as to resolve the discrepancies. That is, a machine in some cases can repair or reject or reconstruct deviant machines (including itself).

(23) A machine encountering an "unknown" machine can observe the behavior of that machine and compare this to the behavior of other machines, both directly and by simulating the behavior of those machines for which it already has or can obtain descriptions.

(24) A machine encountering an unknown machine can examine the structure of the machine and obtain a structural description which can be compared with other structural descriptions.

(25) Encountering an unknown device, a machine can use the structural description of the unknown to simulate its actions. These simulated actions can be compared to those of other machines whose descriptions are stored or which can be made available.

(26) Having the description of an encountered device, a machine can construct a duplicate of it. This duplicate can be set in action and observed, and its behavior compared with the behavior (actual or simulated) of other machines.

(27) The structure and behavior of encountered machines can be compared with those of known useful or benign machines, including that of the inspecting machine itself. This comparison, and the degrees of similarity

discerned, can serve as the basis for a subsequent policy of "friendship," "tolerance," "avoidance," "enmity," etc.

(28) The descriptions of encountered machines can be incorporated into the reproductive construction cycle so that these new machines or their features become part of the continuing and evolving machine system. This is an analogue to biological symbiosis.

Machine multiplication strategies. In describing the logical process of machine reproduction we have concentrated on the means by which the parent system could come to possess the information needed to carry out a replication and the associated question of how offspring would if necessary acquire the programs needed to continue the machine reproduction process. Although these questions, logically, are at the heart of machine replication, they leave open many issues concerning creation and siting of new machine systems as well as the ultimate fate of such systems.

This matter can be approached by considering certain biological analogues to the machine situation. In the known biological realm, all living organisms use the same underlying reproductive logic of protein synthesis and nucleotide sequence copying but employ vastly different broad strategies in producing more of their own kind.

One strategy is seen in the case of seed-bearing plants (as well as most fish and insects), in which vast numbers of "minimal" genetic packets are produced by the parent system and dispersed in the hope that a sufficient number will, largely by chance, find an appropriate site at which to survive and complete growth and development to maturity. At the other end of the scale is human behavior, whereby "construction" and nurture of the offspring may continue under the control and protection of the parent system until near maturity.

The particular multiplication strategy for artificial reproducing systems must of course be adjusted to intentions. The swift utilization of large rich environments might justify a "seed" dispersal strategy, with early maturity of new systems so as to retain a high reproductive rate. On the other hand, an environment consisting of scattered pockets of valuable resources, or situations with less pressure for immediate "explosive" utilization might suggest fewer offspring, possibly more fully developed in regard to their capacity for seeking out and efficiently utilizing the scarce resources available. In this case, the offspring might also be expected to receive longer tutelage from the parent system or from outside controllers (such as humans).

Similarly, the presence of a large contiguous valuable ore body might dictate the extensive ramification of a single machine factory system consisting of many laboring sub-machines. The model of a colonial organism such as coral, or of a social insect such as ants or termites, might make more sense. Zoological and sociobiological studies of animal and plant multiplication strategies may prove

valuable in suggesting optimal machine system growth and reproduction strategies. One important difference must be borne in mind: biological organisms often have adapted their strategies to compete with other organisms, as well as to survive in a world where resources are renewed at certain rates over varying seasons. Some of these factors may be nonexistent or present in very different form in a nonterrestrial machine-inhabited environment.

A few questions that should be considered in determining optimal replicating machine behavior include:

- How large should a system be allowed to grow?
- How large should a system grow before it reproduces?
- What sorts of offspring (e.g., minimal vs mature) should be produced? A mixture?
- How many offspring should be produced? How many offspring should be produced from a single parent machine?
- When should offspring be produced?
- Where and how should offspring be sited? Specific sites? Near? Far? Randomly dispersed?
- What offspring transport mechanisms should be employed? Should new systems be mobile? Under own control? Parent? Human operator?
- When should sited machine systems be turned off? Abandoned? Should lifespan of a machine system be a function of time alone? Reproductive life? Exhaustion of local resources? Work experience and use? Detection of malfunction? When should subsystems be turned off? What growth and death patterns of individual machine systems should be adopted?
- What should be done with unsited offspring systems? Allowed to wander indefinitely?
- What should be done with outmoded machine systems? Dismantle them? Abandon them?

Intergeneration information transmission among replicating machines. Throughout most of the present discussion it has been assumed that the goal was to have the parent machine transmit to its offspring machine the same genetic information it received from its parent, regardless of the logical strategy of reproduction employed. This genetic fidelity is not necessary or even desirable in all cases. Normally the parent should transmit all information necessary for offspring to do their jobs and to construct further offspring in turn, but beyond this simple requirement there are many alternatives. For example, a parent machine might augment its program during its lifetime with some valuable information, and this augmented part of the program could then be transmitted to its offspring.

A few possible variations of interest include:

(1) The parent machine program is not altered in the course of its lifetime and is transmitted unaltered to offspring.

(2) The parent machine program is altered (e.g., by intervention, or by some machine adaptive process of a more or less complex sort) during the course of its lifetime, but again only the program originally received from the parent is transmitted to the offspring.

(3) The parent machine program is altered during the course of its lifetime, and the altered program is transmitted to the offspring machine. The parent machine (being a constructing machine) may make changes in its structure beyond those called for in its received genetic program.

(4) Changes in parent structure are not made part of the offspring structure.

(5) Changes in parent structure are made part of the offspring structure.

(6) Changes in parental structure are not made part of the offspring structure, but are made part of the offspring genetic program. Thus, the offspring can, under its own control, modify its structure to conform to that of its parent machine.

5.2.3 Information and Complexity in Self-Replicating Systems

The design and implementation of a self-replicating lunar factory represents an extremely sophisticated undertaking of the highest order. It is useful to consider the complexity of this enterprise in comparison with the information requirements of other large systems, natural or artificial, replicating or not (Stakem, 1979).

It is not immediately clear what the proper measure should be. One way to look at the problem of machines reproducing themselves is to consider the flow of information that occurs during reproduction. A fully generalized self-replicating system could possess a reproductive behavior of such complexity that the information necessary to describe that behavior is complete to atomic level specifications of machine structure. Such a machine has behavior so complex and complete that it might produce a copy of itself almost from complete chaos — say, a plasma containing equal concentrations of all isotopes. In this case the machine reproduction is essentially complete — given sufficient energy, the system can make copies of itself in any arbitrary environment even if that environment contains virtually no information relevant to replication.

At the other extreme, consider a long row of Unimate PUMA-like industrial robots side by side, each requiring merely the insertion of a single fuse to render it functional. The first working robot, its fuse already in place, seeks to “reproduce” itself from a “substrate” of dormant machines. It accomplishes this by reaching onto a nearby conveyor belt, picking up a passing fuse part, and plugging

it into the neighboring robot. The adjacent machine now begins to function normally as the first (indeed, as an exact duplicate), so it can be said that in some sense the first machine has reproduced itself. Before the reproductive act there was no second working robot; afterwards, one exists. However, this is almost the most trivial case of replication imaginable, since the substrate for reproductive activity — in this case completed machines lacking only fuses — is extremely highly organized. Hence, the operative complexity resides in the substrate, and the action of the machine in “making a new machine” is trivial.

This latter example may be compared to the case of a bacteriophage. The phage particle infects a healthy bacterium, using the captive cellular machinery to manufacture new virus particles. Only the DNA of the virus enters the bacterium, instructing the cellular machinery to make new viral DNA and to interpret the DNA to create protein and polysaccharide components which form the coat or carrier of the viral DNA. Thus the foreign DNA, like the PUMA robot which inserts fuses to “self-replicate,” must situate itself in a very rich complex environment, one already containing a great deal of machinery and information. In this case, the complexity of the virus-making enterprise probably can be gauged by the length of the viral DNA inserted into the host cell, just as the true complexity of the fuse-insertion behavior can be gauged by the length of the program needed to permit location of the supply of fuses and the fuse holder on an adjacent machine in physical space, and to insert the part properly. It is suggested, therefore, that the length of the shortest program which can carry out the process of replication may be an appropriate measure of the complexity of the task.

For instance, in the case of the von Neumann cellular reproducing system each part is already located in its proper place in space, but signals must be injected into that space to cause it to take on the properties desired in the offspring machine. It has been estimated that such a reproducing machine might consist of a minimum of 10^5 cells, with offspring cell type and location the principal parameters which must be specified for each. The length of the shortest program would represent perhaps 10^6 bits of information (Kemeny, 1955).

If the construction of a replicating growing lunar factory was purely a matter of machine parts assembly, then the length of the replication program could be determined by the necessity to locate various required parts in the environment and then to specify and execute the proper placement of each part to construct the desired system (Heiserman, 1976). However, it is likely the reproductive process will be vastly more complicated than this, since it is not likely that all parts can be supplied “free” from Earth. If the lunar factory must begin, not with completed machines or parts, but rather with a raw lunar soil substrate, the task quickly becomes many orders more difficult — though not impossible. Based on the estimates outlined in section 5.3 and the

appendices, the lunar factory replication program length should not exceed roughly 10^{12} bits of information. This compares to about 10^{10} bits coded in the human genome and about 10^{14} bits stored in the human brain. Terabit (10^{12} bits) memories are considered state-of-the-art today.

Complexity of a self-replication program may also be viewed as an index of versatility or system survivability. The more complex the program, the more likely it is that the machine system can bring about its own replication from increasingly disordered substrates. This is an interesting observation because it suggests that reproduction is an activity defined along a broad continuum of complexity rather than as a single well-defined event. Both the chaos-replicator and the fuse-insertion robots described above perform acts of self-reproduction. Fundamentally, these systems differ only in the degree to which they are capable of bringing order to the substrate in which they are embedded.

It is interesting to note that human beings fall somewhere in the middle of this broad reproductive spectrum. A 100 kg body mass, if composed of purely random assortments of the 92 natural elements, would contain roughly 10^{27} atoms and hence require about 10^{28} bits to describe. Yet a 100 kg human body is described by a chromosome set containing just 10^{10} bits. The difference must be made up by the "substrate" in which people are embedded – a highly ordered rich environment, namely, the Earth. Human beings thus are conceptually remarkably similar to von Neumann's kinematic self-reproducing automata, moving around in a "stockroom" searching for "parts."

5.2.4 Conclusions

The Replicating Systems Concepts Team reached the following conclusions concerning the theory of machine reproduction:

(1) John von Neumann and a large number of other researchers in theoretical computer science following him have shown that there are numerous alternative strategies by which a machine system can duplicate itself.

(2) There is a large repertoire of theoretical computer science results showing how machine systems may simulate machine systems (including themselves), construct machine systems (including machine systems similar to or identical with themselves), inspect machine systems (including themselves), and repair machine systems (including, to some extent, themselves). This repertoire of possible capabilities may be useful in the design and construction of replicating machines or factories in space.

5.3 Feasibility

The design and construction of a fully self-replicating factory system will be a tremendously complicated and difficult task. It may also be fairly expensive in the near-term. Before embarking upon such an ambitious undertaking it must first be shown that machine self-replication and growth is a fundamentally feasible goal.

5.3.1 Concept Credibility

The plausibility of the theoretical notion of self-replicating machines already has been reviewed at length (see sec. 5.2). It remains only to demonstrate concept credibility in an engineering sense (Bradley, 1980, unpublished memorandum, and see appendix 5A; Cliff, 1981; Freitas, 1980a; von Tiesenhausen and Darbro, 1980) – that is, is it credible to consider building real physical machines able to replicate themselves?

The credibility of any design proposed for such a machine or machine system depends first and foremost upon whether that design is consistent with reasonably foreseeable automation and materials processing technologies. These technologies need not necessarily be well established or even state-of-the-art, but should at least be conceivable in the context of a dedicated R&D effort spanning the next two decades. It is interesting to note that computer programs capable of self-replication have been written in many different programming languages (Burger et al., 1980; Hay, 1980), and that simple physical machines able to replicate themselves in highly specialized environments have already been designed and constructed (Jacobson, 1958; Morowitz, 1959; Penrose, 1959).

Another major requirement for concept credibility is a plausible system configuration. Proposed designs for self-replicating systems (SRS) must be sufficiently detailed to permit the generation of work breakdown structures, subsystem operational flowcharts, mass and energy throughput calculations, and at least preliminary closure (see sec. 5.3.6) analyses.

A related requirement is plausible mission scenarios. Research and development costs for the proposed design should be many orders of magnitude less than the Gross National Product. The mission must not require launch and support facilities which cannot or will not be available in the next two or three decades. The mission must entail reasonable flight times, system lifetimes, growth rates, production rates, and so forth. The problems of reliability and repair should be addressed.

The final requirement for concept credibility is positive societal impact. A given SRS design must be economically, politically, and socially feasible, or else it may never be translated into reality even if the technology to do so exists. A general discussion of the implications of replicating systems appears in section 5.5, but the team has arrived at no firm conclusions regarding concept feasibility in this area. More research is clearly required.

5.3.2 Concept Definition

In order to demonstrate SRS concept credibility, specific system designs and mission scenarios must be subjected to a detailed feasibility analysis. The first step in this process is to conceptualize the notion of replicating systems in as broad an engineering context as possible. Many kinds

of replicating machine systems have been proposed and considered during the course of the study. Some of these place emphasis on different types of behavior than others.

Consider a "unit machine" which is the automata equivalent of the atom in chemistry or the cell in biology -- the smallest working system able to execute a desired function and which cannot be further subdivided without causing loss of that function. The unit machine may be comprised of a number of subunits, say, A, B, C, and D. These sub-units may be visualized in terms of structural descriptions (girders, gearboxes, generators), functional descriptions (materials processing, parts fabrication, mining, parts assembly), or any other complete subset-level descriptions of the entire system.

SRS may be capable of at least five broad classes of machine behavior:

Production -- Generation of useful output from useful input. The unit machine remains unchanged in the process. This is a "primitive" behavior exhibited by all working machines including replicating systems.

Replication -- Complete manufacture of a physical copy of the original unit machine, by the unit machine.

Growth -- Increase in mass of the original unit machine by its own actions, retaining the physical integrity of the original design.

Evolution -- Increase in complexity of structure or function of the unit machine, by adding to, subtracting from, or changing the character of existing system subunits.

Repair -- Any operation performed by a unit machine upon itself, which does not alter unit population, designed unit mass, or unit complexity. Includes reconstruction, reconfiguration, or replacement of existing subunits.

These five basic classes of SRS behavior are illustrated in figure 5.5.

Replicating systems, in principle, may be designed which can exhibit any or all of these machine behaviors. In actual practice, however, it is likely that a given SRS format will emphasize one or more kinds of behaviors even if capable of displaying all of them. The team has considered two specific replicating systems designs in some detail. The first (cf. von Tiesenhausen and Darbro, 1980), which may be characterized as a unit replication system, is described in section 5.3.3. The second (cf. Freitas, 1980a; Freitas and Zachary, 1981), which can be characterized as a unit growth system, is outlined in section 5.3.4. The team decided to concentrate on the possibility of fully autonomous or "unmanned" SRS, both because these are more challenging from a technical standpoint than either manned

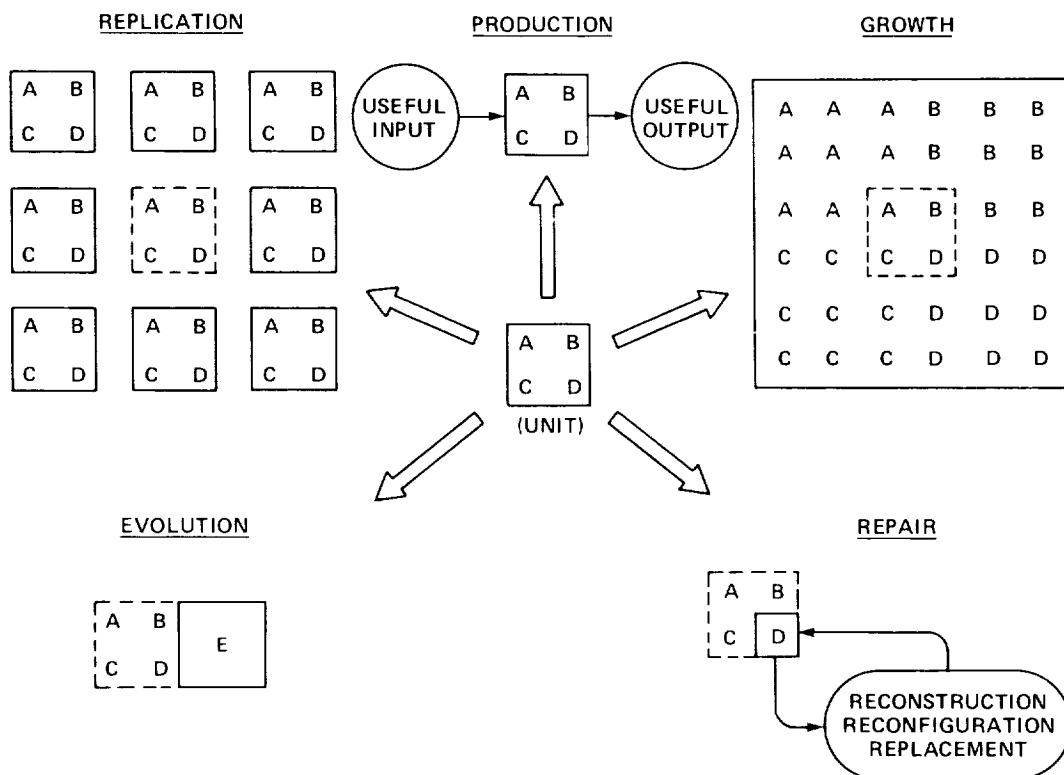


Figure 5.5. -- Five basic classes of SRS behavior.

or teleoperated systems and also because the latter has already been treated to some degree elsewhere in this report (see chap. 4).

5.3.3 Unit Replication: A Self-Replicating System Design

The SRS design for unit replication is intended to be a fully autonomous, general-purpose self-replicating factory to be deployed on the surface of planetary bodies or moons. The anatomy of an SRS is defined by two end conditions: (1) the type and quantity of products required within a certain time, and (2) the available material needed to manufacture these products as well as the SRS itself.

There are four major subsystems which comprise each SRS unit, as shown in figure 5.6. First, a materials processing subsystem acquires raw materials from the environment and prepares industrial feedstock from these substances. Second, a parts production subsystem uses the feedstock to make machines or other parts. At this point SRS output may take two forms. Parts may flow to the universal constructor subsystem, where they are used to construct a new SRS (replication). Or, parts may flow to a production facility subsystem to be made into commercially useful products. The SRS also has a number of other important but subsidiary subsystems, including a materials depot, parts depots, product depot, control and command, and an energy system.

The work breakdown structure given in figure 5.7 lists all SRS elements studied, and each is briefly described below.

Materials processing and feedstock production. In this subsystem, raw materials are gathered by strip or deep mining. They are then analyzed, separated, and processed into industrial feedstock components such as sheets, bars, ingots, castings, and so forth, which are laid out and stored in the materials depot. The processing subsystem has a high degree of autonomy including self-maintenance and repair. It is linked to a central supervisory control system (see below).

The materials processing subsystem is shown schematically in figure 5.8.

Materials depot. The materials depot collects and deposits in proper storage locations the various feedstock categories according to a predetermined plan. This plan ensures that the subsequent fabrication of parts proceeds in the most efficient and expeditious manner possible. The depot also serves as a buffer during interruptions in normal operations caused by failures in either the materials processing subsystem (depot input) or in the parts production subsystem (at depot output).

Parts production plant. The parts production plant selects and transports industrial feedstock from the materials depot into the plant, then fabricates all parts required for SRS production or replication activities. Finished parts are stored in the production parts and the replication parts depots, respectively. The parts production plant is highly automated in materials transport and in distribution, production, control, and subassembly operations.

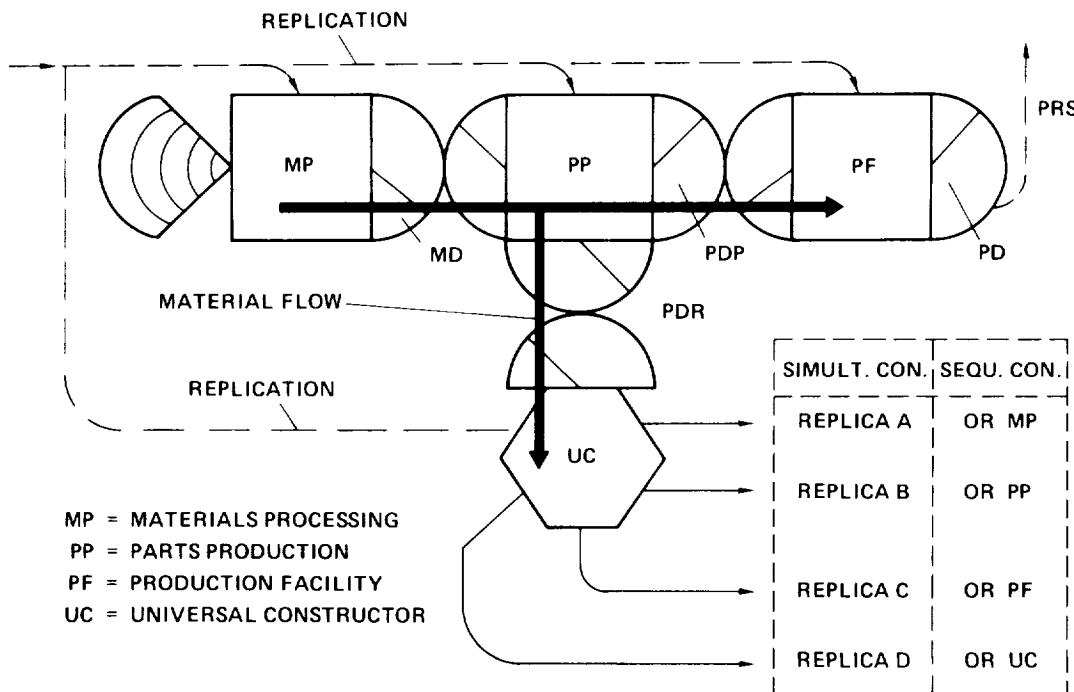


Figure 5.6. -- Functional schematic of unit replication SRS.

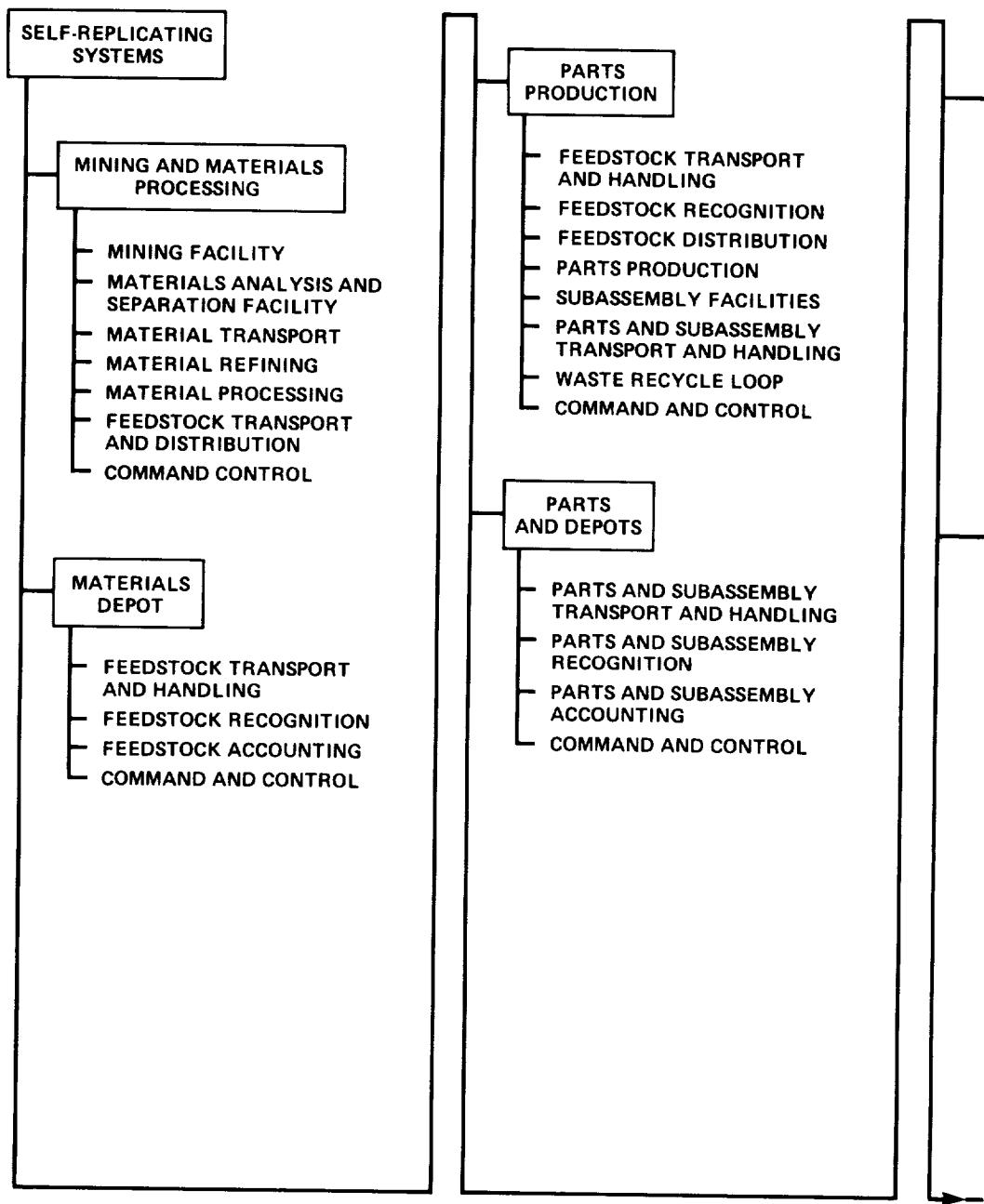


Figure 5.7. – Work breakdown structure for SRS.

The parts production plant subsystem is shown schematically in figure 5.9.

Parts depots. There are two parts depots in the present design. These are called the production parts depot and the replication parts depot.

Parts are stored in the production parts depot exclusively for use in the manufacture of useful products in the production facility. If certain raw materials other than parts and subassemblies are required for production, these

materials are simply passed from the materials depot through the parts production plant unchanged. The parts production depot also acts as a buffer during interruptions in normal operations caused by temporary failures in either the parts production plant or the production facility.

Parts and subassemblies are stored in the replication parts depot exclusively for use in the replication of complete SRS units. Storage is in lots earmarked for specific facility construction sites. The replication parts depot also serves as buffer during interruptions in parts production plant or universal constructor operations.

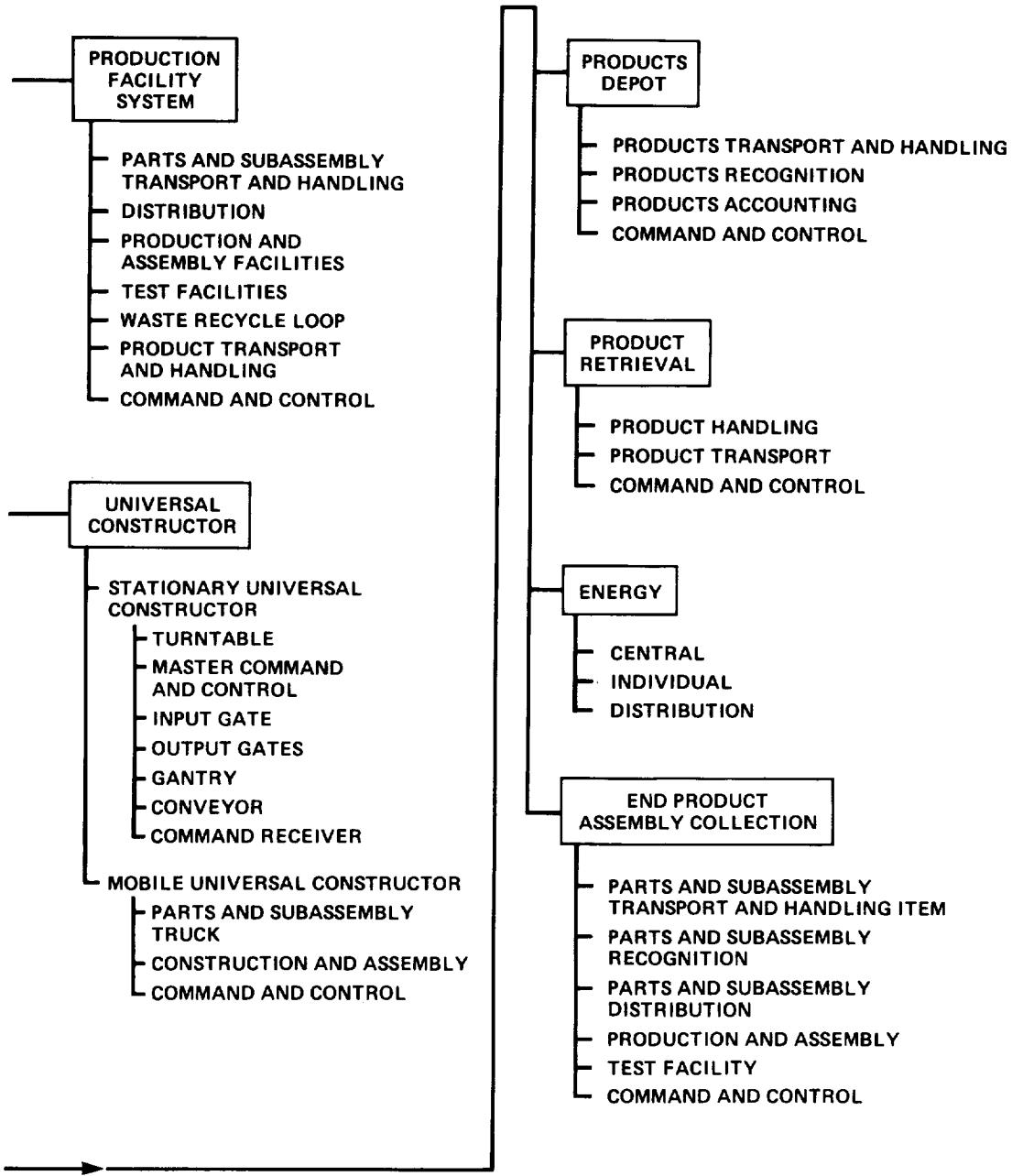


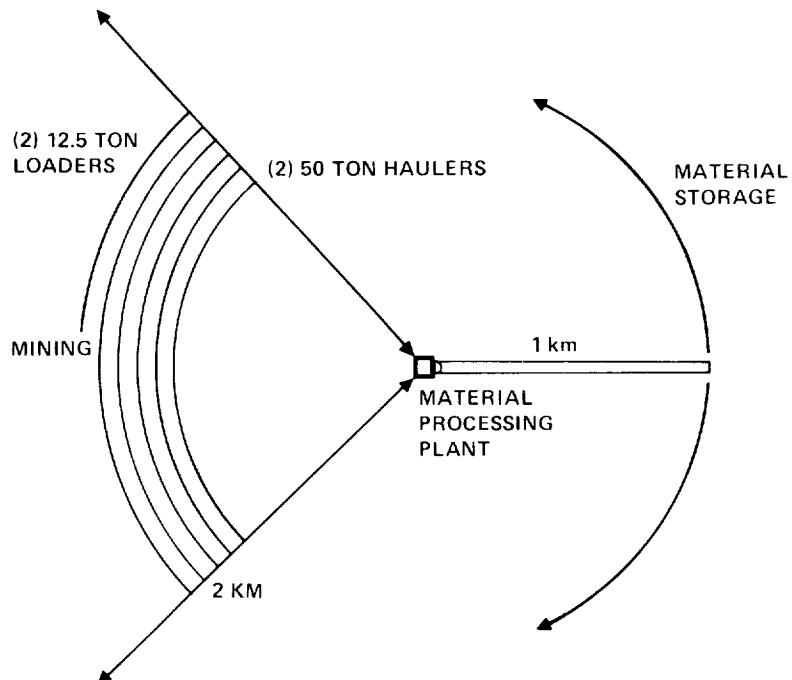
Figure 5.7.— Concluded.

Production facility. The production facility manufactures the desired products. Parts and subassemblies are picked up at the production parts depot and are transported to the production facility to be assembled into specific useful products. Finished products are then stored in the products depot. Ultimately these are collected by the product retrieval system for outshipment.

Universal constructor. The universal constructor manufactures complete SRS units which are exact duplicates of

the original system. Each replica can then, in turn, construct more replicas of itself, and so on. The universal constructor retains overall control and command responsibility for its own SRS as well as its replicas, until the control and command functions have also been replicated and transferred to the replicas. These functions can be overridden at any time by external means.

The universal constructor subsystem consists of two major, separate elements — the stationary universal constructor (fig. 5.10) and the mobile universal constructors



REFERENCE: "LUNAR STRIP MINING SYSTEM"
W. DAVID CARRIER III, JUNE 1978

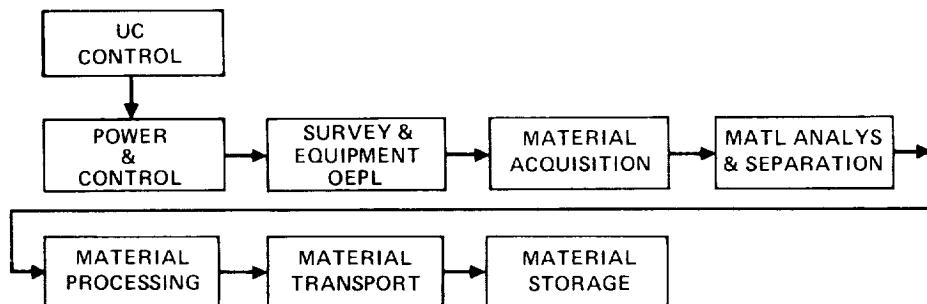


Figure 5.8. – SRS materials processing subsystem.

(fig. 5.11). This composite subsystem must successfully perform a number of fundamental tasks, including receiving, sorting, loading, and transporting parts and subassemblies; assembling, constructing, installing, integrating, and testing SRS systems; starting and controlling SRS operations; and copying and transferring instructions between system components.

Products depot. The outputs of the production facility are stored in the products depot, ready for retrieval. Major hardware components are neatly stacked for ready access by the product retrieval system. Consumables such as elemental oxygen are stored in reusable containers that are returned empty to the production facility. The products depot also serves as a buffer against variable output and retrieval rates.

Product retrieval system. The product retrieval system collects the outputs of all SRS units in an "SRS field" and carries them to an outside distribution point for immediate use or for subsequent outshipment. The dashed lines in figure 5.11 indicate one possible solution to this problem in a typical SRS field. Other solutions are possible -- careful consideration must be given to SRS field configuration to arrive at an optimum product retrieval system design.

Command and control systems. The master control and command system, located within the stationary universal constructor, is programmed to supervise the total SRS operation and to communicate both with the peripheral controls of the mobile universal constructors during the self-replication phase and with the replicated stationary universal constructor during the transfer of command and control for the operation of the new SRS unit.

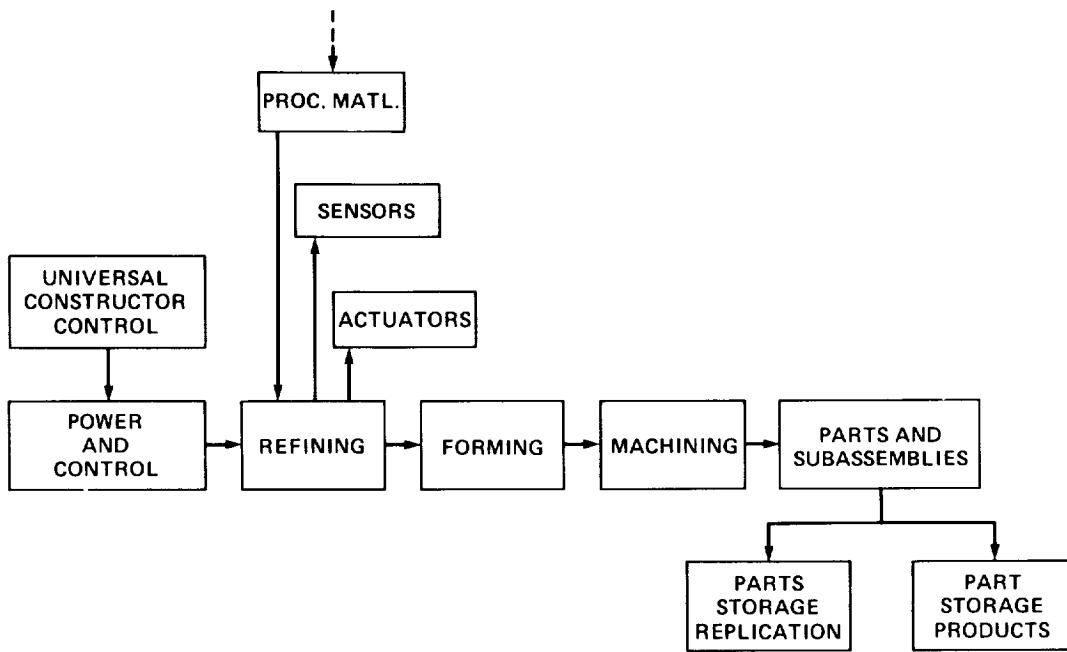


Figure 5.9. – SRS parts production plant subsystem.

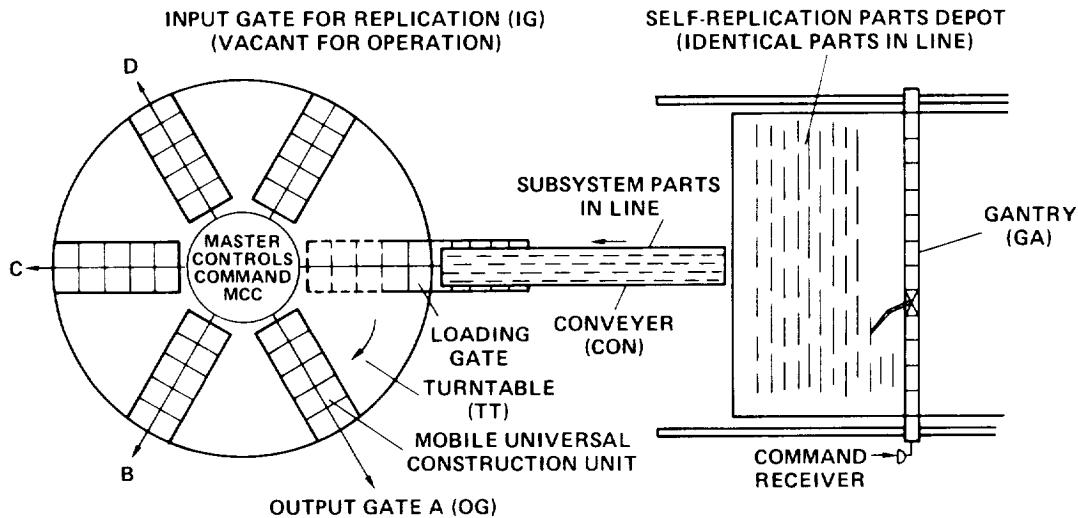


Figure 5.10. – SRS stationary universal constructor.

The master control and command system operates its own SRS unit through individual communication links which address the local control and command systems of individual SRS elements. In this way the master control and command system supervises the condition and operations of its own system elements, from materials acquisition through end product retrieval.

Energy system. The power requirements for the present design may be in gigawatt range. Hence, a single energy source (such as a nuclear power plant) would be excessively massive, and would be difficult to replicate in any case. This leaves solar energy as the lone viable alternative. Daylight options include: (1) central photovoltaic with a ground cable network, (2) distributed photovoltaic with

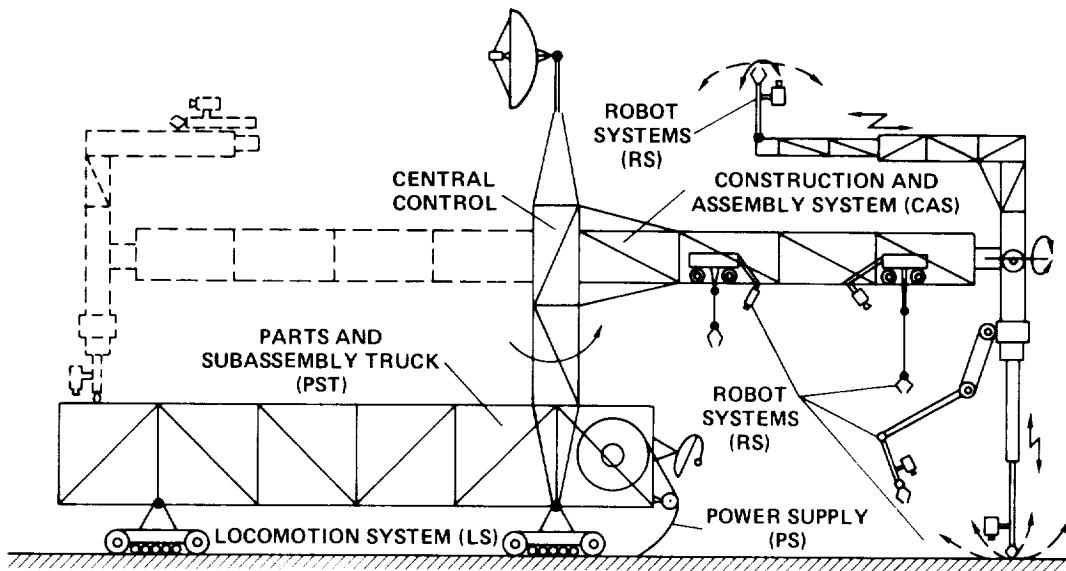


Figure 5.11.—SRS mobile universal constructors.

local distribution system, (3) individual photovoltaic, and (4) satellite power system, with microwave or laser power transmission to central, local, or individual receivers. Night-time power options include MHD, thermionics, or turbo-generators using fuel generated with excess capacity during daytime. Oxygen plus aluminum, magnesium, or calcium could be used for fuel. A 15% efficient central silicon photovoltaic power station has been assumed in the reference design, with an output of tens of gigawatts and a size on the order of tens of square kilometers.

Each SRS produces, in addition to its scheduled line of regular products, a part of the photovoltaic energy system equal to the energy needs of its replicas. These are retrieved along with the regular products by the product retrieval system and are assembled on-site to increase energy system capacity according to demand during the self-replication phase.

SRS deployment and expansion. A complete SRS factory unit, erected on the surface of the Moon, might appear as illustrated in figure 5.12.

As a unit replication scheme, the multiplication of SRS units proceeds from a single primary system to many hundreds of replica systems. This expansion must be carefully planned to reach the desired factory output capacity without running out of space and materials. Figure 5.13 shows one possible detailed growth plan for the geometry of an SRS field. In this plan, each SRS constructs just three replicas, simultaneously, then abandons replication and goes into full production of useful output. After the three generations depicted, an SRS field factory network 40 units strong is busy manufacturing products for outshipment.

The routes taken by mobile universal constructors are shown as solid lines, the product retrieval routes as dashed lines.

Figure 5.14 shows another possible expansion geometry. Again, each SRS constructs just three replicas, but sequentially rather than simultaneously. The end result is a field of 326 individual units after nine cycles of replication. Output is collected by the product retrieval system and taken to an end product assembly/collection system where end products undergo final assembly and other operations preparatory to outshipment. A more detailed discussion of expansion scenarios for SRS fields may be found in von Tiesenhausen and Darbro (1980).

Proposed development and demonstration scenario. It is proposed that the practical difficulties of machine replication should be confronted directly and promptly by a dedicated development and demonstration program having four distinct phases.

In Phase A of the development scenario, a robot manipulator will be programmed to construct a duplicate of itself from supplied parts and subassemblies. The original robot then makes a copy of its own operating program and inserts this into the replica, then turns it on, thus completing the duplication process (see appendix 5J). To complete Phase A, the replica must construct a replica of *itself*, repeating in every way the actions of the original robot. The rationale for the second construction, called the Fertility Test, is to demonstrate that the capacity for self-replication has in fact been transmitted from parent machine to offspring.

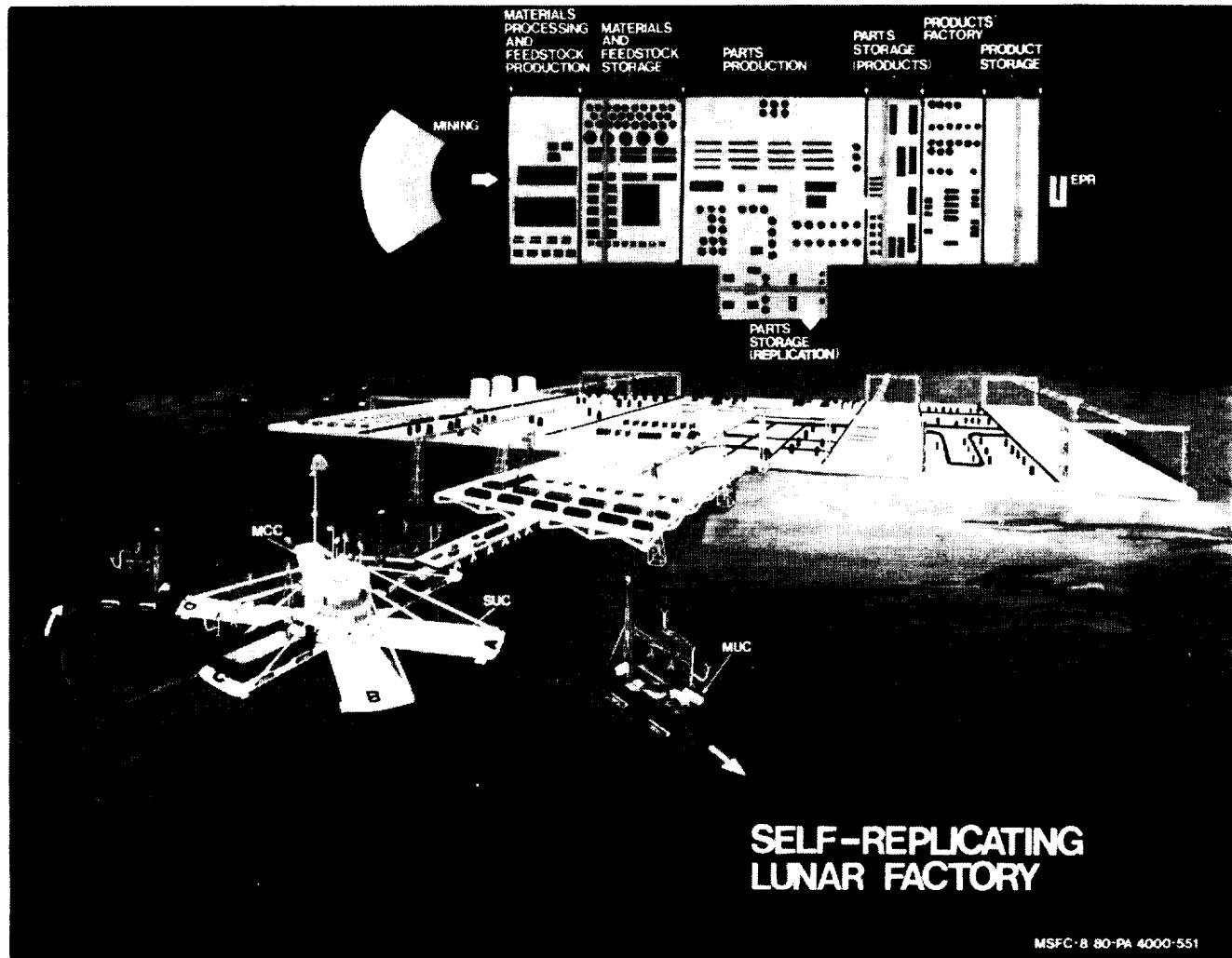


Figure 5.12.— Self-replicating lunar factory.

In Phase B of the development and demonstration scenario, the robot manipulator will be supplied with numerous additional parts so it can assemble objects of interest other than replicas of itself. This is intended to show that the system is able to construct useful products in addition to the line of robot duplicates.

In Phase C the manipulator system is still required to construct replicas and useful products. However, the robot now will be supplied only with industrial feedstock such as metal ingots, bars, and sheets, and must fabricate all necessary parts and subassemblies on its own. Successful completion of Phase C is expected to be much more difficult than the two earlier phases. The reason is that the parts fabrication machines must themselves be constructed by the robot manipulator and, in addition, all parts and subassemblies comprising the newly introduced fabrication machines must also be made available to the manipulator. Fabricator machines thus must be programmed to make not only the parts required for robot manipulators and useful products,

but also their *own* parts and subassemblies as well. This raises the issue of parts closure, a matter which is discussed in section 5.3.6.

In Phase D, the system developed in the previous phase is retained with the exception that only minerals, ores, and soils of the kind naturally occurring on terrestrial or lunar surfaces are provided. In addition to all Phase C capabilities, the Phase D system must be able to prepare industrial feedstock for input to the fabrication machines. Successful completion of Phase D is expected to be the most difficult of all because, in addition to the parts closure problem represented by the addition of materials processing machines, all chemical elements, process chemicals, and alloys necessary for system construction and operation must be extracted and prepared by the materials processing machines. This raises the issue of materials closure (see also sec. 5.3.6). The completion of Phase D will yield an automatic manufacturing facility which, beginning with "natural" substrate, can replicate itself.

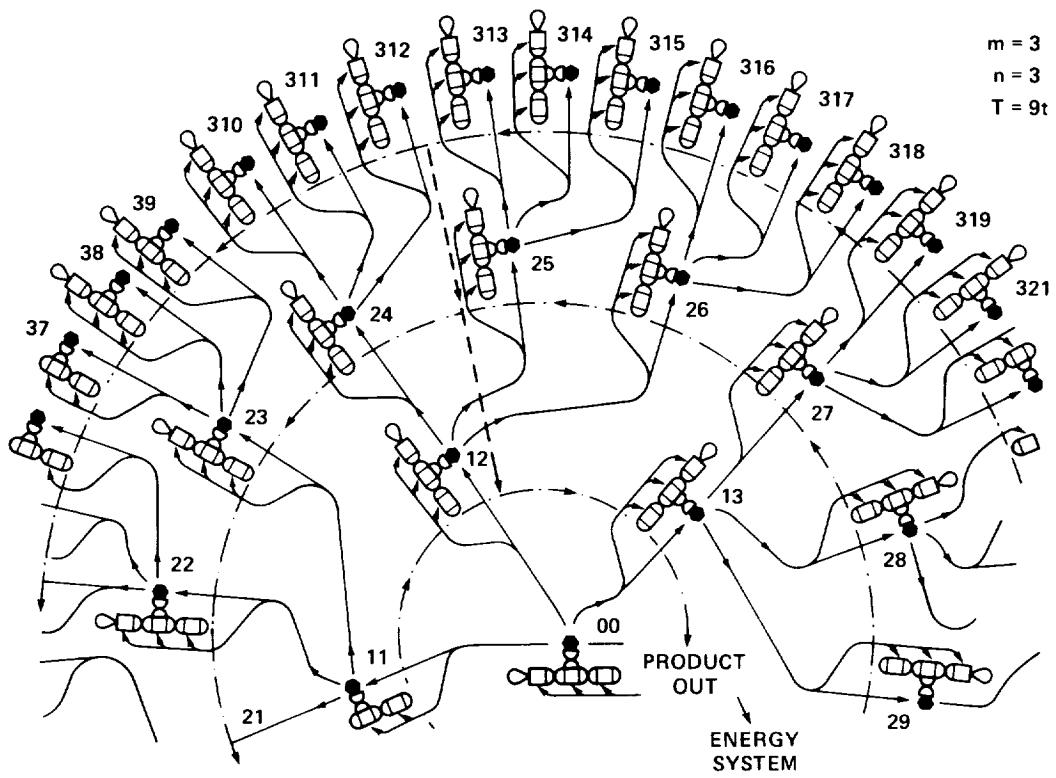


Figure 5.13.— Possible growth plan with simultaneous replica construction, suitable for geometry of an SRS field.

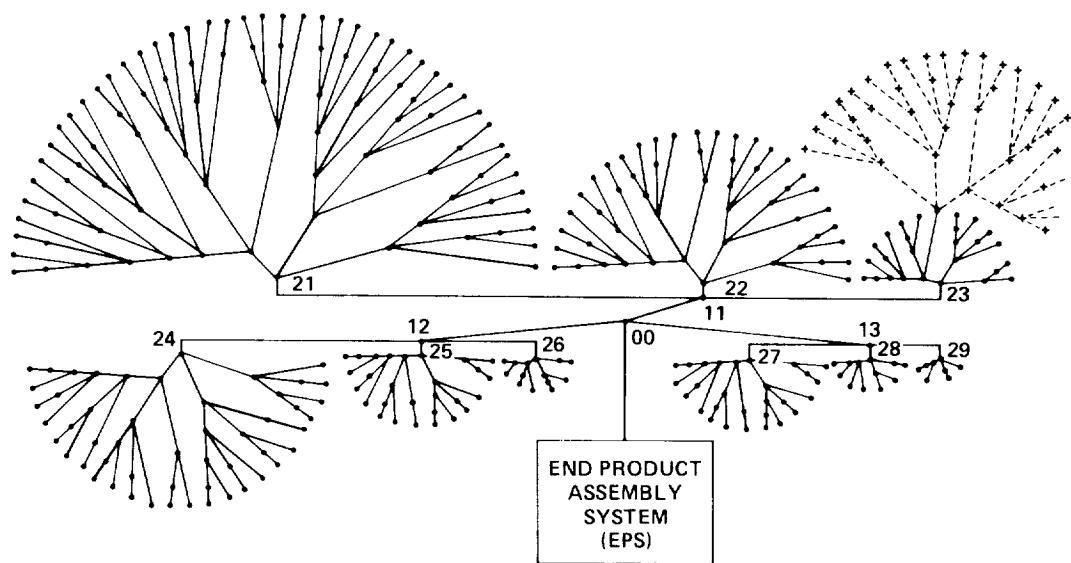


Figure 5.14.— SRS growth plan with sequential replication.

This progressive development of a replicating factory will serve to verify concept feasibility, clarify the functional requirements of such a system, and identify specific technological problem areas where additional research in automation and robotics is needed. A minimum demonstration program should be designed to gain engineering understanding, confidence, and hands-on experience in the design and operation of replicating systems. (See sec. 5.6.) The question of when the results of an Earth-based development and demonstration project should be translated to lunar requirements, designs, and construction remains open. On the one hand, it may be deemed most practical to complete Phase D before attempting a translation to a design better suited to a lunar or orbital environment. On the other hand, major system components for a lunar facility undoubtedly could be undertaken profitably earlier in concert with Phase C and D development.

The proposed development and demonstration scenario is described in greater detail in von Tiesenhausen and Darbro (1980).

5.3.4 Unit Growth: A Growing Lunar Manufacturing Facility

The Lunar Manufacturing Facility (LMF) demonstrating SRS unit growth is intended as a fully automatic general-purpose factory which expands to some predetermined

adult size starting from a relatively tiny "seed" initially deposited on the lunar surface. This seed, once deployed on the Moon, is circular in shape, thus providing the smallest possible perimeter/surface area ratio and minimizing interior transport distances. Expansion is radially outward with an accelerating radius during the growth phase. Original seed mass is 100 tons.

The replicating LMF design encompasses eight fundamental subsystems. Three subsystems are external to the main factory (transponder network, paving, and mining robots). The LMF platform is divided into two identical halves, each comprised of three major production subsystems: (1) the chemical processing sector accepts raw lunar materials, extracts needed elements, and prepares process chemicals and refractories for factory use; (2) the fabrication sector converts these substances into manufactured parts, tools, and electronics components; and (3) the assembly sector, which assembles fabricated parts into complex working machines or useful products of any conceivable design. (Each sector must grow at the same relative rate for uniform and efficient perimeter expansion.) Computer facilities and the energy plant are the two remaining major subsystems. (See fig. 5.15.)

Transponder network. A transponder network operating in the gigahertz range assists mobile LMF robots in accurately fixing their position relative to the main factory

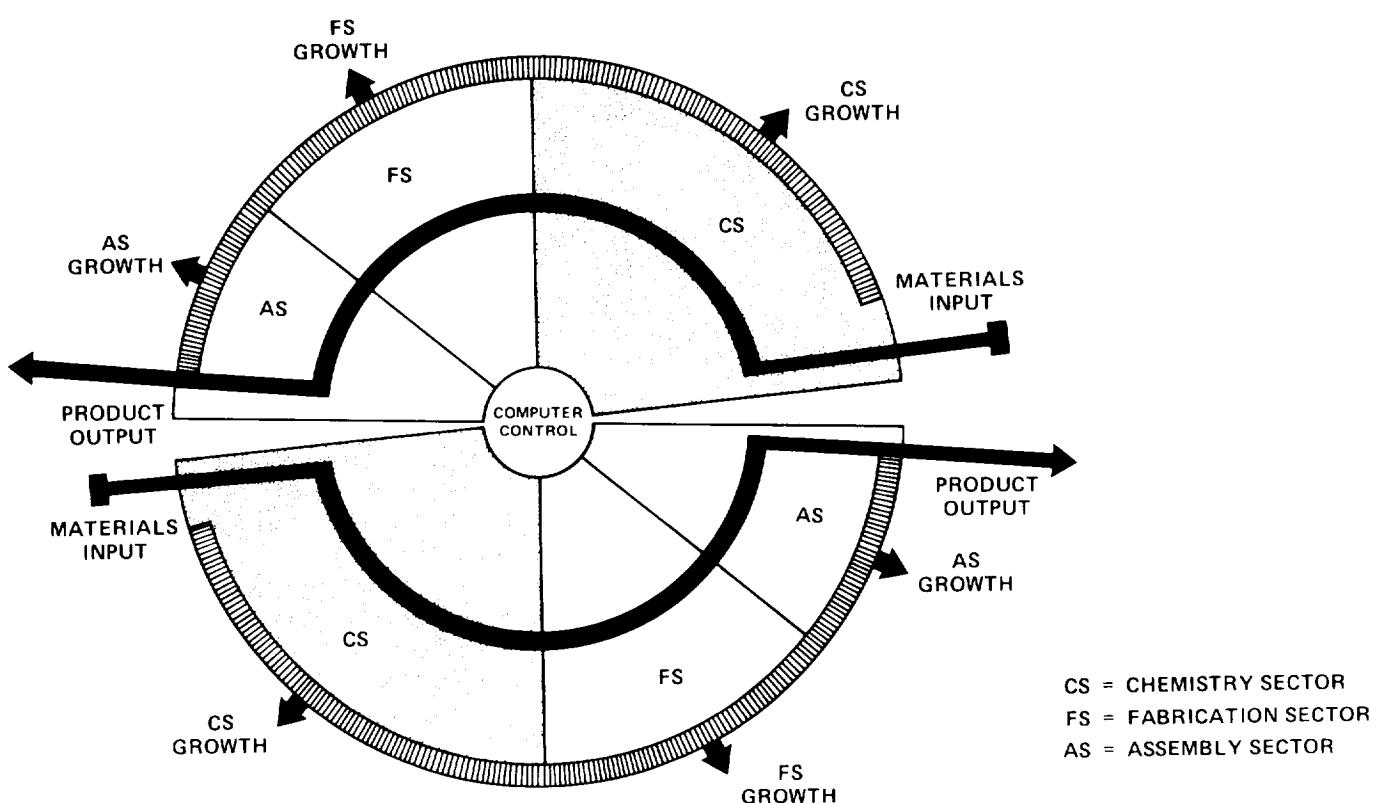


Figure 5.15.—Functional schematic of unit growth SRS.

complex while they are away from it. The network, described briefly in appendix 5B, is comprised of a number of navigation and communication relay stations set up in a well defined regular grid pattern around the initial seed and the growing LMF complex.

Paving robots. In order to secure a firm foundation upon which to erect seed (and later LMF) machinery, a platform of adjoining flat cast basalt slabs is required in the baseline design. A team of five paving robots lays down this foundation in a regular checkerboard pattern, using focused solar energy to melt pregraded lunar soil *in situ*. (See app. 5C.)

Mining robots. As described in appendix 5D, LMF mining robots perform six distinct functions in normal operation: (1) strip mining, (2) hauling, (3) landfilling, (4) grading, (5) cellar-digging, and (6) towing. Lunar soil is strip-mined in a circular pit surrounding the growing LMF. This material is hauled back to the factory for processing, after which the unused slag is returned to the inside edge of the annular pit and used for landfill which may later be paved over to permit additional LMF radial expansion. Paving operations require a well graded surface, and cellar digging is necessary so that the LMF computer may be partially buried a short distance beneath the surface to afford better protection from potentially disabling radiation and particle impacts. Towing is needed for general surface transport and rescue operations to be performed by the mining robots. The robot design selected is a modified front loader with combination roll-back bucket/dozer blade and a capacity for aft attachments including a grading blade, towing platform, and a tow bar.

Chemical processing sectors. Mining robots deliver raw lunar soil strip-mined at the pit into large input hoppers arranged along the edge of entry corridors leading into the chemical processing sectors in either half of the LMF. This material is electrophoretically separated (Dunning and Snyder, 1981; see sec. 4.2.2) into pure minerals or workable mixtures of minerals, then processed using the HF acid-leach method (Arnold et al., 1981; Waldron et al., 1979) and other specialized techniques to recover volatiles, refractories, metals, and nonmetallic elements. Useless residue and wastes are collected in large output hoppers for landfill. Buffer storage of materials output is on site. Chemical processing operations are shown schematically in figure 5.16, and are detailed in appendix 5E.

Fabrication sectors. The LMF fabrication sector outlined in appendix 5F is an integrated system for the production of finished aluminum or magnesium parts, wire stock, cast basalt parts, iron or steel parts, refractories, and electronics parts. Excepting electronics (Zachary, 1981) there are two major subsystems: (1) the casting subsystem, consisting of a casting robot to make molds, mixing and alloying fur-

naces for basalt and metals, and automatic molding machines to manufacture parts to low tolerance using the molds and alloys prepared; and (2) the laser machining and finishing subsystem, which performs final cutting and machining of various complex or very-close-tolerance parts. The basic operational flowchart for parts fabrication is shown in figure 5.17.

Assembly sectors. Finished parts flow into the automated assembly system warehouse, where they are stored and retrieved by warehouse robots as required. This subsystem provides a buffer against system slowdowns or temporary interruptions in service during unforeseen circumstances. The automated assembly subsystem requisitions necessary parts from the warehouse and fits them together to make subassemblies which are inspected for structural and functional integrity. Subassemblies may be returned to the warehouse for storage, or passed to the mobile assembly and repair robots for transport to the LMF perimeter, either for internal repairs or to be incorporated into working machines and automated subsystems which themselves may contribute to further growth. The basic operational flowchart for SRS parts assembly is shown in figure 5.18, and a more detailed presentation may be found in appendix 5G.

Computer control and communications. The seed computers must be capable of deploying and operating a highly complex, completely autonomous factory system. The original computer must erect an automated production facility, and must be expandable in order to retain control as the LMF grows to its full "adult" size. The computer control subsystem coordinates all aspects of production, scheduling, operations, repairs, inspections, maintenance, and reporting, and must stand ready to respond instantly to emergencies and other unexpected events. Computer control is nominally located at the hub of the expanding LMF disk, and commands in hierarchical fashion a distributed information processing system with sector computers at each node and sector subsystems at the next hierarchical level of control. Communications channels include the transponder network, direct data bus links, and E²ROM messenger chips (firmware) for large data block transfers.

Using ideas borrowed from current industrial practice, top-down structured programming, and biology, Cliff (1981) has devised a system architecture which could perform automated design, fabrication, and repair of complex systems. This architecture, presented in appendix 5H, is amenable to straightforward mathematical analysis and should be a highly useful component of the proposed lunar SRS. Further work in this area should probably include a survey of industrial systems management techniques (Carson, 1959) and the theory of control and analysis of large-scale systems (Sandell et al., 1978).

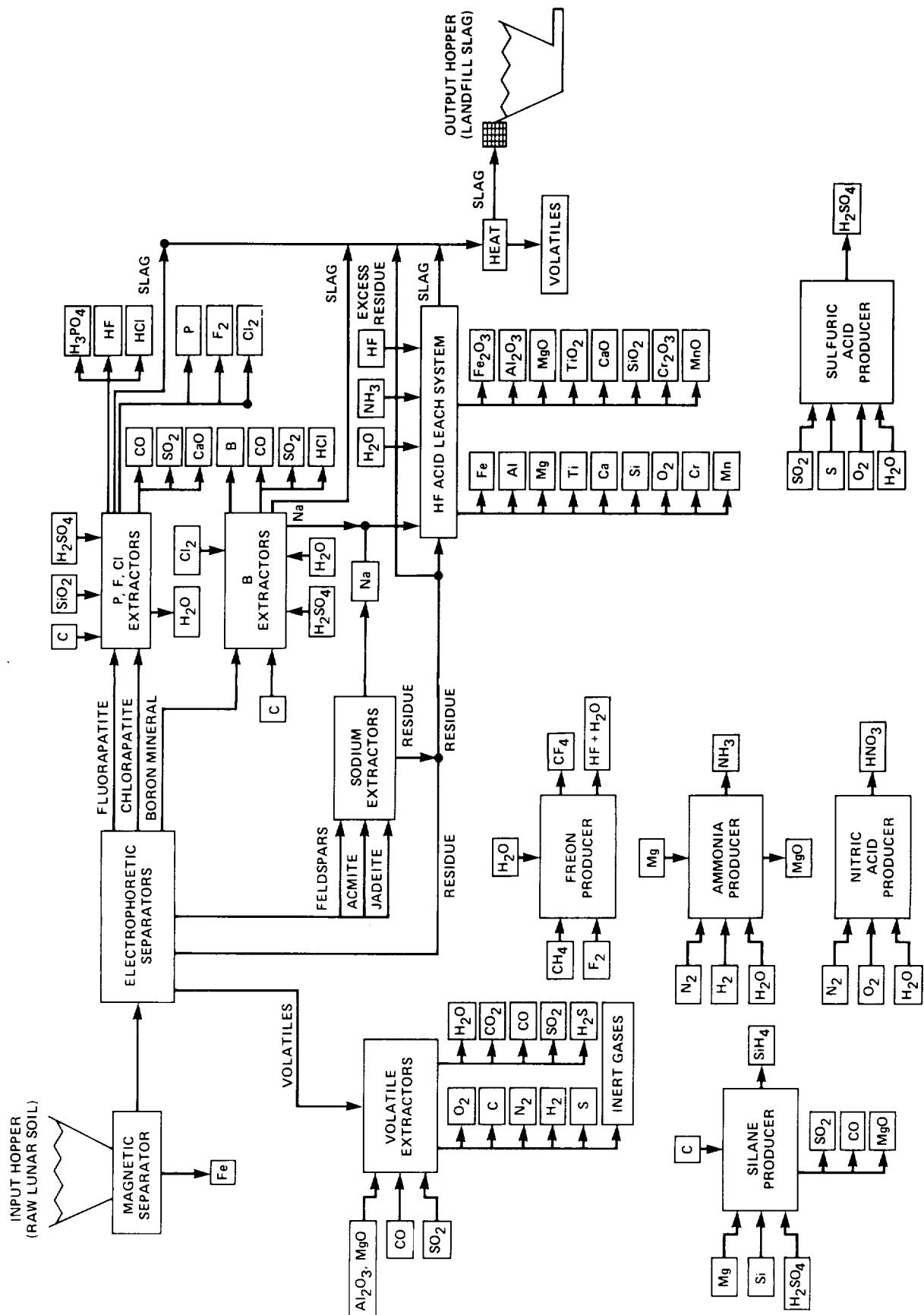


Figure 5.16.—LMF chemical processing sector: Operations.

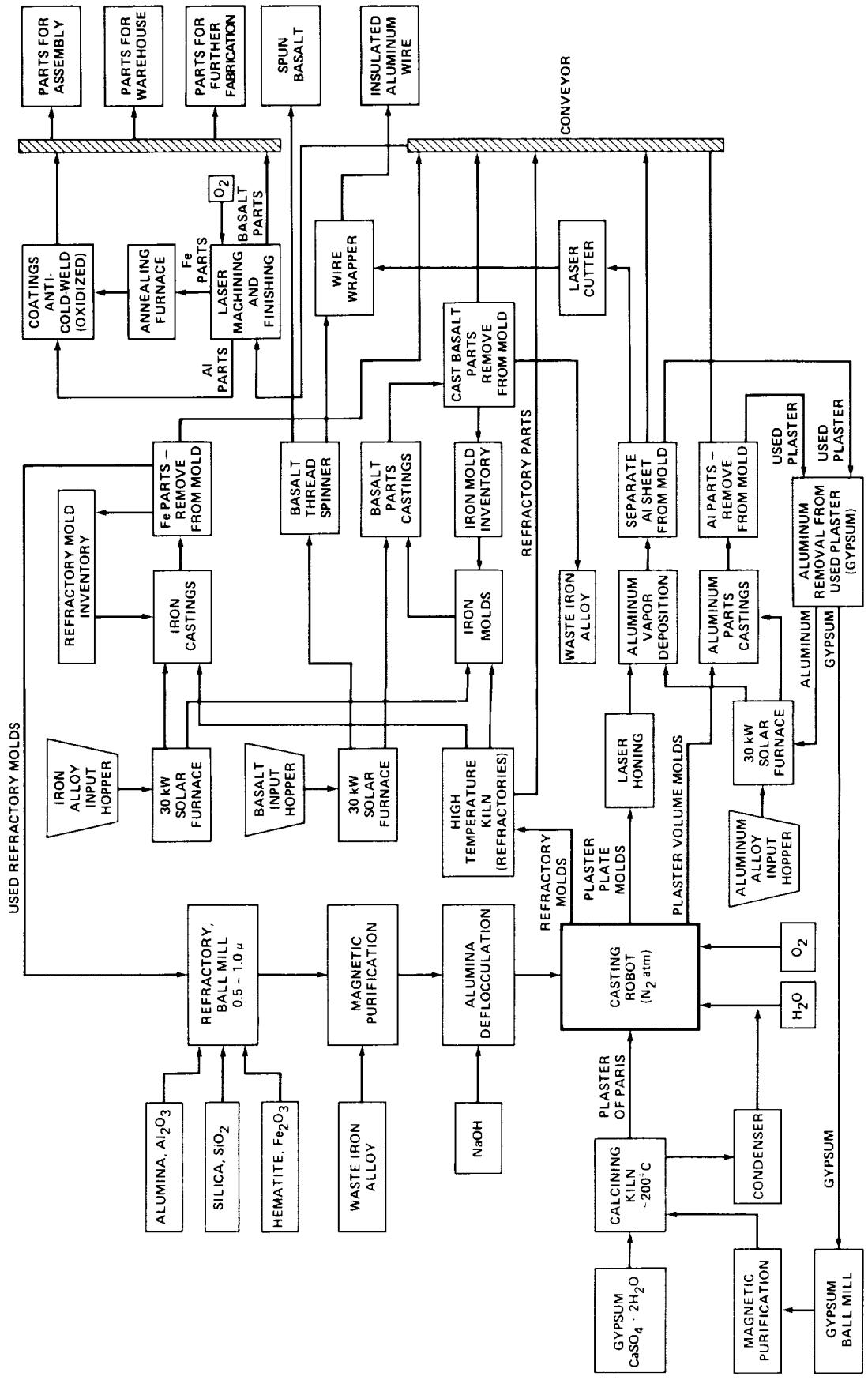


Figure 5.17.—LMF parts fabrication sector: Operations.

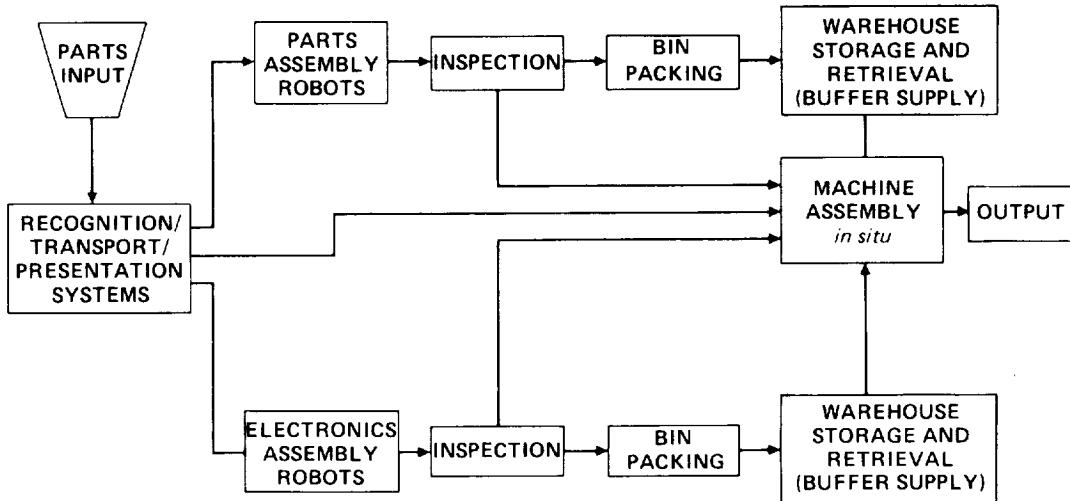


Figure 5.18.—LMF assembly sector: Operations.

In a practical sense, it is quite possible to imagine the lunar SRS operating nonautonomously (Johnsen, 1972). For instance, the *in situ* computer could be used simply as a teleoperation-management system for operations controlled directly by Earth-based workers. Material factory replication would proceed, but information necessary to accomplish this would be supplied from outside. An intermediate alternative would permit the on-site computer to handle mundane tasks and normal functions with humans retaining a higher-level supervisory role. Yet another possibility is that people might actually inhabit the machine factory and help it reproduce manned machine economies can also self-replicate.

Solar canopy: The solar canopy is a “roof” of photovoltaic solar cells, suspended on a relatively flimsy support web of wires, crossbeams and columns perhaps 3–4 m above ground level. The canopy covers the entire LMF platform area and expands outward as the rest of the facility grows. The solar canopy and power grid provide all electrical power for LMF systems. Canopy components may be stationary or may track solar motions using heliostats if greater efficiency is required. A further discussion of canopy design and rationale may be found in appendix 51.

Mass, power, and information requirements. Seed subsystem masses and power requirements scale according to the total system mass assumed. SRS can be reduced indefinitely in size until its components begin to scale nonlinearly. Once this physical or technological limit is reached for any subsystem component, comprehensive redesign of the entire factory may become necessary.

A seed mass of 100 tons was selected in the present study for a number of reasons. First, 100 tons is a credible system mass in terms of foreseeable NASA launch capabilities to the lunar surface, representing very roughly the lunar payload capacity of four Apollo missions to the Moon. Second, after performing the exercise of specifying seed components in some detail it is found that many subsystems are already approaching a nonlinear scaling regime for a 100-ton LMF. For instance, according to Criswell (1980, private communication) the minimum feasible size for a linear-scaling benchtop HF acid-leach plant for materials processing is about 1000 kg; in the present design, two such plants are required with a mass of 1250 kg each. Third, the results of a previous study (Freitas, 1980a) which argued the feasibility of 433-ton seed in the context of an interstellar mission (inherently far more challenging than a lunar factory mission) were compared with preliminary estimates of 15–107 tons for partially self-replicating lunar factories of several different types (O’Neill et al., 1980), and an intermediate trial value of 100 tons selected. The 100-ton figure has appeared in numerous public statements by former NASA Administrator Dr. Robert A. Frosch (lecture delivered at Commonwealth Club, San Francisco, Calif., 1979, and personal communication, 1980) and by others in prior studies (Bekey and Naugle, 1980; Giacconi et al., working paper of the Telefactors Working Group, Woods Hole New Directions Workshop, 1979). Finally, it was decided to use a specific system mass rather than unscaled relative component mass fractions to help develop intuitive understanding of a novel concept which has not been extensively studied before.

For reasons similar to the above, an SRS strawman replication time of 1 year was taken as appropriate. The ranges given in table 5.1, drawn from the analysis presented in appendixes 5B-5I, are estimates of the mass and power requirements of an initial seed system able to manufacture 100 tons of *all* of its own components per working year, hence, to self-replicate. These figures are consistent with the original estimate of a 100 ton circular LMF seed with an initial deployed diameter of 120 m, so feasibility has been at least tentatively demonstrated. However, it must be emphasized that the LMF seed design outlined above is intended primarily as a proof of principle. Numerical values for system components are only crude estimates of what ultimately must become a very complex and exacting design.

Information processing and storage requirements also have been collected and summarized in table 5.1, and lie within the state-of-the-art or foreseeable computer technologies. These calculations, though only rough approximations, quite likely overestimate real needs significantly because of the conservative nature of the assumptions employed. (See also sec. 5.2.3.)

SRS mission overview. In the most general case of fully autonomous operation, a typical LMF deployment scenario might involve the following initial sequence:

(1) The predetermined lunar landing site is mapped from orbit to 1-m resolution across the entire target ellipse.

(2) Seed lands on the Moon, as close to dead center of the mapped target area as possible navigationally.

(3) Mobile assembly and repair robots, assisted by mining robots, emerge from the landing pod and erect a small provisional solar array to provide interim power until the solar canopy is completed.

(4) LMF robots, with the computer, select the precise site where erection of the original seed will commence. This decision will already largely have been made based on orbital mapping data, but ground truth will help refine the estimate of the situation and adjust for unexpected variations.

(5) Mobile robots emplace the first three stations of the transponder network (the minimum necessary for triangulation), calibrate them carefully, and verify that the system is in good working order.

(6) Mining robots equipped with grading tools proceed to the construction site and level the local surface.

TABLE 5.1.— SEED MASS AND POWER REQUIREMENTS ESTIMATES

Seed subsystem	Estimated mass of 100 ton/yr seed, kg	Estimated power of 100 ton/yr seed, W	Computer processor, bits to operate	Computer memory, bits to describe
Transponder network	1,000	---	10^5 ?	10^6 ?
Paving robots	12,000	Up to 10^4	$1\text{-}10 \times 10^6$	$10^7\text{-}10^8$
Mining robots	4,400	Up to 10^4	$4\text{-}7 \times 10^8$	10^9
Chemical processing sector (S)	15,300-76,400	380,000-11,000,000	9.4×10^7	3.1×10^9
Fabrication sector (S)				
Electronics	(3,000)			
Floor map				(10^9)
Totals	137-20,400	270-345,000	10^{10}	10^{11}
Assembly sector (S)				
Assembly robots	83-1,150	83-19,600	10^9	10^{10}
Warehouse subsystem	1,000	10,000	10^7	10^8
Floor map				10^9
Automated transport vehicles	1,000	6,000	10^7	10^8
Mobile assembly and repair robots	4,000	40,000	4×10^9	4×10^{10}
Computer central orbital site map	2,200	37,000	(1.6×10^{10})	1.6×10^{10}
Solar canopy	22,000	---	2×10^7	2×10^8
Totals	63,100-145,600	0.47 MW-11.5 MW	$15.5\text{-}15.8 \times 10^9$	272×10^9
Nominal annual seed output	100,000	1.7 MW		

(7) Five paving robots disembark and begin laying down the seed platform in square grids. This requires one working year for completion.

(8) When a sufficiently large platform section has been completed, seed mobile robots transfer the main computer to a place prepared for it at the center of the expanding platform disk.

(9) Erection of the solar canopy begins, followed by each of the seed sectors in turn, starting with the chemical processing. Total time to unpack the landing pod after moonfall is one working year, conducted in parallel with paving and other activities. The completed seed factory unit, unfurled to a 120 m diam on the surface of the Moon 1 year after landing, might appear as shown in figure 5.19.

The LMF has two primary operational phases — growth and production. The optimal program would probably be to "bootstrap" (grow) up to a production capacity matching current demand, then reconfigure for production until demand increases, thus necessitating yet further growth

(O'Neill et al., 1980). Growth and production of useful output may proceed sequentially, cyclically, or simultaneously, though the former is preferred if large subsystems of the lunar factory must be reconfigured to accommodate the change.

The LMF also may exhibit replicative behavior if and when necessary. Replicas of the original seed could be constructed much like regular products and dispatched to remote areas, either to increase the total area easily subject to utilization or to avoid mortality due to depletion of local resources or physical catastrophes. The scheduling of factory operational phases is very flexible, as shown schematically in figure 5.20, and should be optimized for each mission and each intended use.

5.3.5 Lunar SRS Growth and Productivity

As the study progressed, the team noted a developing convergence between the two designs for SRS described in

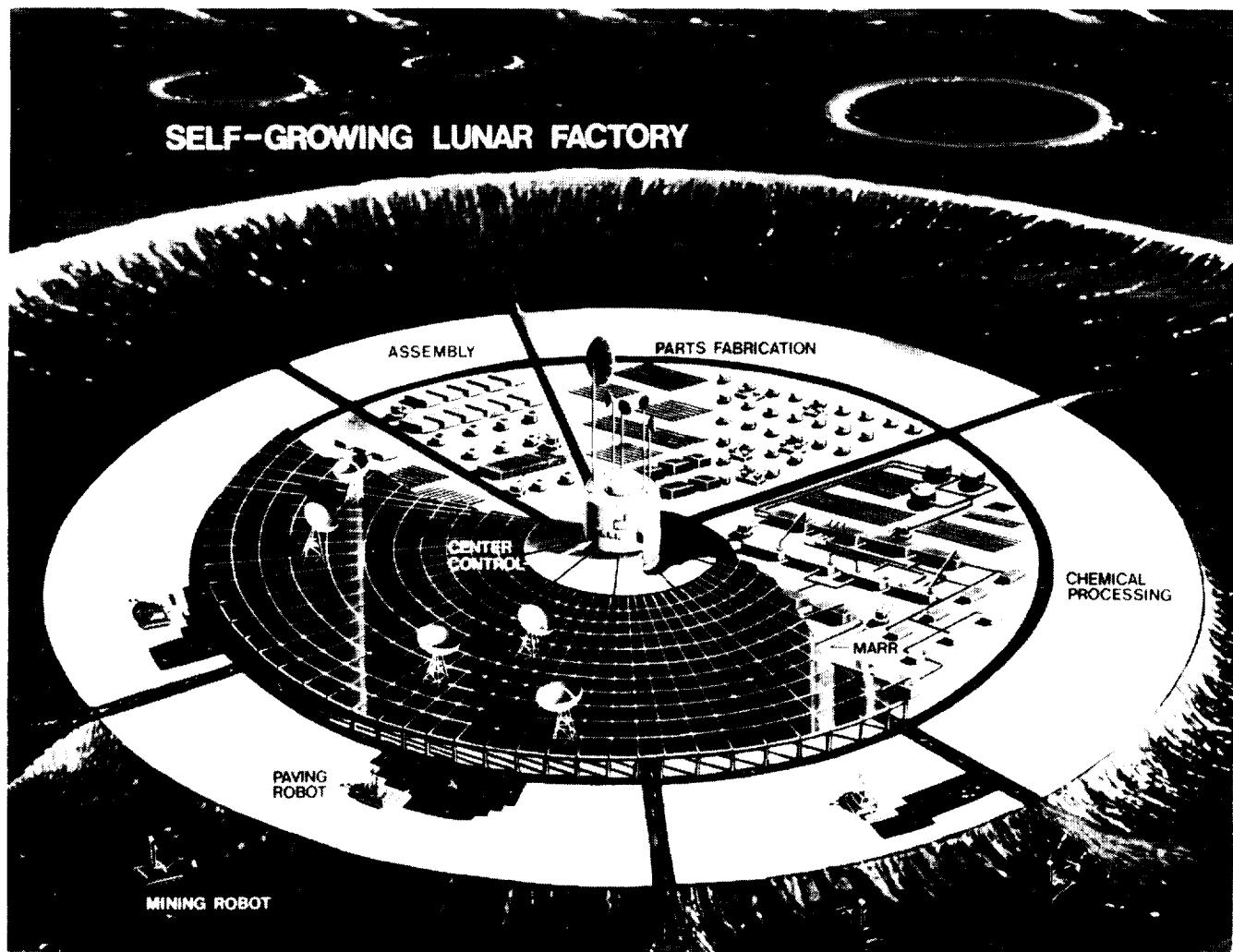


Figure 5.19.— Self-growing lunar factory.

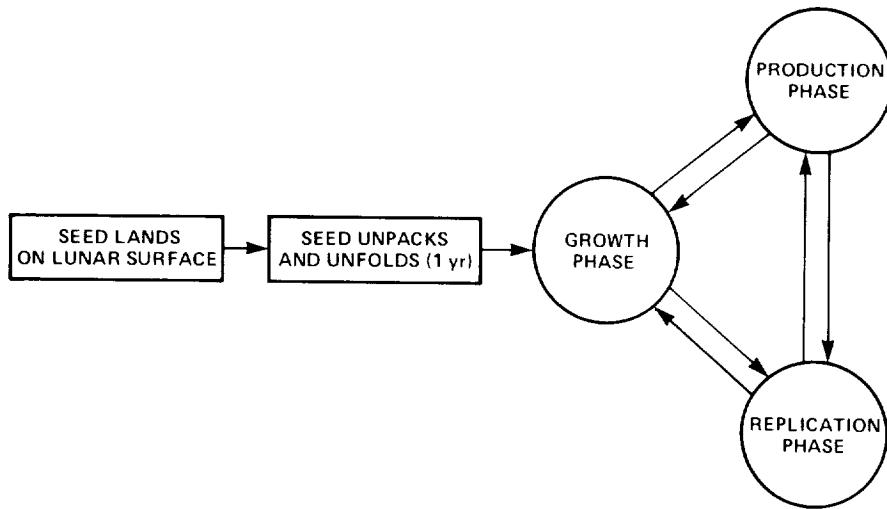


Figure 5.20. – Flexible scheduling of LMF operational phases.

sections 5.3.3 and 5.3.4. Both require three major subsystems — materials processing, fabrication, and assembly — plus a variety of support systems, and each is capable of replication and useful production. Both display exponential expansion patterns.

Of course, in a finite environment exponential growth cannot continue indefinitely. Geometrical arguments by Taneja and Walsh (1980, Summer Study document) suggest that planar packing of triangular, cubic, or hexagonal units can expand exponentially only for as many generations as each unit has sides, assuming that once all sides are used up no further doubling can occur by the enclosed unit. Growth is quadratic from that time on.

However, in real physical systems such as the developing LMF, enclosure need not preclude material communication with exterior units. Selected ramification of communication, control, and materials transportation channels or internal component rearrangement, reconfiguration, or specialization can prevent “starvation” in the inner regions of the expanding system. Hence, SRS exponential growth may continue until limited either by purposeful design or by the specific configuration of the external environment. Assuming that a 100-ton seed produces 100 tons/year of the same materials of which it is composed, then if T is elapsed time and N is number of seed units or seed mass-equivalents generated during this time, $T = 1 + \log_2 N$ for simple exponential “doubling” growth. (There is no replication in the first year, the time required for initial setup.) If P is productivity in tons/year, then $P = 100 \log_2 N$.

However, the above is valid only if each unit works only on its own replica. If two or more units cooperate in the construction of a single replica, still more rapid “fast exponential” growth is possible. This is because new complete replicas or LMF subsystems are brought on line sooner, and hence may begin contributing to the exponentiation earlier than before. Using the above notation, the “fast-exponential” growth rate is given by $T = 1 + 1/2 + \dots + 1/N$

in the optimum case where all available machines contribute directly to the production of the next unit.

Growth rates and productivities are tabulated for exponential and “fast-exponential” expansion in table 5.2. Note that in just 10 years the output of such a facility could grow to approximately one million tons per year. If allowed to expand for 18 years without diversion to production, the factory output could exponentiate to more than 4×10^9 tons per year, roughly the entire annual industrial output of all human civilization.

Useful SRS products may include lunar soil thrown into orbit by mass drivers for orbital processing, construction projects, reaction mass for deep space missions, or as radiation shielding; processed chemicals and elements, such as oxygen to be used in space habitats, as fuel for interorbital vehicles, and as reaction mass for ion thrusters and mass drivers; metals and other feedstock ready-made for space construction or large orbital facilities for human occupation (scientific, commercial, recreational, and medical); components for large deep-space research vessels, radio telescopes, and large high-power satellites; complex devices such as machine shop equipment, integrated circuits, sophisticated electronics gear, or even autonomous robots, teleoperators, or any of their subassemblies; and solar cells, rocket fuels, solar sails, and mass driver subassemblies. Also, a 100-ton seed which has undergone thousand-fold growth or replication represents a 2 GW power generating capacity, plus a computer facility with a 16,000 Gbit processing capability and a total memory capacity of 272,000 Gbits. These should have many useful applications in both terrestrial and space industry.

5.3.6 Closure in Self-Replicating Systems

Fundamental to the problem of designing self-replicating systems is the issue of closure. In its broadest sense, this

TABLE 5.2.— GROWTH RATES AND PRODUCTIVITY FOR EXPONENTIAL SRS EXPANSION

Calendar years	Working years, T	Exponential growth, $\tau = 1$ yr		“Fast-exponential” growth, $\tau = 1$ yr	
		Number of units, N	System productivity, tons/yr	Number of units, N	System productivity, tons/yr
0	0	0	0	0	0
2	1	1	100	1	100
4	2	2	200	4	400
6	3	4	400	11	1,100
8	4	8	800	31	3,100
10	5	16	1,600	83	8,300
12	6	32	3,200	227	22,700
14	7	64	6,400	616	61,600
16	8	128	12,800	1,674	167,400
18	9	256	25,600	4,550	455,000
20	10	512	51,200	12,367	1,236,700
22	11	1,024	102,400	33,617	3,361,700
24	12	2,048	204,800	91,380	9,138,000
26	13	4,096	409,600	248,398	24,839,800
28	14	8,192	819,200	675,215	67,521,500
30	15	16,384	1,638,400	1,835,426	183,542,600
32	16	32,768	3,276,800	4,989,205	498,920,500
34	17	65,536	6,553,600	13,562,066	1,356,206,600
36	18	131,072	13,107,200	36,865,517	3,686,551,700
38	19	262,144	26,214,400	100,210,865	10,021,086,500
40	20	524,288	52,428,800	272,401,372	27,240,137,200
					(~2 km-wide asteroid/yr)

(About 3 billion seed units would completely cover the entire lunar surface)

issue reduces to the following question: Does system *function* (e.g., factory output) equal or exceed system *structure* (e.g., factory components or input needs)? If the answer is negative, the system cannot independently fully replicate itself; if positive, such replication may be possible.

Consider, for example, the problem of parts closure. Imagine that the entire factory and all of its machines are broken down into their component parts. If the original factory cannot fabricate every one of these items, then parts closure does not exist and the system is not fully self-replicating.

In an arbitrary system there are three basic requirements to achieve closure:

(1) Matter closure can the system manipulate matter in all ways necessary for complete self-construction?

(2) Energy closure can the system generate sufficient energy and in the proper format to power the processes of self-construction?

(3) Information closure can the system successfully command and control all processes required for complete self-construction?

Partial closure results in a system which is only partially self-replicating. Some vital matter, energy, or information

must be provided from the outside or the machine system will fail to reproduce. For instance, various preliminary studies of the matter closure problem in connection with the possibility of “bootstrapping” in space manufacturing have concluded that 90–96% closure is attainable in specific nonreplicating production applications (Bock, 1979; Miller and Smith, 1979; O’Neill et al., 1980). The 4–10% that still must be supplied sometimes are called “vitamin parts.” These might include hard-to-manufacture but lightweight items such as microelectronics components, ball bearings, precision instruments and others which may not be cost-effective to produce via automation off-Earth except in the longer term. To take another example, partial information closure would imply that factory-directive control or supervision is provided from the outside, perhaps (in the case of a lunar facility) from Earth-based computers programmed with human-supervised expert systems or from manned remote teleoperation control stations on Earth or in low Earth orbit.

The fraction of total necessary resources that must be supplied by some external agency has been dubbed the “Tukey Ratio” (Heer, 1980). Originally intended simply as an informal measure of basic materials closure, the most

logical form of the Tukey Ratio is computed by dividing the mass of the external supplies per unit time interval by the total mass of all inputs necessary to achieve self-replication. (This is actually the inverse of the original version of the ratio.) In a fully self-replicating system with no external inputs, the Tukey Ratio thus would be zero (0%).

It has been pointed out that if a system is "truly isolated in the thermodynamic sense and also perhaps in a more absolute sense (no exchange of information with the environment) then it cannot be self-replicating without violating the laws of thermodynamics" (Heer, 1980). While this is true, it should be noted that a system which achieves complete "closure" is not "closed" or "isolated" in the classical sense. Materials, energy, and information still flow into the system which is thermodynamically "open"; these flows are of indigenous origin and may be managed autonomously by the SRS itself without need for direct human intervention.

Closure theory. For replicating machine systems, complete closure is theoretically quite plausible; no fundamental or logical impossibilities have yet been identified. Indeed, in many areas automata theory already provides relatively unambiguous conclusions. For example, the theoretical capability of machines to perform "universal computation" and "universal construction" can be demonstrated with mathematical rigor (Turing, 1936; von Neumann, 1966; see also sec. 5.2), so parts assembly closure is certainly theoretically possible.

An approach to the problem of closure in real engineering systems is to begin with the issue of parts closure by asking the question: can a set of machines produce all of its elements? If the manufacture of each part requires, on average, the addition of ≥ 1 new parts to product it, then an infinite number of parts are required in the initial system and complete closure cannot be achieved. On the other hand, if the mean number of new parts per original part is < 1 , then the design sequence converges to some finite

ensemble of elements and bounded replication becomes possible.

The central theoretical issue is: can a real machine system itself produce and assemble *all* the kinds of parts of which it is comprised? In our generalized terrestrial industrial economy manned by humans the answer clearly is yes, since "the set of machines which make all other machines is a subset of the set of all machines" (Freitas et al., 1981). In space a few percent of total system mass could feasibly be supplied from Earth-based manufacturers as "vitamin parts." Alternatively, the system could be designed with components of very limited complexity (Heer, 1980). The minimum size of a self-sufficient "machine economy" remains unknown.

Von Tiesenhausen and Darbro (1980) similarly argue that a finite set of machines can produce any machine element. Their reasoning, outlined in figure 5.21, is as follows:

(1) If all existing machines were disassembled into their individual parts there would obviously be a finite number of parts, many of them identical, and a large number would be of common categories like shafts, motors, wiring, etc. The only differences between the machines would be a different selection, different arrangement, and different dimensions of this finite number of parts.

(2) A finite number of parts involves a finite number of machine operations, this number being less than the number of parts because some machines can make more than one kind of parts.

(3) Therefore, the number of machines is finite and less than the number of operations.

This reasoning can then be generalized to say: "Every existing machine can be reduced to a finite set of machine elements, and there exists a finite set of machine operations." (Still, of course, a limited number of standard elements should be developed and machine operations limited as

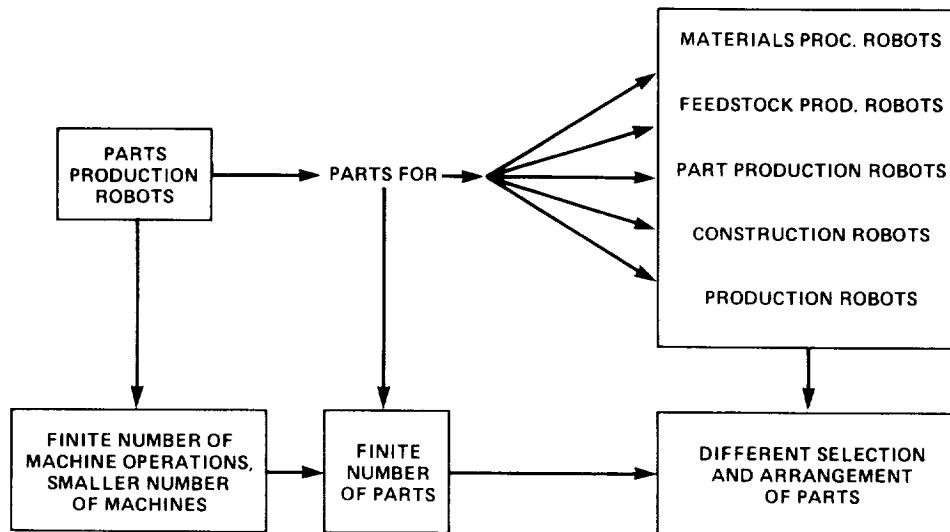


Figure 5.21.- Closure of SRS parts production.

much as practical by substitution, in order to minimize the number of parts and machine operations.)

Similar arguments may be applied to materials processing and feedstock production. There exists a finite number of different materials anywhere. There is a finite number of materials processes which is less than the number of materials because single processes result in various materials (e.g., silicon and oxygen). Hence, there is a finite number of materials processing robot systems needed for an SRS. Also, there is a finite and rather limited number of feedstock requirements such as bars, rods, ingots, plates, etc. The number of materials is much less than the number of parts; therefore, a finite number of parts fabrication robots is required for an SRS.

Closure engineering. In actual practice, the achievement of full closure will be a highly complicated, iterative engineering design process. Every factory system, subsystem, component structure, and input requirement (Miller and Smith, 1979) must be carefully matched against known factory output capabilities. Any gaps in the manufacturing flow must be filled by the introduction of additional machines, whose own construction and operation may create new gaps requiring the introduction of still more machines.

The team developed a simple iterative procedure for generating designs for engineering systems which display complete closure. The procedure must be cumulatively iterated, first to achieve closure starting from some initial design, then again to eliminate overclosure to obtain an optimally efficient design. Each cycle is broken down into a succession of subiterations which ensure three additional dimensions of closure:

- (1) Qualitative closure – can, say, all parts be made?
- (2) Quantitative closure – can, say, enough parts be made?
- (3) Throughput closure – can parts be made fast enough?

In addition, each subiteration sequence is further decomposed into design cycles for each factory subsystem or component, as shown in figure 5.22.

The procedure as outlined, though workable in theory, appears cumbersome. Further work should be done in an attempt to devise a more streamlined, elegant approach.

Quantitative materials closure – numerical results. In the context of materials processing, “closure” is a relationship between a given machine design and a given particular substrate from which the machine’s elemental chemical constituents are to be drawn. Hence the numerical demonstration of closure requires a knowledge of the precise composition both of the intended base substrate to be utilized and of the products which the SRS must manufacture from that substrate. Following a method suggested by the work of Freitas (1980a), a modified “extraction ratio” R_n

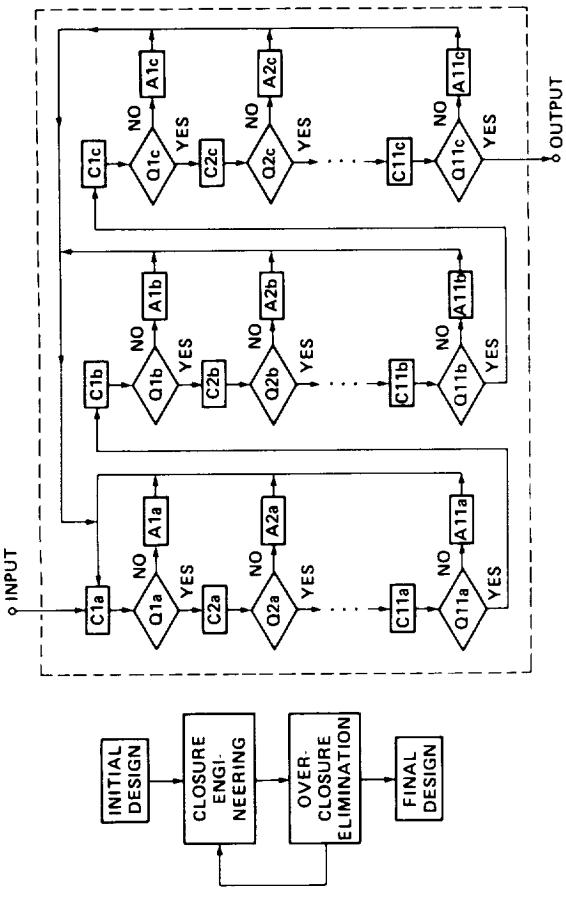
is defined as the mass of raw substrate material which must be processed (input stream) to obtain a unit mass of useful system output having the desired mass fraction of element n (output stream).

Consider the significance of the extraction ratio to the problem of materials closure. Assume that the final product is to be composed of elements x , y , and z . An $R_x = 1$ means that 1 kg of lunar soil contains exactly the mass of element x needed in the manufacture of 1 kg of the desired output product. On the other hand, $R_y = 10$ means that 10 kg of lunar regolith must be processed to extract all of element y required in 1 kg of final product. The difference between R_x and R_y may signify that y is more rare in lunar soil than x , or that the two elements are equally abundant but ten times more y than x is required (by weight) in the final product. When the output stream is identical to the machine processing system itself, then the system is manufacturing more of itself -- self-replicating -- and the extraction ratio becomes an index of system materials closure on an element-by-element basis.

The total net extraction ratio R is some function of the individual extraction ratios R_n , and depends on the methods of materials processing employed. At worst, if only one element is recovered from a given mass of input stream (“parallel processing”), then R is the sum of all R_n . At best, if the input stream is processed sequentially to extract all desired elements in the necessary amounts (“serial processing”), then R is driven solely by the R_n of the element most difficult to extract, say, element z . That is, $R = (R_n)_{\max} = R_z$, which is always equal to or smaller than the sum of all R_n . As serial processing should dominate in the lunar factory the latter formula is assumed for purposes of the present calculations. Note that R_n can be less than 1 for individual elements, but for an entire machine system R must always be greater than or equal to 1.

As a general rule, a low value for R implies that the system is designed for low mass throughput rates and is built from relatively few different chemical elements. A high value of R implies that many more elements are necessary and that a higher mass throughput rate will be accommodated to obtain them.

The “closure” of a given output stream (product) relative to a specified input stream (substrate) is computed by treating R as an independent variable. If I_n is the concentration of element n in mineral form in the input stream of lunar soil (kg/kg), E_n is the efficiency of chemical extraction of pure element n from its mineral form which is present in lunar soil (kg/kg), and O_n is the concentration of element n in the desired factory output stream (kg/kg), then $R_n = O_n/E_n I_n$. Closure C_n for each element is defined as the mass of pure element n available in a system with a total net extraction ratio R per unit mass of output stream. For any given element, if $R \geq R_n$ then all pure element n needed is already available within the system. In this case, $C_n = O_n$. On the other hand, if $R < R_n$ then the choice of R is too low; all the pure element n needed cannot be



EXPLANATION OF SYMBOLS

PROCEDURAL QUERIES = Q

- Q1a – CAN SYSTEM ASSEMBLE ALL MACHINES OF WHICH IT IS COMPRISED?
 Q2a – CAN SYSTEM ASSEMBLE ALL SUBASSEMBLIES OF WHICH ITS MACHINES ARE COMPRISED?
 Q3a – CAN SYSTEM MACHINE ALL PARTS OF WHICH ITS MACHINES ARE COMPRISED?
 Q4a – CAN SYSTEM EXTRACT ALL MATERIALS OF WHICH ITS PARTS ARE COMPRISED?
 Q5a – CAN SYSTEM MINE ALL MINERALS OF WHICH ITS MATERIALS ARE COMPRISED?
 Q6a – CAN SYSTEM TRANSPORT ALL (MACHINES/SUBASSEMBLIES/PARTS/MATERIALS/MINERALS) OF WHICH IT IS COMPRISED?
 Q7a – CAN SYSTEM VERIFY ALL (MACHINES/SUBASSEMBLIES/PARTS/MATERIALS/MINERALS) OF WHICH IT IS COMPRISED?
 Q8a – CAN SYSTEM WAREHOUSE ALL (MACHINES/SUBASSEMBLIES/PARTS/MATERIALS/MINERALS) OF WHICH IT IS COMPRISED?
 Q9a – CAN SYSTEM REPAIR ALL SUBSYSTEMS OF WHICH IT IS COMPRISED?
 Q10a – CAN SYSTEM COMPUTER RECONTROL ALL SUBSYSTEMS OF WHICH IT (THE SYSTEM) IS COMPRISED?
 Q11a – CAN SYSTEM ASSEMBLE ENOUGH MACHINES TO REPLICATE THE ORIGINAL SYSTEM?
 Q11b – CAN SYSTEM ENERGY PLANT PRODUCE ENOUGH ENERGY TO PERMIT REPLICATION OF THE ORIGINAL SYSTEM?
 Q11c – CAN SYSTEM ASSEMBLE MACHINES FAST ENOUGH TO REPLICATE THE ORIGINAL SYSTEM WITHIN ESTABLISHED TIME CONSTRAINTS?
 Q11d – CAN SYSTEM ENERGY PLANT PRODUCE ENOUGH ENERGY FAST ENOUGH (HIGH ENOUGH POWER) TO REPLICATE THE ORIGINAL SYSTEM WITHIN ESTABLISHED TIME CONSTRAINTS?

SYSTEM COMPONENTS LISTS = C

- | | | | |
|------------|---------------------------|---------------|--|
| C1a, b, c | MACHINES | A1a, b, c | – NEW MACHINES |
| C2a, b, c | SUBASSEMBLIES | A2a, b, c | – NEW SUBASSEMBLY-MAKING MACHINES |
| C3a, b, c | PARTS | | |
| C4a, b, c | MATERIALS | | |
| C5a, b, c | MINERALS (SUBSTRATE) | | |
| C6a, b, c | TRANSPORTATION FACILITIES | | |
| C7a, b, c | VERIFICATION FACILITIES | | |
| C8a, b, c | STORAGE FACILITIES | | |
| C9a, b, c | REPAIR FACILITIES | | |
| C10a, b, c | COMPUTER FACILITIES | A11a, b, c | – NEW ENERGY FACILITY-MAKING MACHINES |
| C11a, b, c | ENERGY FACILITIES | A \approx a | – Add new machines to make ... |
| | | A=b | – Add new machines to make ... |
| | | A \neq c | – increase replication time available or change machine design for ... |

Figure 5.22. Generalized closure engineering design cycles.

recovered, and more lunar soil must be processed to make up the difference if 100% closure is to be achieved. In this case, $C_n = O_n(R/R_n)$, since the closure deficit is measured by the ratio of the chosen R to the actual R_n of the given element (i.e., how much the factory has, divided by how much the factory actually needs). Total net system closure C is simply the sum of all C_n for all elements n required in the output stream of the SRS factory (Freitas and Zachary, 1981).

To estimate the quantitative materials closure for the lunar SRS baseline designs proposed in sections 5.3.3 and 5.3.4, three different approaches were taken in an attempt to converge on a useful estimate of the composition of the output stream necessary for LMF self-replication. First, the "seed" element distribution given by Freitas (1980a) in the context of a self-reproducing exploratory spaceprobe was adopted. These figures are derived from published data on the material consumption of the United States (the world's largest factory) during the years 1972-1976 (U.S. Bureau of Mines, 1978; U.S. Bureau of the Census, 1977, 1978). A second but less comprehensive measure called "demandite" is based on 1968 U.S. consumption data (Goeller and Weinberg, 1976). A molecule of "nonfuel demandite" is the average nonrenewable resource used by humans, less fuel resources (Waldron et al., 1979). Third, the direct estimate of LMF elemental composition presented in appendix 5E was used to obtain additional trial values for O_n . (Appendix 5E also represents a first attempt to deal with qualitative materials closure for SRS.) In all cases the input stream was assumed to consist of lunar maria regolith, with values for I_n averaged from published data (Phinney et al., 1977) and listed in table 5.3. Following earlier work, for simplicity all efficiencies E_n were taken to be 0.93 (Rao et al., 1979; Williams et al., 1979).

The closures calculated from these data are plotted against extraction ratio in figure 5.23. (Data for the human body are included for purposes of comparison.) Note that 100% closure ($C = 1$) is achieved for the "U.S. Industrial" estimate (84 elements of the spaceprobe "seed") at $R = 2984$; for "Demandite" (28 elements) at $R = 1631$; and for the appendix 5E "LMF" (18 elements) at $R = 45$. This suggests that the fewer the number of different elements, and the more common and more efficiently extractable are the elements the factory system needs for replication to occur, the lower will be the total mass of raw materials which must be processed by the LMF.

Note also that in all three cases, virtually complete (>90%) closure is achieved for extraction ratios of 2 to 14. The incremental gains in closure after 90% are purchased only at great price -- from 1 to 3 orders of magnitude more raw materials mass must be processed to achieve the last bit of full materials autonomy. Two conclusions may be drawn from this observation. First, for any given SRS design it may well be more economical to settle for 90-95% system closure and then import the remaining 5-10% as "vitamins"

TABLE 5.3. AVERAGE CHEMICAL ELEMENT ABUNDANCES IN LUNAR MARIA

Element	Abundance	Element	Abundance
Al	6.80%	Ho	3.73 ppm
Ca	7.88%	I	2.00 ppb
Cr	0.264%	In	32.9 ppb
Fe	13.2%	Ir	6.32 ppb
K	0.113%	La	17.2 ppm
Mg	5.76%	Li	12.9 ppm
Mn	0.174%	Lu	1.22 ppm
Na	0.290%	Mo	0.520 ppm
O	41.3%	N	95.4 ppm
P	0.066%	Nb	19.6 ppm
S	0.125%	Nd	38.2 ppm
Si	20.4%	Ne	2.75 ppm
Ti	3.10%	Ni	169 ppm
		Os	12.9 ppb
Ag	45.2 ppb	Pb	3.11 ppm
Ar	0.800 ppm	Pd	12.3 ppb
As	0.206 ppm	Pr	7.20 ppm
Au	2.66 ppb	Rb	3.21 ppm
B	4.78 ppm	Re	1.36 ppb
Ba	195 ppm	Rh	0.192 ppm
Be	2.63 ppm	Ru	0.231 ppm
Bi	3.19 ppb	Sb	22.1 ppb
Br	0.178 ppm	Sc	48.8 ppm
C	104 ppm	Se	0.306 ppm
Cd	0.197 ppm	Sm	10.9 ppm
Ce	48.8 ppm	Sn	0.900 ppm
Cl	25.6 ppm	Sr	167 ppm
Co	40.3 ppm	Ta	1.26 ppm
Cs	0.392 ppm	Tb	2.58 ppm
Cu	14.4 ppm	Te	0.0545 ppm
Dy	15.3 ppm	Th	2.50 ppm
Er	9.24 ppm	Tl	1.61 ppb
Eu	1.77 ppm	Tm	1.42 ppm
F	174 ppm	U	0.805 ppm
Ga	4.99 ppm	V	114 ppm
Gd	14.3 ppm	W	0.358 ppm
Ge	0.626 ppm	Y	84.2 ppm
H	54.8 ppm	Yb	8.40 ppm
He	28.5 ppm	Zn	23.4 ppm
Hf	7.77 ppm	Zr	311 ppm
Hg	0.019 ppm		

from Earth. Second, in those applications where 100% closure (full materials autonomy) is desirable or required, great care must be taken to engineer the self-replicating system to match the expected input substrate as closely as possible. This demands, in the case of quantitative materials closure, a design which minimizes the value of R , thus optimizing the use of abundantly available, easily extractable elements.

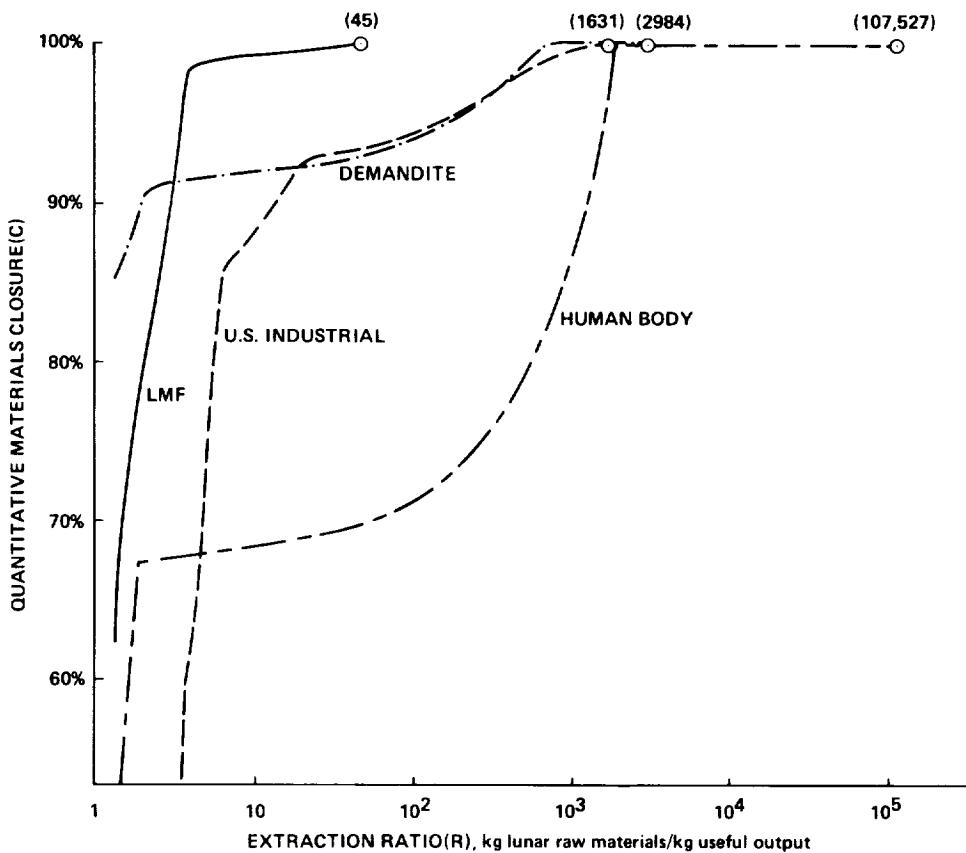


Figure 5.23.— Quantitative materials closure data for various self-replicating systems.

5.3.7 Conclusions

The team reached the following major conclusions regarding the feasibility of self-replicating machine systems:

- The basic concept of physical machine systems capable of self-replication appears credible both from a theoretical and a practical engineering standpoint.
- It is reasonable to begin designing replicating systems based on current knowledge and state-of-the-art technology, but final design definition will require significant further research.
- Complete systems closure is achievable in principle, though partial closure may be more feasible from an economic and pragmatic engineering standpoint in the near term.
- It is feasible to begin immediate work on the development of a simple demonstration SRS on a laboratory scale, with phased steps to more sophisticated levels as the technology is proven and matures.

5.4 Applications

Having shown that machine SRS is, in principle, both theoretically possible and feasible in terms of engineering

systems design, their usefulness in some economic or commercial sense remains to be demonstrated. That is, what might such systems permit humankind to do that could not be done before? The main advantage in using SRS over other methods of space exploration and industrialization is that a very large capability for performing any desired task can be rapidly achieved at arbitrary remote locations, starting with a relatively small investment of time, money, energy, and mass in the original "seed" mechanism.

The team has identified four general criteria for determining the most probable and profitable application of replicating systems technology:

- (1) A large number of identical or similar products is required;
- (2) Excessively long production periods for alternate approaches are required;
- (3) Raw materials or parts are available onsite; and
- (4) Sufficient physical space is available for replication.

Each of these criteria should be applicable, or largely so, in a specific case before the use of SRS technologies is considered.

Replicating systems will find many applications on Earth, in near-Earth and lunar space, throughout the Solar System, and in the interstellar realm, for both exploration and utilization. SRS also provides a number of fascinating applications in basic and applied research in automata theory, theoretical biology, experimental evolution, and machine intelligence and robotics architecture.

5.4.1 Terrestrial Applications

The early development of replicating systems technology on Earth will be the history of modern industrial automation. The United States at one time enjoyed the highest productivity in the world, and still partakes of the prosperity that has brought. Recently, however, competition from other nations who are more rapidly automating their industries is seriously eroding the U.S. position of leadership. The resulting economic forces are impelling domestic industry to accelerate the automation of its factories.

The space program is viewed by many as a high-technology venture which predominantly makes use of computers, robot spacecraft, and other trappings of automation. In reality, NASA's activities are strongly people-intensive. For example, large teams of trained technicians and scientists are required to operate a robot space probe by remote control. The same economic forces at work in the marketplace are forcing NASA to rethink its traditional way of doing business. Not only will there be more automation in the space program for this reason, but also there will be missions that are difficult or impossible to conduct without using advanced machine intelligence and robotics technologies.

The harsh environment of space, the significant costs of life support systems for human beings and of "man-rating" space systems for safety, and the communications problems caused by the immense distances involved in interplanetary travel have given NASA additional incentives to develop systems of total automation beyond those commonly employed in industry. The sheer magnitude of many potentially interesting missions requires massive automation. Accordingly, NASA should strongly participate in automation research and development in anticipation of spinoffs to industry of great potential value. The agency also should closely monitor industrial R&D efforts, remaining alert for new developments on the commercial front which might prove beneficial to the space program. The infusion of NASA funds at critical points could allow the agency to exert subtle influence on industrial development so as to provide for NASA's special needs at less cost than an independent program to achieve the same ends.

Similarly, the Department of Defense (DOD) is embarking upon an ambitious program of industrial automation. The aim is to produce war materiel in the most economical and flexible manner possible, and to shorten the time between concept and field deployment of weapons systems.

Much of the DOD effort will produce results useful for the space program. To take maximum advantage of this, NASA should maintain close liaison with DOD and should join in various cooperative efforts in areas of overlapping interests.

Computer-aided design (CAD), manufacturing (CAM), and testing (CAT), and robotics. Automation for replication will require extensive application of computer science and robotics. At the initial stage of development, and during periods when repair or reconstruction operations must be performed, computers can be used in many ways to aid the design process (CAD). They are excellent for generating and maintaining documentation. Computer-executed graphics are invaluable in assisting human operators to visualize complex objects in the absence of a real, physical construction. Simulation using computer models is used in place of, or as a cost-saving adjunct to, physical models or prototypes. Recent developments in machine intelligence research has made far easier the complete automation of the entire design process. Ultimately, the capability will exist for a human to carry on a dialog with a computer system in which the person merely defines the functional specifications of the desired product and the computer determines the remaining design details autonomously.

Computers have been used in manufacturing (CAM) for more than two decades. The most common modern application is business data processing. Computerized inventory control and scheduling are two promising uses rapidly gaining prominence today. Process control using analog computers began many years ago in chemical plants, steel mills, and paper mills. Newer facilities rely instead upon digital computing. An important subset of process control is numerical control (N/C) of machine tools, with instructions traditionally recorded on punched paper tape. Today it is feasible to connect N/C machine tools directly to a computer able to generate and store instructions in electronic memory, and increasingly this is being done, especially in the aerospace industry.

Computers can also be used to great advantage in the testing of products (CAT). (This is distinguished from measurements of process variables, which is considered a process control function.) Highly complex products such as microprocessor integrated circuits cannot realistically be tested without the aid of computer technology. A standard interface protocol (the IEEE-488 bus) has been defined for the interfacing of test instrumentation to a host computer.

In the context of a factory, robotics generally is understood to refer to materials handling and assembly functions. Typical operations include loading/unloading machine tools and spot-welding automobile bodies. Hard automation (special-purpose robots of very limited versatility) commonly are used in applications requiring high volume output. But computer-controlled general-purpose robot manipulators are becoming increasingly popular, as exemplified by the rather anthropomorphic PUMA device (a robot arm system manufactured by Unimation).

Repetitive automation. CAD, CAM, CAT, and robotics technologies could be combined to produce an almost totally automated factory. The Department of Defense has instituted an ongoing program designed to promote this very concept, called Integrated Computer-Aided Manufacturing or ICAM. The technology now exists to design integrated circuits in one location (CAD), then fabricate the masks for microelectronic manufacture in another (CAM) under the direction of several intercommunicating computers. Further developments and advances in ICAM techniques are imminent.

In a very real sense, an industrialized nation is a symbiotic self-replicating, growing "organism" consisting of humans and machines working together. At the beginning of the industrial revolution the "organism" consisted chiefly of human beings, who, aided by a few machines, performed logical and physical functions. In later years more and more of the heavy and most dangerous work was delegated to machines. As ICAM increasingly enters the mainstream of industrial automation, the logical processes of man-machine manufacturing "organisms" will begin to be taken over by sophisticated computer systems and the physical functions will be dominated by commercial robot devices.

When ICAM techniques are directed toward the production of components of their own systems (CAD, CAM, CAT, and robot machines), a regenerative effect occurs in which each generation of automated factories is cheaper to construct than the preceding one. By the time this regeneration, which has been termed "superautomation" (Albus, 1976), is achieved on Earth, there may be very little human intervention in the replication process except for supervisory and top-level guidance functions. The final step in achieving totally autonomous machine replication requires the replacement of the human top-level managers with computers and turning over any remaining physical tasks to robot devices.

The near-term removal of all human intervention from the industrial "organisms" on Earth is highly unlikely. Certainly people may want to continue to perform various logical and physical functions for social or psychological reasons, and man may always remain the decisionmaker in control of which products are produced. Certain tasks are likely to prove more difficult to automate than expected, and human beings will continue to perform these jobs for economic reasons for a long time to come. Superautomation on Earth will proceed only as far and as fast as is economically advantageous.

The long-term future almost certainly will see the development of full replicative automation capability on Earth. Whether it is economical remains an open question at present. The main advantage of pure machine replicating systems over man-machine symbiotic systems is that autonomous factories can be sent to locations where there is not, or cannot be without great expense, a population of

human workers adequate to operate and maintain the factory complex.

Prime candidates for terrestrial replicating systems applications will most likely be mass-produced products for use in inaccessible or hostile places requiring large spaces to perform the specified tasks. Possibilities include large photovoltaic arrays for centralized power plants in the southwestern regions of the United States (Leonard, in-house document, Bechtel Natl. Inc., San Francisco, Calif., 1980), desert irrigation and soil conditioning equipment covering vast areas, agricultural or military robots, ocean-bottom roving mineral retrievers and seawater extractors patrolling the vast continental shelves, or solar-power satellite ground receiver (*rectennae*) devices. Each of these machine systems could probably be made to self-replicate from a basic feedstock substrate, possibly even from a raw material substrate ultimately.

A few somewhat more speculative terrestrial applications have been proposed by imaginative writers. For instance, Moore (1956) suggested the idea of an artificial living plant able to extract its own nutrients from the sea. These machines could obtain energy from sunlight to refine and purify materials, manufacture them into parts, and then assemble the parts to make duplicates of themselves. Such plants could be harvested for a material they extracted or synthesized. Thus, an artificial plant which used magnesium as its chief structural material could be cannibalized for its metal content. Like lemmings, schools of artificial living machines could be programmed to swim to a harvesting factory when they reached adulthood.

Clearly there would be need for international controls and allocation of areas for production and harvesting. This would involve not only the political rights of nations but also questions of natural conservation. Social problems could arise in connection with the selection of products to be manufactured. An artificial plant might be designed to make a product useless to the plant itself. It might extract gold from seawater, refine it, and cast it into an ingot, which would be harvested as the crop from the plant. But this would be a shortsighted choice. Multiplying at an exponential rate, the gold-making plant would soon produce so much that gold would lose its scarcity value and probably end up being worth very little. An excellent candidate for production by an artificial plant is fresh water, which is needed in great quantities in various parts of the world.

Dyson (1979) suggests a small self-reproducing automaton well adapted to function in terrestrial deserts. It builds itself mainly out of silicon and aluminum which it extracts from ordinary rocks wherever it happens to be. Its source of energy is sunlight, its output electricity and high-tension transmission lines. There is bitter debate in Congress over licensing this machine to proliferate over our Western states. The progeny of one robot can easily produce ten times the present total power output of the United States. Legislation is finally passed authorizing the automaton to

multiply, with the proviso that each machine shall retain a memory of the original landscape at its site, and if for any reason the site is abandoned the device is programmed to restore it to its original appearance.

After its success with the rock-eating automaton in the United States, the company places on the market an industrial development kit, designed for the needs of developing countries. For a small down payment, a country can buy an egg machine which will mature within a few years into a complete system of basic industries together with the associated transportation and communication networks, custom made to suit the specifications of the purchaser. The vendor's guarantee is conditional only on the purchaser's excluding human population from the construction area during the period of growth. After the system is complete, the purchaser is free to interfere with its operation or to modify it as he sees fit. (A technological spinoff is the Urban Renewal Kit -- a city's architects and planners work out a design for urban rebuilding, then the kit is programmed to do the job for a fixed fee.)

Theodore Taylor calls all such devices "Santa Claus Machines" because of their almost "magical" behavior (Calder, 1978). In his version of SRS, a fully automatic mining, refining, and manufacturing facility gathers scoopsfuls of raw lunar materials and then processes them by means of a giant mass spectrograph with huge superconducting magnets. This device converts mined material into an ionized atomic beam which is deflected by the magnetic field. Lighter elements curve more than heavier atomic species, so the material is sorted into stockpiles of constituent elements atom by atom. To manufacture any item, the Santa Claus Machine selects the necessary metals and plastics, then vaporizes and sprays them onto a mold. Instructions for manufacturing, including directions for adapting to new processes and replication, are stored on magnetic tapes in the machine, perhaps activated by radio command from Earth. Conceivably, costs eventually could fall to zero; and if the workload grows too large, the machine simply reproduces itself.

5.4.2 Near-Earth and Lunar Space Applications

While terrestrial self-replicating systems may be limited for some time to coevolution with Earth-based industry constrained by normal economic factors, the prospect for extraterrestrial applications is quite different. The difficulty of surmounting the Earth's gravitational potential makes it more efficient to consider sending information in preference to matter into space whenever possible. Once a small number of self-replicating facilities has been established in space, each able to feed upon nonterrestrial materials, further exports of mass from Earth will dwindle and eventually cease. The replicative feature is unique in its ability to grow, *in situ*, a vastly larger production facility than could reasonably be transported from Earth. Thus, the time

required to organize extraordinarily large amounts of mass in space and to set up and perform various ambitious future missions can be greatly shortened by using a self-replicating factory that expands to the desired manufacturing capacity.

In the not-too-distant future such facilities could be sited either in Earth or lunar orbit, or on the surface of the Moon. The chief advantages of orbital factories are near-zero gravity, absence of lunar dust or atmosphere, convenience in choice of orbit, proximity to Earth (relative ease of transport of finished products), and unobstructed view of virtually the entire celestial sphere. For some applications, however, the lunar surface may be the preferred location. Many manufacturing processes require at least small amounts of gravity, and the availability of solid ground for physical support may be important too. The main advantage to factories on the lunar surface is that the raw materials to be processed into finished products are right at hand -- only relatively low-mass final products need be lifted from the lunar surface, rather than bulky raw materials as in the case of an orbital factory. The Moon can also be used as a shield to block sunlight or electromagnetic interference from Earth during highly sensitive observations.

The useful applications of replicating factories with facilities for manufacturing products other than their own components are virtually limitless.

Manufacturing. Huge solar power satellites with dimensions 1–10 km on a side could be constructed in Earth orbit by a fleet of free-flying assembly robots or teleoperators manufactured by a replicating factory complex using material from the Moon. Components for very large structures, including communications, storage, recreational, penal, or even military platforms could be fabricated, and later assembled, by an SRS. Another exciting mass-production possibility is the notion of orbital habitats, or "space colonies" (O'Neill, 1974, 1976), by which increasingly large populations of human beings could be safely and comfortably maintained in a support capacity for the space program. Additionally, a replicating factory could build more copies of itself, or new variants of itself capable of manifesting different behaviors and producing different outputs, in almost any desired location. Possible useful output of such facilities already has been summarized in section 5.3.4.

Observation. Exceedingly large sensor arrays for Earth or astronomical observations could be rapidly constructed from nonterrestrial materials by a self-replicating manufacturing facility. This technology could be used to make feasible such advanced missions as optical extrasolar planet imaging (using millions of stationkeeping mirror assemblies arranged in an array with an aperture diameter on the order of kilometers); complex multisensor arrays; very large, high-resolution x-ray telescropy; and other self-organizing optical or radio telescopic arrays of grand proportions to

permit such ambitious undertakings as galactic core mapping, continuous observation of large numbers of passive fiducial markers for Earth crustal plate motion monitoring, and various SETI (Search for Extraterrestrial Intelligence) observations including beacon acquisition, radio "eavesdropping," or, ultimately, active communication. Automated mass production will make possible arrays with heretofore unattainable sensitivity and spatial resolution.

Experimentation. Replicative automation technology will permit a tremendous expansion of the concept of a "laboratory" to include the Earth-Moon system and ultimately all of the bodies and fields in the Solar System. A number of grand experiments could be undertaken which would prove too costly if attempted by any other means. For example, an Earth orbital cyclotron could be constructed as a series of thousands of robot-controlled focusing coils and stationkeeping target assemblies within the terrestrial magnetosphere, with operating energies possibly as high as TeV for electrons and GeV for protons. Additional experiments on magnetospheric propulsion and energy generation could be conducted by free-flying robot drones manufactured on and launched *en masse* from the lunar surface. Gravity field probes, including mascon mappers and drag-free satellites, could be coordinated to perform complex experiments in kinematics, special and general relativity, and celestial mechanics. Investigations of artificial *in situ* lunar crater formation dynamics, solar wind composition and utilization, unmanned ecological simulation modules, and isolation or "hot lab" module manufacturing for conducting dangerous experiments with explosive, radioactive, or biologically engineered materials are still further possibilities.

Exploration. The Moon is largely unexplored. A growing, self-replicating factory could be reprogrammed to mass-produce modified mining or other mobile robots, including orbiters and rovers, for detailed investigation of the lunar surface. This would augment orbital sensing and intelligent image processing systems (see chap. 2) around the Moon, and could be linked to lunar subsurface explorers and other automated surface prospecting equipment to assist in new resource location, colony siting, and the further acquisition of scientific knowledge. Subselene or subterrene (see discussion of the "Coal Mole" in Heer, unpublished draft notes, Pajaro Dunes Workshop, 1980) mining robots could burrow deep into the lunar or terrestrial crust in search of pockets or veins of useful substances, and then dig them out. A self-replicating manufacturing facility could produce thousands of meter-long robot rovers equipped with cameras, core samplers, and other instrumentation which could survey the entire Moon — or any other planet, for that matter — in just a few years. Such exploration would take a century by more conventional methods. Similarly, due to the low gravity, lack of atmosphere, and relative abundance

of energy and raw materials, the Moon is an excellent location for the construction and launching of future generations of interplanetary exploratory spacecraft.

Human resources. The augmentation of human services and the extension and safety of the human habitat is yet another near-term application of self-replicating systems. In principle, it is possible to construct a completely autonomous lunar-based facility, but it may turn out to be inefficient or uneconomical in the future unless a few human beings are present onsite to handle unforeseen problems with the machinery. (Humans are the most compact and efficient general-purpose self-replicating systems of which we have certain knowledge.) Initial crew quarters and supplies can be transported from Earth, but much larger and more pleasant living accommodations could be manufactured *in situ* by lunar or orbital replicating systems. The inexpensive mass-production of habitation and agricultural modules (or their components) could help open the door to more extensive lunar and space colonization by people, including recreational, industrial, medical, and educational uses, especially because of the abundant solar energy and the expected ability of replicating factories to manufacture and implement a low-cost lunar-surface-to-orbit launch capability. A comprehensive, highly sophisticated automated astronaut search and rescue system may also become necessary as the human population in space begins to grow, with system components mass-produced by SRS.

Presently, there are about 6000 known and tracked pieces of debris orbiting the Earth at various altitudes and inclinations, and countless additional shards which lie below observational thresholds in near-Earth space. These represent an ever-increasing danger of collision with spacecraft. Debris-catchers or "scavengers" mass-produced by SRS technology could be automatically launched into various Earth orbits, seek out and recognize space debris, report ephemerides in the case of satellite-like objects to avoid destruction of operational equipment and, upon go-ahead, collect the debris. Scavengers would be programmed either to enter the Earth's atmosphere after a specified time in orbit and self-destruct, or to return their collections to orbital manufacturing facilities for recycling of high-level components and materials to help build new robots. A more advanced network could offer protection from possible ecological disasters caused by terrestrial meteorite impacts (Alvarez et al., 1980).

Another possibility, however controversial, is meteorological and climatological intervention on both a local and global scale. A number of interesting alternatives were discussed by the participants of the recent Pajaro Dunes Workshop (Heer, unpublished draft notes, 1980), including:

- Manufacture of 10^7 copies of a 1-km² sunshade to achieve global cooling, if required, which could be

deployed most effectively for the polar regions at Earth-Sun L1 (losses due to image diffusion) or in LEO (serious orbital problems).

- Deployment of 1 to 10 million copies of 1-km² mirrors in LEO, to cause localized heating effects by concentrating incident solar radiation.
- A system of several 1 to 10 GW microwave frequency solar power satellites to add 100 to 200 W/m² to selected terrestrial ground spots 10 km diam, to be deployed in geosynchronous Earth orbit (GEO).

The replicative manufacturing facility needed to economically produce such large numbers of similar system elements would make possible at least a rudimentary global homeostatic environmental control by humanity.

Given the exotic conditions prevailing on the lunar surface and in space, and the novel materials and processes that may become available, it is highly probable that a self-replicating growing lunar facility will be able to economically produce many goods directly for use in space and for export to Earth. What these goods might be is not now certain. However, the economic importance of the telephone, steamboat, airplane, television, office copying machine, etc., during their early stages of development likewise were not at all obvious to most people.

5.4.3 Solar System Applications

The technology of replicating systems will become increasingly important as humanity expands its theater of operations from near-Earth space out to encompass the entire Solar System. Mankind has fallen heir to an incredible treasure trove of nonterrestrial energy and material resources (see sec. 4.2.1). It is likely that replicating machines will provide the only “lever” large enough to explore, and ultimately manipulate and utilize in a responsible fashion, such tremendous quantities of organizable matter. Lacking this advanced automation capability, most of the more ambitious Solar System applications appear uneconomical at best, fanciful at worst.

Observation. Exceedingly far-reaching observational possibilities may become feasible with the advent of SRS technology. Very large baseline interferometry (VLBI) may be attempted with components distributed across the entire Solar System, perhaps located at the stable Trojan points of the Jovian planets or their moons, providing multiplanar baselines of from 1 to 100 AU and complete spherical coverage with the use of out-of-ecliptic robot sensor devices that are mass-manufactured by replicating factories. The solar wind could also be mapped in three dimensions, and by using the entire Sun as a gravitational lens focal lengths on the order of the size of the Solar System can in theory be obtained (Ingel, 1974). This may permit simultaneous observation of the entire celestial

sphere across the full spectrum of gravitational radiation using fleets of gravity-wave detectors manufactured by SRS and stationed along the focal plane. A Solar System surveillance network could be constructed to track and warn of objects approaching human habitats, facilities, or the Earth on collision courses, allowing mankind to avoid potentially severe catastrophes.

Exploration. The technologies developed for a generalized lunar autonomous replicative manufacturing facility should be directly applicable in the exploration of all planetary and satellite surfaces. One early possibility is a mission to land a single replicative “seed” on Mars which would then use local materials to produce large numbers of rovers (including, perhaps, fliers, crawlers, walkers, or rollers) and orbiters. A population of 1000 to 10,000 surface rovers each perhaps 100 kg in mass, coupled with a chain of orbital monitors, might continuously monitor and explore the planetary surface and leave stationary probes (active or passive) behind in permanent emplacements. The probes need only have lifetimes on the order of a year or so, since they could constantly be repaired and replenished by the rovers (each of which could last 10 years or more). This system would provide complete surface exploration and continuous status monitoring of all areas on the planet, including temperatures, pressures, wind velocities, seismic events and crustal creeps, meteorite impacts, surface and subsurface compositions, illumination, precipitation, and numerous other phenomena of interest. Automated balloon explorers could be mass-produced and released in Jovian atmospheres, and “trains” of deep solar probes (Heer, unpublished draft notes, 1980) could be hurled into the Sun to obtain direct information on internal conditions there.

Materials retrieval. Replicating systems would make possible very large-scale interplanetary mining and resource retrieval ventures. Nonterrestrial materials could be discovered, mapped, and mined using teams of surface and subsurface prospector robots manufactured *en masse* in an SRS factory complex. Raw materials could be dug up and sent back to wherever they were needed in the Solar System, or could be refined along the way and the waste slag used as reaction mass, or could be utilized *in situ* for manufacturing useful products which would then be exported. Atmospheric mining stations could be established on many different planets – Jupiter and Saturn for hydrogen, helium (and rare isotopes potentially useful for fusion power generation, Martin, 1978), and hydrocarbons, using “aerostats” (Parkinson, 1978); Venus for carbon extraction; Europa for water; Titan for hydrocarbons; etc. Comets could be intercepted to obtain large quantities of useful volatiles, and Saturn’s rings could be mined for water-ice by large fleets of mass-produced robot craft. Heavy metals may be

retrieved in great quantities from asteroids. Replicating systems might manufacture huge mining, processing, even ground-to-orbit and interplanetary transportation capabilities using local materials in surprisingly short periods of time.

The general product factory. The team has proposed the design and construction of an automatic multiproduct replicating lunar factory. The reason for the factory having multiproduct capability is to permit it to be able to respond to any changing requirements in kind or amount of product output. This leads to a still broader concept — the notion of a general product factory.

A general product factory is one which can be instructed to manufacture anything which is physically possible to make. Such a system is the physical realization of von Neumann's "universal constructor" automaton, which can construct anything constructable, given an adequate substrate and the rules of operation of his artificial cell-space universe. In the context of drawing upon planetary resources, we should think of each celestial body in terms of its menu of possible materials and the repertoire of processes theoretically available there (see sec. 4.5.4). The following questions should then be considered:

- What is the total range of things which can be made using these processes acting upon these material resources? (See sec. 5.3.6.) This query should be viewed in the broadest possible fashion, including biological as well as mechanical entities.
- Does there exist, for this planetary environment, a factory design which is capable of making all of these entities?
- Can an initial system be designed which, when introduced into the target environment, will yield such a general product factory? A few important developmental milestones are suggested in table 5.4.

The notion of a general product factory using asteroidal material was briefly considered at the Pajaro Dunes Workshop. The "Hive," as it was called, would consist of "an autonomous space island 'beehive' of independently intelligent machines . . . specialized in mining and production, experts in planning, navigation and repair." The product of the Hive would be solar power satellites, "asteroids turned into space colonies, vacuum-filled balloons of nickel floated down to a resource-hungry Earth, spaceships, telescopes, or even another Hive." The Hive was envisioned as an independent economy, using raw materials gathered from the Asteroid Belt, refined and processed with solar or fusion energy, then fashioned into useful output by robot hands. Workshop participants suggested a timetable in which the first fully autonomous replicating system could be in operation in the Asteroid Belt by 2040, commencing expo-

TABLE 5.4. DEVELOPMENTAL MILESTONES FOR A GENERAL PRODUCT FACTORY

1. Design and construct a system which, when supplied only with parts and subassemblies, can duplicate itself.
2. Design and construct a system which can duplicate itself, and in addition produce some useful product.
3. Design and construct a system which, when supplied only with feedstock, can duplicate itself.
4. Design and construct a system which, when supplied with raw materials only, can duplicate itself.
5. Design and construct an automated, reprogrammable, multiproduct system which can, from raw materials, duplicate itself.
6. Design and construct an automated, reprogrammable, multiproduct system which, using only lunar materials and employing only those processes possible in the lunar environment, can duplicate itself.
7. Design and construct an initial automatic "seed" system which, if placed on the lunar surface, could unpack itself and develop into an automated, reprogrammable, multiproduct replicating system, using lunar resources and lunar processing modes only.
8. Design and construct an initial seed which can, in the lunar environment, develop and augment itself so as to become a *general*-product factory, relative to the lunar environment.
9. Design and construct a seed which can, in an *arbitrary* planetary environment, develop into a general-product factory.

nential growth with a replication time of 5 years, resulting in a total of 1000 new Hives available for production by the year 2080.

Human resources. From the human standpoint, perhaps the most exciting consequence of self-replicating systems is that they provide a means for organizing potentially infinite quantities of matter. This mass could be so organized as to produce an ever-widening habitat for man throughout the Solar System. Self-replicating homes, O'Neill-style space colonies, or great domed cities on the surfaces of other worlds would allow a niche diversification of such grand proportions as never before experienced by the human species.

SRS provides such a large amplification of matter-manipulating capability that it is possible even to consider the "terraforming" of the Moon, Mars, Venus, and other worlds. Terraforming is a theoretical concept in which a planetary environment with otherwise inhospitable conditions for life is purposefully and artificially altered so that

humans may live there with little or no life support equipment. The "traditional" approach is to suggest biological means, such as the proposal to seed the atmosphere of Venus with genetically tailored algae to convert excess carbon dioxide into combined carbon and free oxygen. This would have the incidental salutary effect of lowering the planetary surface temperature so that people could live unaided on the surface. However, it is not known whether biological organisms can be found or developed which are able to withstand present conditions in the Venusian atmosphere.

An alternative approach is to use nonbiological replicating systems which may be far more durable under extreme conditions. A few simple calculations reveal the approximate magnitude and duration of such an enterprise. Consider the terraforming of Mars. For simplicity it is assumed that the planetary crust is largely silicon dioxide and that a general-purpose 100-ton SRS factory "seed" which lands there can replicate itself in 1 year. In just 36 years such a system could theoretically manufacture an SiO_2 reduction capability able to release 220,000 tons/sec of pure oxygen into the Martian atmosphere, which in only 60 years is sufficient to produce 4×10^{17} kg O_2 . Assuming negligible leakage through the Martian exosphere, this is enough oxygen to establish a 0.1 bar breathable oxygen atmosphere planet-wide – approximately equivalent to normal air on Earth at an altitude of 3000 m (16,000 ft). This plan requires a solar power satellite system in near-Mars orbit with a total generating capacity of about 10^{17} W, a network which would take less than a year for the finished replicating factory system to produce. The total material thus excavated to terraform Mars is of the order of 10^{18} kg SiO_2 , enough to fill a surface depression 1 km deep and 600 km diam. This is roughly the size of the crater Edom near the Martian equator, or Mare Crisium on the Moon.

Of course, far more efficient methods for terraforming planets may eventually be found, such as Dyson's proposal to mine the Saturnian moon Enceladus for its water-ice and return the material to Mars (Dyson, 1979). But the utility of self-replicating systems is clear, and it appears that terraforming times on the order of one century are conceivable using the SRS approach.

Technology requirements. Additional technology over and above "superautomation" (sec. 5.4.1) will be required for the highly ambitious ventures described in this section using advanced space-based self-replicating systems. The most important new technology in this regard is "closure engineering," discussed in section 5.3.6. Some of the enterprises proposed above are of such large scale that it is difficult to envision a feasible mode of operation with anything less than 100% materials and energy closure and virtually 100% information closure as well. No doubt there exist manufacturing operations which are not economically

viable candidates for total automation in terrestrial industry – in these instances the functions either must be redesigned for full automation or else people must be permanently incorporated as symbionts of a locally teleoperated or remotely human-supervised system. Manufacturing processes developed for terrestrial environments must be re-engineered to accommodate the input and production environments found in space or on the surfaces of other planets, and output streams must be sufficiently flexible to make feasible the notion of a general products factory.

5.4.4 *Interstellar and Galactic Applications*

Replicating systems technology is the key to exploration and human habitat expansion beyond the confines of the Solar System. Although these kinds of missions necessarily are highly speculative, and admittedly exceed the limits of current or projected technology in many areas, a consideration of possible interstellar and galactic applications is nonetheless a useful exercise because it serves to illustrate the fantastic power and virtually limitless potential of the SRS concept.

Extrasolar exploration. Before humankind can move out into interstellar space, automated probes will scout the way ahead. The distances are so large and the volumes so vast that self-replicating probes are highly desirable, even essential, to adequately and efficiently perform a reconnaissance of extrasolar star systems in a search for human habitable worlds and extraterrestrial life. A preliminary design for a self-reproducing interstellar probe has been presented in the scientific literature (Freitas, 1980a), and another study of the comparative benefits of reproducing and nonreproducing galactic exploration strategies by unmanned probes suggests that search patterns using semi-intelligent automata involving more than about the nearest 100 stars would probably be optimized (in terms of economy and productivity) if self-replicating systems are employed (Valdes and Freitas, 1980). Reproductive probes could permit the direct investigation of the nearest million stars in about 10,000 years and the entire Milky Way galaxy in less than 10^6 years, starting with a total investment by humanity of a single self-replicating exploratory spacecraft.

The problems in keeping track of, controlling, and assimilating data returned by an exponentially growing number of self-reproducing space probes are staggering. Part of the solution may lie in the use of an extremely high level of autonomy in operations management and reasoning such as discussed in chapter 3 of this report; part may lie in the utilization of high levels of abstraction in the information returned to Earth after the fashion of the World Model sensing and data-processing philosophy articulated in chapter 2. Another major piece of the solution is the development of a hierarchical command, control, and information-gathering architecture in which any given

probe communicates directly only with its own parent and offspring. Control messages and exploration reports would pass up and down the chain of ancestral repeater stations erected by earlier generations (Valdes and Freitas, 1980). Certain highly critical but low probability signals might perhaps be broadcast in an omnidirectional alarm mode to all members of the expanding network (and to Earth) by individual probes which encountered specific phenomena or events – such as the discovery of an extrasolar planet suitable for human habitation or a confrontation with intelligent alien lifeforms or their artifacts.

Extrasolar utilization. Before mankind can venture out among the stars, his artifacts and replicating machines must blaze the trail. Ultimately, however, one can envision free-flying space colonies journeying through interstellar space (Matloff, 1976). Upon reaching some new solar system or other convenient source of raw materials, these mobile habitats would reproduce themselves with the human passengers redistributed among the offspring colonies. The original space habitats would serve as extraterrestrial refuges for humanity and for other terrestrial lifeforms that man might choose to bring along. This dispersal of mankind to many spatially separated ecosystems would ensure that no planetary-scale disaster, and, as people travel to other stars, no stellar-scale or (ultimately) galactic-scale event, could threaten the destruction of all mankind and his accomplishments. Replicating systems may be the only rational means to attempt large-scale astroengineering projects usually relegated to the domain of science fiction, such as the construction of “Dyson Spheres” which enclose and utilize the energy output of entire suns (Dyson, 1959).

The limits of expansion. The expansion of a population of replicating systems in any environment is restricted largely by two factors: (1) replication time, and (2) maximum velocity of the outer “envelope” which defines the physical extent or dispersion of the population. No population can accrue at a faster rate than its components can reproduce themselves. Similarly, no population can disperse faster than its medium will permit, no matter how fast components are manufactured – assuming number density remains essentially constant, corresponding to continuous maximum utilization of the environment. Neither factor may be ignored during any phase of population growth.

If envelope expansion velocity does not constrain a population because components are produced only relatively very slowly, then that population will experience exponential multiplication according to:

$$N(T) = \exp(T/t) \quad (1)$$

where $N(T)$ is the number of replicating units comprising the population at time T (replication starts at $T = 0$) and t is the replication time per unit, assumed constant. On the

other hand, if unit reproduction is so swift that multiplication is not constrained by replication time, then the population can grow only as fast as it can physically disperse – that is, as fast as the expansion velocity of the surface of its spherical outer envelope – according to:

$$N(T) = 4/3 \pi d(VT)^3 \quad (2)$$

where V is peak dispersion velocity for individual replicating units at the periphery and d is the number density of useful sites for reproduction. Expansion cannot exceed the values for $N(T)$ given either by equations (1) or (2) at any time T , provided all replication sites receive maximum utilization as stipulated (e.g., constant number density of units).

Populations of machines expanding across the surfaces of worlds with replication times on the order of 1 year will not achieve mean envelope growth speeds in excess of a few meters per hour, even in later phases of extreme enlargement when the population of SRS covers a large fraction of the available planetary surface. This figure is well within anticipated nominal ground transport capabilities, so exponential extension should remain largely velocity-unconstrained on such bodies if replication time remains constant at greater population sizes.

Similarly, three-dimensional populations of replicating systems in interplanetary space using Solar System materials and solar energy ultimately are restricted to spherical circumstellar shells where SRS units can collect virtually all energy radiated by the Sun. If a “Dyson Sphere” of 100-ton replicating “seed” units is assembled near the orbit of Earth, approximately one terrestrial mass is required to manufacture the more than 10^{19} individual units needed to completely enclose the star. But maximum expansion velocity even in this case never exceeds about 100 m/sec, hence interplanetary replicating systems as well in theory may spread at purely exponential rates.

In the interstellar realm, however, the situation is far more complex. Depending on the maximum dispersal velocity and interstellar probe replication time, either equation (1) or (2) may control. Figure 5.24 compares pure exponentiation and dispersal speed effects for $t = 1$ year (see sec. 5.3.4) and $t = 500$ years (Freitas, 1980a), and for $V = c$ (since the theoretical maximum envelope expansion rate is the speed of light) and $V = 10\% c$ (Martin, 1978) for an assumed homogeneous stellar distribution of “habitable” star systems (taken as 10% of the total) in the galactic disk. In most cases, exponential multiplication soon is halted by the speed-of-light barrier to dispersion, after which the SRS population expansion proceeds only polynomially.

Technology requirements. In order to sustain the expansion of a potentially infinite replicating system, new dispersal mechanisms must be developed. Initially, self-replicating machines or their “seeds” must be capable of motion across

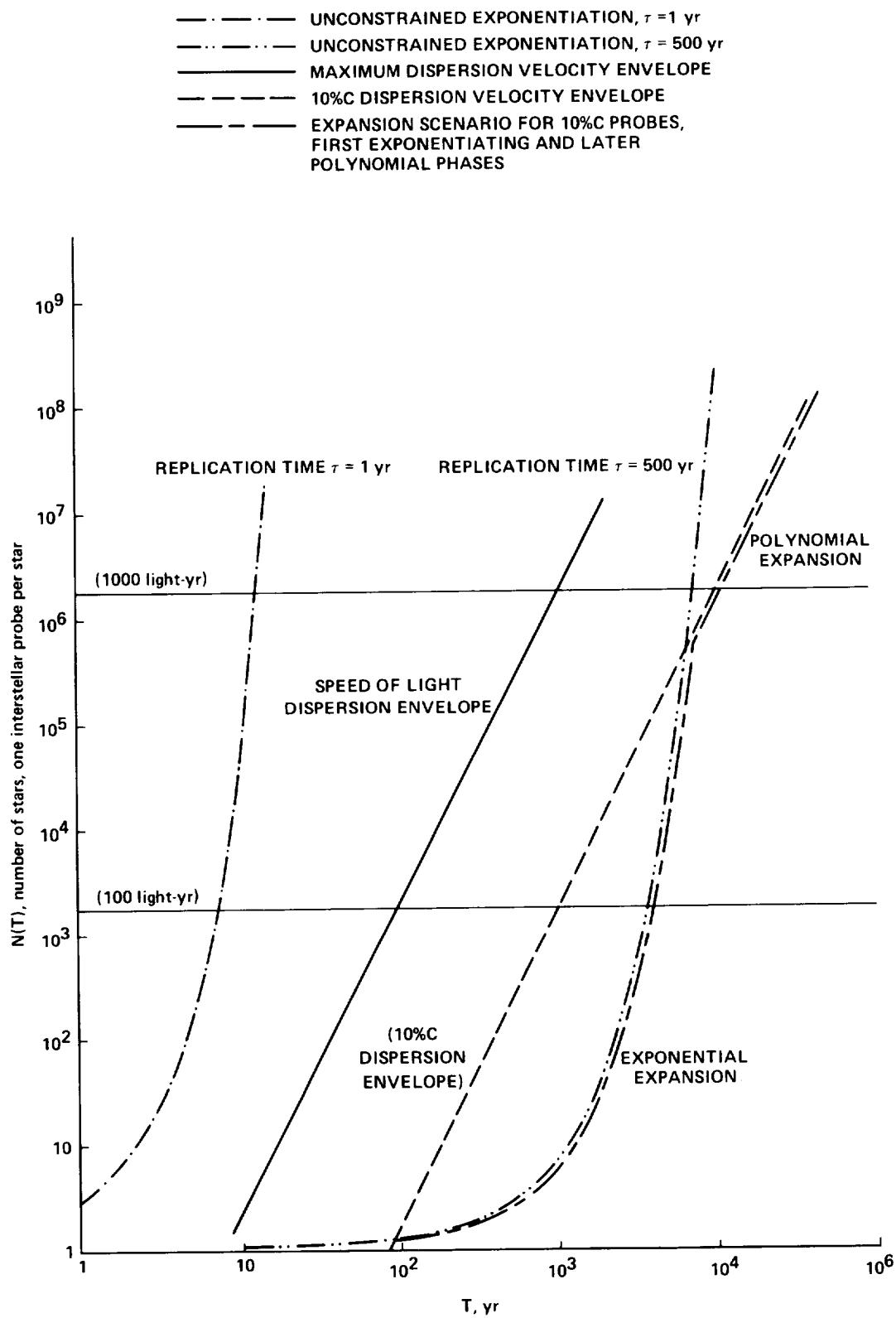


Figure 5.24.— Limits to exponential and polynomial expansion of self-replicating interstellar probe populations dispersing throughout the galactic disk.

a planetary surface or through its atmosphere or seas. Later, interplanetary, interstellar, and, ultimately, intergalactic dispersal mechanisms must be devised. Supplies of energy, stored and generated, must be established if extrasolar spacecraft are to survive in the depths of interstellar space far from convenient sources of power (such as stars) for a major portion of their lives. The technologies of command, control, and communication over stellar and galactic distances ultimately also must be developed.

5.4.5 Applications to Basic Research

In addition to specific applications of replicating systems technology to future missions in space, a number of applications to basic research in biology, computer science, and automata theory have been identified by the team. These are discussed below.

Automaton theory. Automaton theory is the abstract and precise study of all mechanistic devices and processes. At times this has been restricted to the theory of discrete and deterministic machines with a fixed finite number of states. In this narrow sense it is the abstract mathematical counterpart of physical devices such as existing digital computers. In the broadest sense, though, automaton theory can include the study of all mechanisms, discrete or continuous, deterministic or probabilistic or even indeterministic, with a fixed, variable, or indefinitely large number of possible states. Included in this wider definition is the notion of devices which can alter the number of their states by growth or by contraction in respect to certain of their organs, much like the way a Turing machine or a push-down automaton (or a linear-bounded automaton) can increase or decrease the number of its states by increasing or decreasing the length of its memory tape — but also can grow by increasing or decreasing the numbers of its more active computing components. This is representative of machines which can construct or dismantle other machines.

These machines can not only increase their memory capacity but can augment their computing power by the construction of additional active computing organs (registers, control units, etc.) and by constructing machines separate from themselves, including duplicates of themselves. Von Neumann had begun to develop a general and logical theory of automata which would have embraced all these machine types. Automaton theory has, however, never achieved the generality he sought, at least not in the sense he seems to have intended.

The very general theory of automata has become increasingly abstract, moving from describing mechanistic processes in terms of algebraic concepts such as groups and semigroups to employing category theory, the most abstract and general of algebraic theories. Although a certain level of understanding of what mechanisms might exist has thereby been developed, the applicability of such

approaches to the design of complex systems of automata is very slight or nonexistent. In this regard, von Neumann once lamented that "...at a great distance from its empirical source, or after much abstract inbreeding, a mathematical subject is in danger of degeneration Whenever this stage is reached, the only remedy seems to me to be the rejuvenating return to the source — the reinjection of more or less directly empirical ideas." (von Neumann, 1966).

It may be that an effort to actually design and implement a system of machines which can construct more machines like themselves would encourage theorists again to attempt to develop a very general automaton theory, including as a part of its subject matter the spatial and communicatory interactions of vast and increasing numbers of submachines. (Perhaps the automatic telephone system provides us with the closest physical analogy to such systems, aside from the analogy of human societies themselves.)

Such a theory would enable one to ask what is the best organization of a system of (potentially) arbitrarily increasing numbers of active components, arranged in various spatial geometries. How might the interacting activites of vast numbers of submachines be optimized? What rules of interaction and of interconnection can be imposed on such a system in order to attain efficient and stable behaviors? What are the safest physical and behavioral interactions, and which lead to instabilities and pathologies?

A general theory would also take as part of its subject matter the flow of parts and materials. It might, like the von Neumann cellular system, treat the creation and flow of materials and the movement of machinery as a form of information flow. It might distinguish information, materials (raw materials, feedstock, and parts) and the movement and siting of machines, but treat them in an identical format so that tradeoffs and exchanges in these categories could be computed (while retaining the essential differences among these types of flow important to the working of the system). The theory would answer such questions as: When will more information be the best substitute for more parts or more feedstock? Under what conditions in the vast assemblage of machines should parts be made anew, from raw materials and feedstock, and when should information or already finished parts be employed to the same purpose? When should machines which are likely to fail be abandoned? When should machines in the assemblage which are still in good condition nevertheless be shut down, moved, sacrificed for parts or dismantled, or sealed off? Under what local and global conditions should submachines be retired, repaired, or replaced?

Theoretical biology. Machines which can construct machines, and machines which can construct replicas of themselves, display behavior which in many ways is analogous to that of natural organisms. Furthermore, as machines are designed to examine their own structure and

the structures of other machines, to repair themselves and other machines, and generally to become more autonomous and more reliable, the analogies become even more apparent.

The ways in which machines carry out these processes of growing, repairing, regenerating, and reproducing may or may not be similar to those carried out by natural organisms – which, in many cases, are not yet even known.

One goal of theoretical biology is to develop an understanding of the mechanisms of living systems, to the point where these systems can be characterized in a precise mathematical fashion (Miller, 1978). To attain such a characterization one needs a good intuitive feeling for the full possible range of lifelike forms. For example, a theory of biology that takes as its subject matter only Earth-evolved forms would be as unlikely to be capable of providing adequate explanation for non-Earth forms as were attempts to account for the forms of extant organisms quite apart from their extinct progenitors.

It seems, therefore, likely that an adequate explanatory theory of biology of any elegance and simplicity must embrace not only all biological forms which presently exist, but all those which have ever existed, or will exist, or could exist. Indeed, the proper subject matter for a true theoretical biology in its broadest sense would be the study of lifelike behavior wherever it occurs whether now, or in the past, or the future; whether on Earth or elsewhere in the universe and whether it is exemplified in artificial or natural forms (Freitas, 1980b), a field of study termed “xenobiology” by one author (Freitas, 1981). This suggests that research on complex automata able to reconstruct, reproduce, and repair themselves might serve as a fertile source of hypotheses as to the logical control and organizational aspects of how living organisms in fact carry out these processes. Such explanatory hypotheses can apply to lifelike systems generally and have the advantage that they are likely to be simpler and more elegant than the necessarily *ad hoc* explanations of behavior for the particular organisms of particular worlds, at particular times.

Thus, research in self-growing and self-replicating machine systems can be viewed as a contribution to, even as a central part of, a true theoretical biology which takes as its subject matter not merely the evolved, naturally occurring living organisms of Earth, but lifelike mechanisms, natural or artificial, having existed or possible, wherever in the universe they might arise.

Design of biological and hybrid organisms. The forms and processes of artificial organism-like systems are not bound to follow the particular structure and logical organizations of known naturally evolved organisms. As the design of increasingly complex artificial systems capable of drawing materials and energy from natural surroundings and possessing more and more organism-like properties proceeds, it may become apparent that there are artificial

organism functions which, if embodied in biological organisms, would be of value. With advances in “genetic engineering” it may become possible to create new biological forms, possessing the desired features.

Just as the design of artificial mechanisms can be inspired by contemplation of evolution’s apparent solutions to various design problems, so might new *biological* systems also be created, drawing upon designs originally conceived for artificial systems – a kind of inverse bionics. Taking this a step further, one can envisage as a research goal the gradual elimination of the perhaps arbitrary line now drawn between artificial and natural organisms, and the consideration of a more deliberate systematic investigation of the creation of hybrid biological-mechanical systems.

Experimental evolution. Studies of form and function in biological and artificial systems may contribute to an understanding of the design and construction of both biological and mechanical organisms. This interdisciplinary exchange should not be limited to studies of the relationship between individual classes of lifelike entities, but should also extend to studies of the consequences of large numbers of such entities interacting and competing for resources. Replications of programs and creation of new machines (including replicas), and compounds and combinations of initially existing machines, can be a feature of the proposed machine replicating systems. It seems clear that development of a science of evolving systems is needed (Miller, 1978). (This would again be a part of a very general “true” theoretical biology, which takes all possible lifelike systems as its subject matter.)

For example, one putative value of sexual over asexual reproduction is the enormously increased mobility of genetic variation in the species population. This widely available variation tends to ensure that environmental changes can be accommodated or exploited with great swiftness by at least some members of the population (Smith, 1978). In a “designed” universe, one is free to consider the advantages (if any) of three or more sexes (Freitas, 1980c; Smith, 1978) or of the consequences of other, even more radical redesignings of existing natural systems. In particular, the actual behavior of largely autonomous growing replicating machine “species” with differing capabilities and reproducing strategies certainly should be an object of study by evolutionary biologists who might be able to predict the forms which would persist and come to dominate in systems left unperturbed by external pressures and commands.

The existence of large interacting populations of entities whose “genetics” are precisely known, but whose global behavior over time cannot readily be predicted, may be of great experimental value to evolutionary biologists. At present, computer simulation is the usual tool of choice for such problems. However, if the physical creation of machine populations becomes sufficiently inexpensive,

experimental situations might be created in remote nonterrestrial regions. Machine growth and population changes could be monitored over time for their adherence (or not) to hypothesized consequences. The advantage of this approach over the computer simulation would be in the much greater detail and fidelity to real situations, and the consequent likelihood of serendipitous useful observation.

Machine intelligence architecture. Very general symbol manipulating devices (such as stored program computers) are at the heart of efforts to demonstrate that machines can exhibit behavior which in humans or other animals would be considered intelligent. In one sense, such devices are computationally universal. That is, certain mathematical technicalities aside, they can carry out any arbitrary Turing machine computation and, accepting the Church-Turing Thesis, can also carry out any algorithmic process. Thus, if *any* machine can be intelligent one need look no further than to a general-purpose computer, for there is some program which will cause the machine to display the desired intelligent behavior. This is so even if one insists that brains, for example, are machines, but are not at all like digital computers. This is because digital computers, again accepting the Church-Turing Thesis, can be programmed to simulate any known mechanistic process to any fineness of detail, whether the process of interest be analog, frequency coded, probabilistic or other.

Even though ordinary computing machines do not, for example, reproduce themselves, they can be programmed to simulate the behavior of machines which do in fact reproduce. From this point of view, the concept of machines which possess the power to construct other machines and to replicate themselves can be represented to any degree of detail in the computation of an ordinary general-purpose computing machine which cannot itself reproduce. Even though existing general-purpose machines cannot generally inspect themselves and draw conclusions therefrom, computers can be programmed to simulate such unlikely machine actions if such a simulation is thought useful or interesting. Hence, the construction of the kinds of machines considered here – machines that can compute, construct, reproduce, and inspect, repair, simulate, and observe both themselves and other machines – would not enlarge what a general-purpose device can in theory already do but rather our perception of their capability to exhibit more sophisticated mindlike behavior.

It should also be noted that machines can be designed and constructed so as to do things beyond what any known evolved organism (including man) can do. We are already aware of this superiority of machines in regard to strength, speed, accuracy, flight, and the like. There are already many ways in which machines can be designed and constructed so as to exceed human mental capabilities for specific tasks.

For example, though we are constantly reminded of the social value of being able “to see ourselves as others see us,” our evolutionary history has left us with only a very limited capacity for accurate introspection and self-examination – though in this respect we admittedly exceed all other known evolved creatures. Machines, however, can be designed to secure far greater access to their internal structure and states than we are ever likely to possess as individuals, and this capacity might mean that machines can be programmed to achieve mindlike powers far beyond ours. A trivial case of this “introspective” superiority of machines is seen in their ability to “remember.” Computers can be programmed to methodically search all of their memory with a thoroughness that can evoke human envy.

5.5 Implications

It appears that self-replicating systems may have numerous economical applications on Earth, in near-Earth and lunar space, throughout the Solar System, and perhaps even in the interstellar realm, for future exploration and utilization. The main advantage of SRS is their tremendous capability for performing any desired task at almost any remote location, starting with a relatively small investment of time, money, energy, and mass. This suggests that replication technology may have significant social and economic impacts on American and human society, as discussed below. A number of philosophical and ethical implications may derive from replicating systems techniques. Various issues regarding the future of human and machine evolution must be addressed, together with the “cosmological” implications of SRS.

As the time allotted to consideration of the implications of machine replication was relatively small, the team was not able to examine many intriguing questions in depth. In many cases, it was possible only to frame questions regarding general classes of social and cultural impacts, as no satisfactory answers were immediately apparent. Consequently, this section must be regarded simply as a blueprint for further study in this area.

5.5.1 Socio-Economic Implications of Self-Replicating Systems

The history of technology on this planet is the record of man’s constant attempts to control his environment through the use of extrasomatic tools. The development of SRS in this context will be revolutionary, with impacts equal to or exceeding those engendered by other “revolutions” in human history. For the first time, mankind will be creating, not merely a useful paradigmatic tool (e.g., the scientific method, Copernican revolution), organizational tool (e.g., centralized cultivation, agricultural revolution), or energy-harnessing tool (e.g., steam power, industrial revolution), but rather a wholly new category of “tool” - a

device able to *use itself* intelligently and with minimum human intervention. In many respects, with SRS mankind is creating a technological partner rather than a mere technical implement.

Superautomation on Earth and in space. The use of self-replicating systems on Earth poses many problems. A compact, freely replicating system released on the surface of the Earth potentially could compete with humans for resources such as living space and energy. It could also smother us in its waste products. Even if kept under control, a terrestrial SRS could wreak economic havoc by manufacturing products for which the consumers who will use the products will not have to pay. Unfortunately, we will probably have to deal with this problem regardless of whether replicating systems technology *per se* is ever developed. If industrial automation continues in the direction it seems to be headed now, global commerce soon will reach a state of "superautomation" (Albus, 1976) in which an entire national industrial base has become automated and is, for all practical purposes, a terrestrial SRS. Such a system may function without the need for significant inputs of human labor. Eventually it should be possible to deal with the attendant economic dislocations, but the transition is certain to be excruciatingly painful.

In Earth orbit and on the lunar surface, however, the situation is quite different. In the environment of space SRS would not be in competition with an established human presence. Instead, they would provide a powerful "tool" by which humans can manipulate that environment to their advantage. One can envision building vast antenna arrays (for radio astronomy and SETI), solar power satellites, or even lunar, orbital, or free-flying habitations. These applications should enhance, rather than destroy, the economic fabric of terrestrial civilization, just as colonies in the New World enhanced the economies of their parent nations. By expanding into space, mankind has the potential to gain, rather than lose, from extensive automation. Instead of doing the same amount of "work" that is required to sustain terrestrial existence (and doing it with fewer and fewer people), by moving into space even more people can be kept occupied than before while at the same time extending into a redundant habitat. This seems perhaps the best way to sustain the least trauma in the years ahead.

The development of the necessary artificial intelligence, robotics, and automation techniques will likely have enormous short range impacts on Earthbound activities. If our economy is to be transformed by such revolutionary technologies in a fairly short period of time, how can the United States (and the entire industrialized global community) prepare for and avoid or mitigate potentially vast dislocations? Will we need a new academic discipline of "revolution management"?

Economics of replicating systems. Whether supported by public or private sources, the development of SRS must make good economic sense or else it will never be attempted. Self-replicating factories on Earth or in space may appear theoretically capable of creating bountiful wealth and endless supplies of goods and services for all (Bekey and Naugle, 1980; Heer, 1979). However, this utopian ideal must be tempered with the cold logic of cost-benefit analyses and indices of profitability if it is to gain some measure of credibility in the business world.

Let us assume that a financial consortium invests a sufficient quantity of capital to research, design, build, and successfully deploy the first SRS. This consortium may represent an association of private businesses (e.g., the Alaskan Pipeline), an intergovernmental entity (e.g., the International Monetary Fund), or individual public agencies (e.g., NASA). Deployment may occur on Earth, in orbit, on the Moon, or even on the surfaces of other planets or the asteroids. After a relatively brief period (T years) of growth, the capacity of the initial SRS expands a thousandfold by self-replication, and commercial production begins.

Assume that the original investment is $\$X$ and the original factory could produce useful manufacturing output with an annual value of $\alpha\$X$. After the SRS undergoes thousandfold expansion, its output is worth $1000\alpha\$X$ per year (provided demand remained unaffected). The value of the original investment after T years is $\$X(1 + I)^T$, where I is the mean annual inflation rate during the period of investment. Thus, to repay the original investment and achieve economic breakeven will require approximately $\$X(1 + I)^T / 1000\alpha\X years of production following the period of nonproductive factory growth. The results of this simple calculation for $T = 20$ years are shown in table 5.5 for several representative values of α and I .

What is a reasonable value for α ? The Lunar Manufacturing Facility developed in an earlier section replicates its own mass (of similar components) in one year, or $\alpha = 1$. Waldron et al. (1979) propose a semireplicating factory which can produce its own mass in metal products in less than 6 days, for a maximum $\alpha = 60$. Nevertheless, table 5.5 shows that even if $\alpha = 0.01$ (corresponding to extraordinarily low productivity) the repayment time is still less than a year in a national or global economy with low-to-moderate inflation or interest rates (10% or less). In an economy with interest rates up to 50%, reasonable repayment times -- on the order of typical plant lifetime, about 30 years in usual industrial practice -- remain available for $\alpha > 0.1$ (also a fairly pessimistic lower limit on productivity). Under conditions of hyperinflation (100% and higher) a 30-year breakeven can be obtained only for highly robust, productive systems with $\alpha > 35$.

Economic feasibility, however, is not limited to amortization of costs. A net profit must be made, markets established and maintained, production managed in a reliable

TABLE 5.5.— ECONOMICS OF SELF-REPLICATING FACTORIES

Relative specific productivity, α (\$/yr-\$) ^a	Repayment period of original investment, for an adult seed ^b			
	Inflation = 0%	Inflation = 10%	Inflation = 50%	Inflation = 100%
0.01	1 mo	8 mo	330 yr	100,000 yr
.1	4 d	1 mo	33 yr	10,000 yr
1.0	9 hr	2 d	3 yr	1,000 yr
10	50 min	6 hr	4 mo	100 yr
100	5 min	35 min	12 d	10 yr
1000	30 sec	4 min	1 d	1 yr

^a α = fraction of original value of seed that the adult LMF can produce per year.

^bRepayment period = $\$X \cdot (1 + I)^{20} / (1000\alpha \cdot \$X)$, assuming an initial 20 year nonproductive period.

and flexible manner, and so forth. Given the tremendous power of SRS, severe economic distortions are conceivable across the board. If a replicating factory system is used to flood a market with products, the prices of these products will fall, carrying profits downward as demand saturates in an unregulated economic environment. On the other hand, in a tightly controlled economy the well-known problem of inferior production control feedback would be exacerbated, leading possibly to wild fluctuations in supply and demand for SRS products. These relationships should be investigated more thoroughly by economists.

If control of Earth-deployed replicating factories is retained by national or subnational entities, governments lacking this technology will seek equitable licensing agreements. One interesting problem is ownership of SRS offspring grown from the soil of one country but generated by a leased parent machine owned by another. Should licensing arrangements require return of offspring? Perhaps the offspring should be allowed to remain the property of the licensee, but with royalties levied against production in favor of the owner of the parent machine? Clearly such arrangements could become quite complex in just a few generations of cross-licensing. (SRS capable of "sexual" reproduction present a host of additional theoretical complications.) From the businessman's point of view, it might be better just to sell a "mule SRS" — an infertile factory with the capacity for rapid automated manufacturing but which lacks some vital software element or process necessary for replication. Of course, this is an open invitation to a black market traffic in "bootstrap kits" which allow users to restore fertility to their neutered systems. It is difficult to see how the rapid spread of such technology, once introduced in any form, could be held in check for long by any governmental, corporate, or private entity.

Social aspects of SRS cornucopia. How will humankind deal with what has been termed, with some justification, "the last machine we need ever build?" How might people's lives be changed by a replicative universal constructor system capable of absorbing solar energy and raw dirt and manufacturing, as if by magic, a steady stream of finished building materials, television sets and cars, sheet metal, computer components, and more robots — with little or no human intervention required? Just as the invention of the telephone provided virtually instantaneous long-distance communication, and television permits instant knowledge of remote events, and the automobile allows great individual mobility, the autonomous SRS has the potential to provide humanity with virtually any desired product or service and in almost unlimited quantities. Assuming that global human population does not simply rise in response to the new-found replicative cornucopia and recreate another equilibrium of scarcity at the original per capita levels, supply may be decoupled from demand to permit each person to possess all he wants, and more. The problems of social adjustment to extreme sudden wealth have been documented in certain OPEC nations in recent years. Much attention has also been given to the coming "age of leisure" to be caused by superautomation. What more difficult psychological and social problems might emerge in an era of global material hyperabundance?

If the enterprise of establishing an automated lunar mining and manufacturing facility is successful, there might thereby be made available to humanity a vast supply of energy and useful products. By exporting heavy industry to the Moon, the Earth might be allowed to revert to a more nearly natural state of "controlled wilderness." This should permit the preservation of the animals and plants which people have for so long enjoyed. Although contrary to the

historical evidence, on the negative side people may take their new prosperity as license to exercise their natural biological proclivities and yet further overwhelm this planet with teeming human billions. If this occurs, eventually we shall find that although we might make our Earth into a parkland, the actual effect will be more like Yosemite National Park on a midsummer weekend. This is one problem we must not export to other worlds.

Is there a similar danger that the SRS project, though completely successful as a technological and financial venture, will (much like penny-per-gallon gasoline) encourage profligate behavior heedless of catastrophic negative consequences? What unfortunate things might we do, possessing almost unlimited energy and material resources? Will the possibility of hyperabundance lead not to continued national resolve and focus, but rather to a pervasive national complacency, making us think that all is well, that all has been solved, that things always get solved, and that henceforth we need do little or nothing more to improve our lot? If the system works, and we come to depend on it, growing once more to the limits of our productive inventiveness, will we not be dangerously subject to catastrophic damage as a vital, progressive race?

If space offers any solution to this contradiction between the "good life" and our innate breeding proclivities, it probably will involve the establishment of orbital human colonies. To be practical, these habitats must approach replicating factories in the range of goods and services which they produce. The expense of maintaining a large human colony with direct Earth-based support would be immense, so automated factories most likely must provide the goods and services to support such an operation. Once more the need for SRS facilities in the future of humanity becomes apparent.

Replicating factory systems have the potential to severely disrupt or disable most all modern national economies. The concept of "rate of return" on investments may have to be replaced with the notion of "acceleration of return" for nonterrestrial exponentiating SRS. Will present-day governments and other national and international economic entities support the replicating factory concept if it is seen as a potential threat, capable of rendering obsolete the entire global economic order which now exists and under which they now operate?

Environmental impacts. It has been suggested by Dyson (1979) that it might be possible to design a compact replicating robot which can itself serve as part of an enormous energy-collecting grid. Each machine consists of solar panels on top, power transformers and a universal power grid bus connector, some means of mobility such as tracks or wheels, and manipulators and other subsystems necessary for self-replication. Released, say, in the Arizona desert, one or two SRS could rapidly multiply into a "free" gigawatt-capacity generating system in just a few years.

This could then be tapped by power-hungry municipal utilities or even by individual users.

Moore (1956) also discussed the possibility of replicating machine "plants" turned loose on Earth. In Moore's scenario, a single floating self-reproducing barge is released into the oceans; a few years later, it has multiplied itself into a population of millions, with each unit periodically commuting to shore bearing useful products for mankind derived from the sea (salts, minerals, gold). Reviewing this scenario, Dyson noted that such seagoing SRS might become so numerous that frequent crowding and collisions would occur between them. The "dead bodies" of machines involved in major accidents could slowly accumulate on the ocean floor and along the coastline, causing congestion and representing a menace to navigation. The introduction of machine cannibalism to clean up the mess introduces fresh complications into an already difficult situation — ownership and proper recognition of "dead" machines, destruction control and failsafe mechanisms, nonrecyclable parts, violations of national economic zones, and military applications of the technology.

Environmentalists might perhaps regard SRS released on Earth merely as automated strip-mining robots — yet another sophisticated instrumentality in the hands of those who would mercilessly rape the Earth of its limited resources, leaving behind the ugly scars of profit. There are two responses to this shortsighted view of SRS. First, in the Age of Plenty ushered in by these machines, human society will be sufficiently wealthy to regard environmental integrity and beauty as indispensable outputs of any manufacturing system. These functions may be designed into machines as a matter of course; SRS can be preprogrammed first to strip mine, then reclaim, the land they work. Second, machine replication will make possible significant advances in recycling technology. Given sufficient energy, junkpiles and city dumps may be regarded as low grade "ores" — materials processing robots could be turned loose to analyze, separate, and extract valuable resources. Collection and distribution systems would be streamlined by the use of robot workers constructed at an enormous rate by a sessile self-growing factory complex.

Utilization of the Moon by SRS as proposed in earlier sections may be viewed with outrage by other nations as a predatory attempt to secure a part of the "common heritage of all mankind" for the benefit of America alone. Very drastic alteration of the lunar surface is proposed, raising a question of whether there ought to be reserved areas. Should there be more exploration to determine which regions should be exploited and which should not? Must an environmental impact statement be prepared? As on Earth, lunar surface despoilment in theory may be largely reversed — the machines could be programmed to photograph the original landscape in detail and to restore it after mining operations are finished in that area. A potentially more serious environmental impact is the possible creation

of an appreciable lunar atmosphere during the course of industrial operations conducted on the Moon (Johnson and Holbrow, 1977; Vondrak, 1976, 1977). Even small leakages of gas from millions of SRS could create enough atmosphere to disable or seriously disrupt the operation of mass drivers and other manufacturing facilities requiring vacuum conditions.

5.5.2 Implications for Human Evolution

When contemplating the creation of large, imperfectly understood systems with which we have no prior experience, it is prudent to inquire as to the possibility of unforeseen dangers to our continued existence. In particular, artificial intelligences could conceivably become adversaries, whether they reproduce or not. Similarly SRS might become a threat, independent of their intelligence. Because of the imminence of advanced AI and replicating systems technologies in the next several decades, such questions are no longer merely theoretical but have a very pragmatic aspect.

We must begin to examine the possible problems in creating artificial intelligences or replicating systems which could conceivably become our adversaries or competitors for scarce resources. It is not too early to begin considering the possible kinds of behaviors which advanced machines might display, and the "machine sociobiology" which may emerge. It seems wise to try to identify early any fundamental distinctions between intelligent "natural" biological and advanced "artificial" machine systems. Finally, we should consider the significance of the development of advanced machine technologies to the future of human evolution and also to the broader sweep of cosmic evolution.

To serve mankind. The most immediate, urgent impetus for the development of automation and machine replicative techniques is to improve, protect, and increase the productivity of human society. One way of achieving the goal of human preservation and improvement is to make our mechanical creations intelligent, so that they can automatically do what is good for us. We want them to do this even if we have forgotten to specify what "good" is in each instance. Perhaps we don't even know in all cases how to define "good." For example, consider what would happen if a physically capable, literal-minded idiot were put at the controls of a bulldozer (e.g., Pvt. Zero in the "Beetle Bailey" comic strip, present-day computers, etc.). If told to "drive the bulldozer into the parking lot," the idiot would do exactly that, regardless of whether or not the lot happened to be full of automobiles.

One rather compact statement of what is required for our protection already exists. This has come to be known as "Asimov's Three Laws of Robotics":

(1) A robot may not injure a human being, or, through inaction, allow a human being to come to harm.

(2) A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.

(3) A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws (Asimov, 1950).

This is an excellent prescription of what is required but not of how to accomplish it. Exactly what do these laws entail? The following list of conditions is certainly necessary for the Three Laws to hold:

(1) A robot may not injure a human being.

(2) A robot must use common sense.

(3) A robot must be intelligent.

(4) A robot must be conscious.

Common sense, intelligence, and consciousness are the essence of artificial intelligence research. Even if we cannot exhaustively enumerate the ways to harm a human in any and all circumstances, a robot with the above four properties would protect people to the best of its ability. If it ever did injure a human being it would be because neither we, nor it, foresaw that possibility. But it would immediately perceive its error and would never make the same kind of mistake again. We can do no better ourselves, and usually we do worse.

At the present time we have only the most rudimentary knowledge of what common sense, intelligence, and consciousness are, let alone how to insert these qualities into a robot (Good, 1965). As our computers become ever more complex and pervasive, there is the distinct possibility that these characteristics will arise spontaneously. In this case we would be involved in a totally uncontrolled experiment (Hogan, 1979). If conditions 1-3 were not yet fulfilled, but condition 4 was, the outcome could be catastrophic for mankind. For reasons of self-preservation, we must pursue AI research with the goal of ensuring that capabilities 1-3 are achieved *first*.

The problem with this entire approach is that any machine sufficiently sophisticated to engage in reproduction in largely unstructured environments and having, in general, the capacity for survival probably must also be capable of a certain amount of automatic or self-reprogramming. Such SRS in theory may be able to "program around" any normative rules of behavior (such as the Three Laws) with which it is endowed by its creators. Thus, it might modify its own patterns of behavior, as determined by its basic goals and learned motivational structure.

It is possible to conceive of a machine design containing "read-only" hard-wired goal structures. But hardware specialists will admit that such procedures can be circumvented by sufficiently clever software in large, complex systems. Further, since SRS must be quite adept at physical manipulation it is likely that it will be able to re-wire its own

structure in accordance with perceived operational objectives — assuming it can analyze the functions of its own components as needed for repair or maintenance operations. It may be of no use to try to distribute the hard-wired functions throughout the whole machine, or a large subset thereof, in hopes that the system will be unable to comprehend such a large fraction of itself simultaneously. Omitting the special functions from the machine's stored genetic description of itself would probably be equally ineffectual. Laing (1975, 1977) has shown that machine reproduction by complete self-inspection — wherein the parent knows virtually nothing about its own structure to begin with — is quite possible, and has provided several logical designs for such machines. Consequently, it is not possible to logically exclude the possibility of conscious alteration of hard-wired robot "commandments" by intelligent self-replicating machines desirous of doing so.

It would therefore appear nearly impossible, as with people, to absolutely guarantee the "good" behavior of any common-sense, intelligent, conscious SRS. However, also like people, machines can be taught "right" and "wrong," as well as the logical rationales for various codes of behavior. And they can probably be expected to remain even more faithful to a given moral standard than people.

SRS population control. An exponentially increasing number of factories (even if the rate is not sustained indefinitely) will seem especially threatening and psychologically alarming to many. Such a situation will draw forth visions of a "population explosion," heated discussions of *lebensraum*, cancerous growth, and the like. Nations not possessing replicating systems technology will fear an accelerating economic and cultural gulf between the haves and the have-nots. On another level altogether, humankind as a species may regard the burgeoning machine population as competitors for scarce energy and material resources, even if the net return from the SRS population is positive.

Of course, self-replicating factories are not ends in themselves but have specific purposes — say, to produce certain desired products. The quantity of these products is determined by needs and requirements and is the basis for designing an SRS. Depending on the type of product, factors such as the time when these products need to be available, the production time, and replication time per replica determine the optimum number of replica factories per primary and the number of generations required. The following controls might be used to achieve this condition:

(1) The "genetic" instructions contain a cutoff command after a predetermined number of replicas. After each replica has been constructed one generation command is marked off until at the last predetermined generation the whole process is terminated after the final replica is completed. Besides all this, engineers may have their hands full keeping the SRS replicating on schedule and function-

ing properly. It is not likely that they will soon be able to do much more than we expect.

(2) A predetermined remote signal from Earth control over a special channel can easily cut the power of the main bus for individual, groups, or all SRS at any time. Replication energy production shows one of the fundamental differences between biological and mechanical replicating systems as presently conceived. In biological systems energy is generated in distributed form (in each living cell throughout the entire organism) whereas in mechanical systems such as SRS energy is produced centrally in special parts (e.g., power plant, solar cells) and then is distributed to wherever it is needed. This should make control of mechanical systems comparatively easy.

For replicating systems much smaller than factories, say, in the 10^2 – 10^4 kg category, the situation may be somewhat different. One potential problem with such devices is that once started, their multiplication may be difficult to stop. As a reasonably large population accumulates, it may become almost physically impossible for humans to maintain any semblance of control unless certain precautions are taken to severely limit small-machine population expansion. In many ways a large population of low-mass SRS resembles a biological ecology. While the analogy is imperfect, it serves to suggest some useful ideas for automata population control once people determine that direct control of the situation has somehow been lost.

Predation is one interesting possibility. Much as predator animals are frequently introduced in National Parks as a population control measure, we might design predator machines which are either "species specific" (attacking only one kind of SRS whose numbers must be reduced) or a kind of "universal destructor" (able to take apart any foreign machine encountered, stockpiling the parts and banking the acquired information). Such devices are logically possible, as discussed earlier in section 5.2, but would themselves have to be carefully controlled. Note that a linear supply of predators can control an exponentiating population of prey if the process of destruction is, as expected, far more rapid than that of replication.

Clearly it is easier to design the solution to this problem into the SRS from the start, as suggested above in reference to larger factory systems. For instance, machines might be keyed to respond to population density, becoming infertile, ceasing operations, reporting (like lemmings) to a central disassembly facility, or even resorting to dueling or cannibalism when crowding becomes too severe. However, a method by which the materials and information accumulated by SRS units during their lifespans can be preserved would be in the best interests of human society.

The unpluggability problem. Many people, suspicious of modern computers and robotics technology, take solace in the fact that "no matter what goes wrong, we can always

pull out the plug." Such individuals might insist that humankind always retain ultimate life-and-death control over its machines, as part of the social price to be paid to permit their development. Whether this is advisable, or even necessary, is a question which requires further study. Certainly it is true that our civilization all too easily becomes habituated to its machines, institutions, and large organizations. Could we unplug all our computers today? Could we "unplug" the Social Security Administration? It is difficult, or impossible, and in many cases ill-advised, to retreat from a social or mechanical technology once it has been widely introduced and a really significant change has taken place. Many individuals in our society would prefer to turn back the clock on the industrial revolution, but today this could not be done without the sacrifice of hundreds of millions of human lives and extreme trauma to global civilization.

Further, we must assume that we cannot necessarily pull the plug on our autonomous artificially intelligent species once they have gotten beyond a certain point of intelligent development. The one thing the artificial system may learn is how to avoid a human being "pulling its plug out," in the same way that human beings come to understand how to defend themselves against other people (George, 1977). Consequently, it is imperative that we study ways to assure ourselves that our technological creations will serve to our benefit rather than to our detriment, as best we can, prior to their widespread adoption.

Assuming we wish to retain ultimate control over our creations (by no means a foregone conclusion), the team first considered, as a theoretical issue, the following intriguing problem:

Is it logically possible to design an internal mechanism which permits normal SRS functioning to be interrupted by some external agency, yet which is impossible for the SRS itself to circumvent either by automatic reprogramming or by physical self-reconstruction?

That is, is it impossible to build a machine whose "plug" cannot be "pulled"?

Machine capabilities of the future span a wide spectrum of sophistication of operation. As systems become more complex, individual human beings will come to understand decreasing fractions of the entire machine function. In addition, very advanced devices such as SRS may need to be programmed with primitive survival instincts and repair capabilities if they are to function autonomously. At some point the depth of analysis and sophistication of action available to a robot system may exceed the abilities of human minds to defeat it (should this become necessary). If there is even the slightest possibility that this is so, it becomes imperative that we learn exactly what constellation of machine capabilities might enable an SRS to cross the subtle threshold into "theoretical unpluggability."

To this end, the team subsequently reformulated the unpluggability question as follows:

What is the *least* sophisticated machine system capable of *discovering* and *circumventing* a disabling mechanism placed within it?

While no specific firm conclusions were reached, the team concluded that the simplest machine capable of thus evading human control must incorporate at least four basic categories of intelligence or AI capabilities (Gravander, personal communication, 1980):

- (1) Class invention, concept formation, or "abduction"
- (2) Self-inspection
- (3) Automatic programming
- (4) Re-configuration or re-instrumentation capability (especially if the "plug" is in hardware, rather than software)

These four characteristics are necessary preconditions for theoretical unpluggability -- a machine lacking any one of them probably could not figure out how to prevent its own deactivation from an external source. Whether the conditions are sufficient is an urgent subject for further research.

Sociobiology of machines. The creation of replicating manufacturing facilities, remotely sited, and for long times left under their own control, poses some very special problems. In order to eliminate the use of humans as much as possible in a harsh environment, these systems of machines should be designed to seek out their own sources of materials; to decide on this basis to invest in duplicates of themselves; to determine their power requirements and see to the construction of requisite new power sources; to monitor their own behaviors; to undertake the repair of machines, including themselves; to determine when machines have, under the conditions obtaining, reached the end of their useful working lives; and so forth. They must operate reliably and resist corrupting signals and commands from within their own numbers, and from without. They must be able to discern when machines (whether of their own sort or not) in their neighborhood are, by their behavior, disrupting or endangering the proper functioning of the system. Since we cannot foresee all of the ways in which the system may be perturbed, we shall have to supply it with goals, as well as some problem-solving or homeostatic capabilities, enabling the machines to solve their own difficulties and restore themselves to proper working order with little or no human assistance.

As SRS make duplicates of themselves, the offspring will, if suited to surroundings different from those of their parents, differ somewhat from them. More of one sort of submachine or subordinate machine may be required, fewer of another. The main "genetic" program will undoubtedly increase in size, generation by generation. At removed locations constructor-replicators may symbiotically combine with miners, surveyors, and fabricators, to form satellite

machine communities differing considerably from the original population.

At this stage it may be that some of the claims made by evolutionary biologists as to the likely origin of complex, social behavior of animal populations may begin to apply to machine populations. Indeed, it may be that the arguments of the sociobiologists will be *more* applicable to machines than to animals and humans. In the case of animals, and especially in regard to humans, the opponents of the evolutionary biologists insist on the priority of alternative sources for social behavior — namely, individual learning. Behavior need not have its origins in the genome. These opponents of the evolutionary biologists constantly challenge them to specify where in the genome is the locus of selfishness, distrust of strangers, aggression, and the like. This is not really readily done.

However, in the case of machines the locus of behavior can indeed be specified: It is in the program of instructions, and these programs can, like genes, be modified and transmitted to offspring. Though we may not be mere machines driven by our genes, mere real machines are indeed driven by their gene-like programs, and for them, some of the evolutionary biological predictions of the likely resulting system behaviors may apply.

Thus, at the most elementary level, if some one of the SRS machines capable of duplicating itself begins to concentrate on this reproductive activity to the neglect of all other tasks we intend for it, its progeny (possessing the same trait) might soon become dominant in the machine population. But far more complex aberrations and consequent elaboration of machine behavior can arise and be propagated in machine populations of the sophistication we may be forced to employ.

Thus, our machines can reproduce themselves as well as tell whether an encountered machine is “like” themselves or could be one of their offspring. If the structure of an encountered machine is examined and found to be similar to or identical to the machines of one’s own making or of one’s own system of machines, then such machine should be welcomed, tolerated, repaired, supplied with energy and consumables, and put to work in the common enterprise. If, on the other hand, the structure of an encountered machine deviates greatly from that of any of one’s own system of machines — even if it is in fact a device of one’s own construction which has suffered severe damage or defect of construction — then prudence suggests it should be disabled and dismantled.

It is interesting to note that this “reasonable” kin-preferring behavior could arise generally throughout the machine population quite without it having been made a deliberate part of the programs of machines of the system (Hamilton, 1964). If a single machine of the sort which reproduces ever chances upon the program “trait” of tolerating machines like itself, or aiding or repairing them

while ignoring, disabling, or dismantling machines unlike itself and its offspring, then this machine species will tend to increase its numbers at the expense of other reproducing machines (all other things being equal) so that after a few generations all machines, quite without having been given the goal or purpose of preferring their own kind, will have this kin-preferring property. Other types of machines that are less kin-supportive would not leave relatively so many of their kind to further propagate. This is the familiar biological selection principle of differential reproduction.

This argument can be carried further. In a society of machines in which it “pays” to know which machines are your “relations,” it will become risky to undertake or to submit to close structural inspection as this will reveal what sort of machine you really are — friend or foe. Instead, behavioral cues will likely develop that signal whether a machine is kin or not. Unfortunately, such signals can equally well be used to deceive. A machine could learn to give the kinship sign even though it is not at all a relation to the encountered machine, or friendly either. It may use the conventional sign of friendship or kinship merely as a means of soliciting undeserved assistance (e.g., repair, materials, energy) from the deceived machine and the system of subordinate machines with which it is associated, or may even use the signals of kinship or friendship as a means of approaching close enough to disable and dismantle the deceived machine.

The evolutionary argument should be cast as follows. Any machine which chances upon a behavioral sign that secures the assistance of a machine or a population of machines will be spared efforts at survival it would otherwise have to undertake on its own, and thus will possess extra resources which can be utilized to undertake the construction of more machines like itself. If the “deceitful signal” behavior is transmitted in the genetic-construction program, then its offspring will also be able to employ the deceitful signal, and will thus produce proportionately more of their kind. The deceitful gene-program machines will increase their numbers, relative to the others, in the machine population. In turn, those machines which chance upon ways of detecting this deceit will be protected against the cheating machines, and will themselves increase their numbers vis-a-vis their “sucker” related machines who will soon be spending more and more time aiding, servicing, and supplying cheaters (thus have fewer resources in the form of time, energy, and materials to reproduce their own kind).

It is even possible that in a largely autonomous system of reproducing machines a form of *reciprocal altruism* will arise, in which machines behave in seemingly unselfish fashion toward other machines which are not kin (and are not deceitfully posing as kin). The evolutionary biologists, especially Trivers (1971), have argued that in situations where the reproducing entities have (1) long lifespans,

(2) remain in contact with others of their group, and (3) experience situations in which they are mutually dependent, reciprocal altruism may arise out of chance variation and evolutionary selection. In human terms, if helpful actions can be taken which are low risk to the giver and have a high value to the receiver (high/low risk defined relative to the impact on individual reproductive potentials) and there is the likelihood that the individuals will remain in fairly close association for a long time, then any genetic predisposition to take altruistic actions will tend to spread in the population. For, in effect, it will lead to reciprocal assistance in times of need, to the greater survival (and hence increased breeding opportunity) of those members of the populations bearing this genetic trait. A good example is that of an individual saving another from drowning by reaching out a branch. The risk to the giver is small, and the benefit to the receiver is great, and over a long time the benefits (in terms of increased numbers of offspring) are likely to be great, to those members of the population genetically predisposed to behave in this reciprocally altruistic fashion.

Needless to say, the opportunities for deceit and cheating in the case of hoped-for altruistic reciprocity are even more numerous and complex than for kin selection strategies. In particular, each individual (animal or machine) must possess the memory capacity to remember the altruistic acts and the partners in them, since the opportunity for reciprocity may not arise for some time. Also some cost-benefit analysis must take place in which the value of the act, the character of the reciprocity partner, the capacity of this partner to repay, and the likely lifespan of the giver and receiver all must be carefully weighed. Some evolutionary biologists would go so far as to claim that purely genetic (and hence "mechanical") workings out of such subtle relationships drove the hominid brain, in a few million years, from dullness to sophistication. A few even suggest that the origins of human language lie in the process of making claims of kinship (while possibly being no relation at all), of offering friendship (while possibly intending harm), and promising future assistance (while intending, when called upon, to turn away).

If our machines attain this level of behavioral sophistication, it may finally not be amiss to ask whether they have not become so like us that we have no further right to command them for our own purposes, and so should quietly emancipate them.

Entropy, SRS and biology. Nature has provided on Earth an example of the primary generation of self-replicating biological systems from energy and matter alone. The second law of thermodynamics states that the entropy of *energy* continually increases. At the moment of the Big Bang, it may have been zero and today it spreads between a lower boundary that covers neutron stars and black holes and an upper boundary indicated by the 3 K

background radiation. At the same time *matter* decreased in entropy from practical infinity at the moment of the "Big Bang" to a lower boundary evolving from hydrogen atoms to light elements, heavy elements, life, to the human brain – towards ever more complex structures, generally more intelligent matter, limited by the upper boundary of elemental particles. Matter tends to evolve toward greater complexity at the expense of energy, which in turn acquires increasing entropy. (See fig. 5.25.)

The generation of a desired material order which may represent an SRS and its self-description would recapitulate biology-like evolution in engineering terms. However, there may be one fundamental difference between the two. Living organisms have two separate information systems that help determine their behavior: DNA and the brain. Between these two there is no direct information transfer, perhaps instead only indirect sociobiological influences. DNA information is initially provided to the organism, whereas brain information is gradually acquired through diverse environmental interactions. In SRS there is not this differentiation – initially provided information is the principal driver of actions and is accessible to the SRS intelligence (fig. 5.26).

Man-machine co-evolution. In the very long term, there are two possibilities for the future of mankind: Either we are a biological "waystation" in the evolutionary scheme of the universe, or else we are an evolutionary dead end. If we continue to be limited to our exceedingly fragile existence on spaceship Earth, a natural disaster or our own jingoistic or ecological foolhardiness is almost certain to terminate our existence perhaps centuries or millenia from today. Barring these unpleasant circumstances, our civilization, without the challenge of a frontier, may stagnate while other beings flourish and populate the universe.

Replicating systems technology gives humanity other options for continued and fruitful evolution. We can create autonomous (unmanned) SRS – in a very real intellectual and material sense our offspring – and send them out into the cosmos. Alternatively, we could create a symbiotic human-machine system in which people would inhabit a vast self-reproducing habitat. This is analogous to creating an artificial Earth which replicates itself whenever its population of humans fills the available space or saturates the energy supply or waste disposal facilities. In the process of working to achieve the second goal, mankind could use SRS to attempt terraforming other worlds. Experiments could be performed on planetary-scale weather modification with relevance to maintaining or changing the Earth's climate.

At present, machines already "reproduce" themselves but only with human help. Such mutualism is commonplace in biology – many flowering plants require cross-pollination by insects to survive. The most successful organism is one which can enlist the most reproductive assistance from other creatures. It has often been suggested

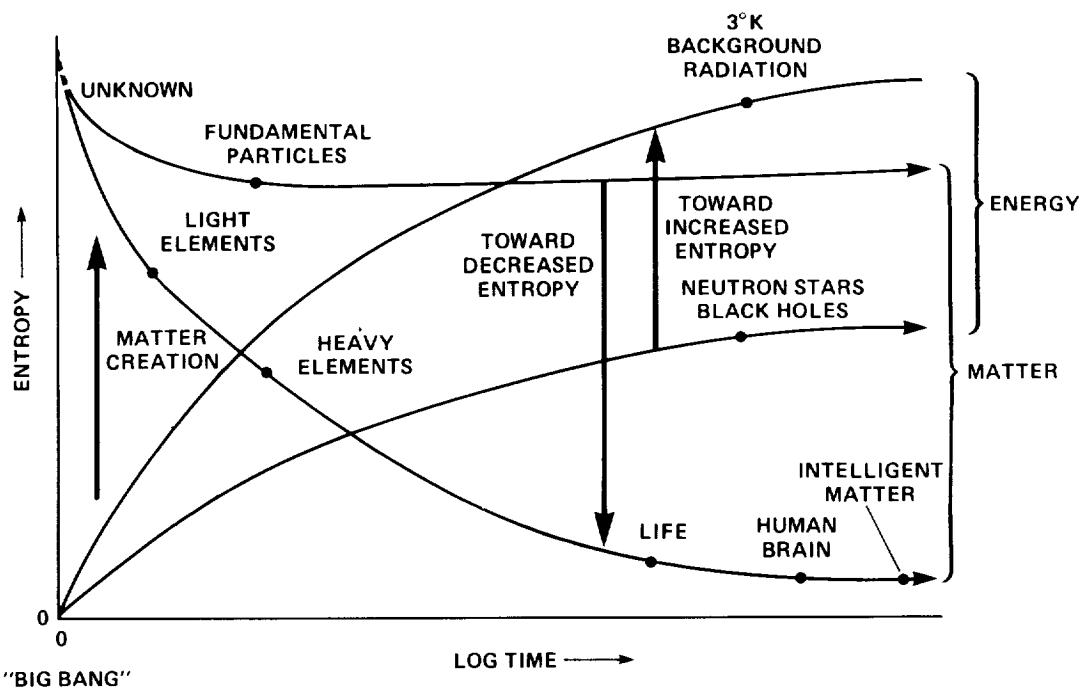


Figure 5.25. Natural evolution of complexity of matter in the cosmos.

that an extraterrestrial biologist who chose Los Angeles as the site for his field study of Earth might well conclude that the automobile was the dominant lifeform on this planet and that humans represented its detachable brains and reproductive organs. Indeed, further observation might suggest that many people are redundant although the human population of Los Angeles has remained relatively constant during the past decade, the car population has continued to increase.

This issue has tremendous importance to the question of human survival and long-term evolution. Asks Burhoe (1971): "Will we become the 'contented cows' or the 'household pets' of the new computer kingdom of life? Or will *Homo sapiens* be exterminated as *Homo sapiens* has apparently exterminated all the other species of *Homo*?" Perhaps machine-wrecking New Luddites of the future will band together to form secret organizations devoted to "carbon power" and the destruction of all silicon micro-electronic chips and robotic devices.

Are we creating a new "kingdom of life," as significant as the emergence and separation of plant and animal kingdoms billions of years ago on Earth? Or perhaps such an event has even greater import, since "machine life" is of a totally different material substance than either animal or plant life, and because "machine life" very possibly is a form which cannot evolve by direct natural routes but instead requires a naturally evolved biological creator. In addition, while human brains process data at a rate of about 10^{10} bits/sec/kg, silicon computer microprocessors operate

at 10^{16} - 10^{22} bits/sec/kg. This enormous disparity in potential intelligence has given some people great cause for alarm. For example, according to Wesley (1974):

In terms of the 4.5 billion years of carbon-based life on Earth, the advent of machines has been amazingly abrupt. Yet the evolution of machines is subject to the same laws as the evolution of ordinary carbon-based life. Machines have also evolved toward an increased biomass, increased ecological efficiency, maximal reproduction rate, proliferation of species, mobility, and a longer lifespan. Machines, being a form of life, are in competition with carbon-based life. Machines will make carbon-based life extinct.

Not everyone is so unduly pessimistic. Of course, if we create SRS then we will find ourselves co-inhabiting the universe with an alien race of beings. But the ultimate outcome is unknown: we could dominate them, they could dominate us, we could co-exist as separate species, or we could form a symbiotic relationship. This last is the most exciting possibility. Humankind could achieve the simultaneous perpetuation and development of its being and expansion of its niche of the Universe. At the price of being a part of a larger system mankind could achieve immortality for itself. The Earth was a gift of creation, but someday people may have the opportunity to make many more such systems of their own choosing.

Automated space habitats could serve as extraterrestrial refuges for humanity and other terrestrial lifeforms that

man might choose to bring along as insurance against global terrestrial catastrophes. The dispersal of humankind to many spatially separated ecosystems would ensure that no planetary-scale disaster, and, as people travel to other stars, no stellar-scale or (ultimately) galactic-scale event could threaten the destruction of the entire species and its accomplishments.

5.5.3 Philosophical, Ethical and Religious Questions

New developments in science and technology frequently have profound religious and philosophical consequences. The observation that, rather than being the center of a rather small universe, the Earth is but a small frail speck of a spacecraft in an unimaginably enormous universe is only

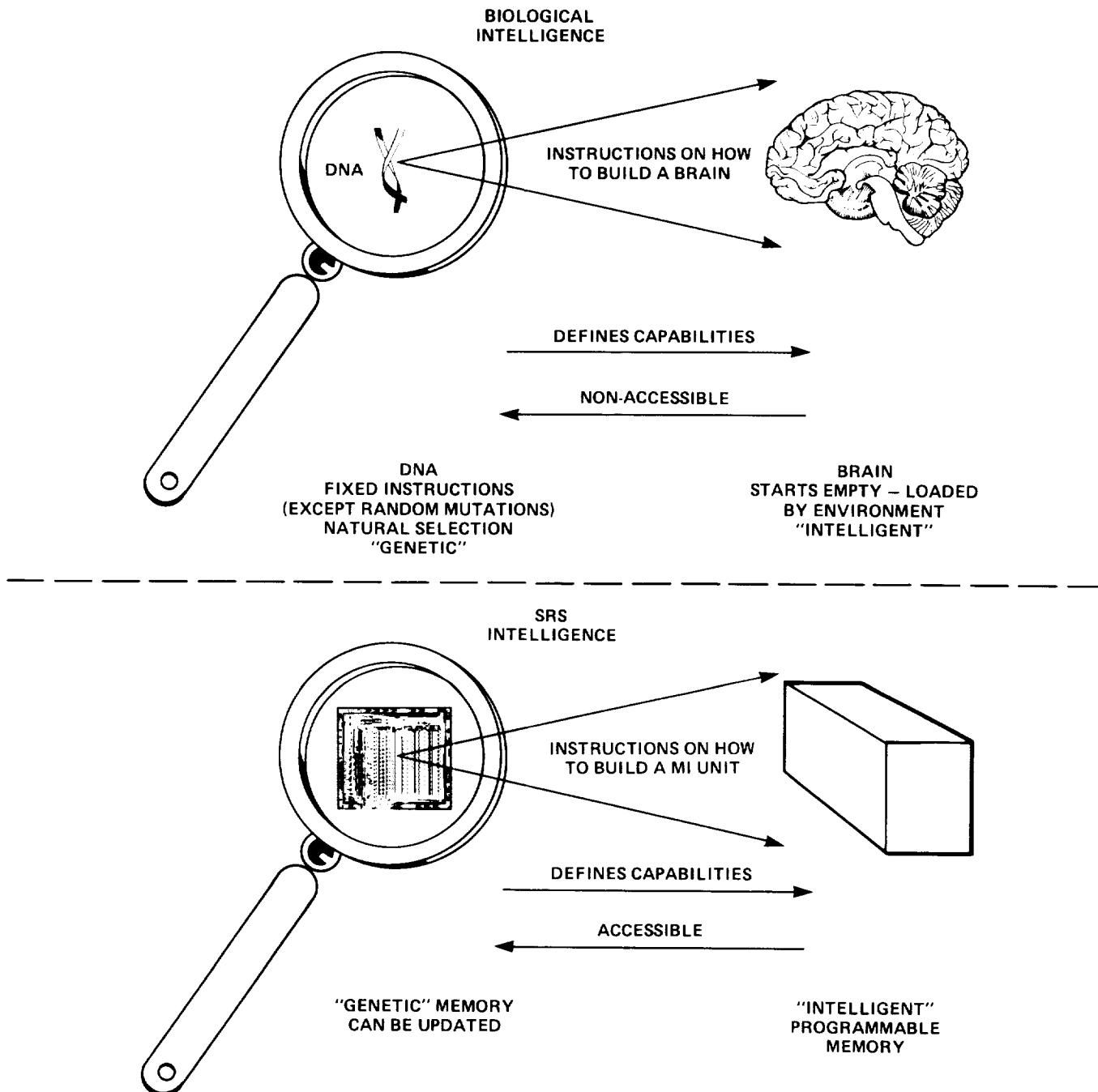


Figure 5.26.-- Accessibility of biological and machine-stored information.

just now beginning to be appreciated and woven into the fabric of human religion, philosophy, and culture. The existence of an alien race of beings, as alive as we are, would similarly challenge our old beliefs. We may encounter this alien race either through SETI or through our own technological creation.

According to British Agriculture Minister Peter Walker, "Uniquely in history, we have the circumstances in which we can create Athens without the slaves." However, if robots gain intelligence, sensitivity, and the ability to replicate, might not they be considered legal persons, hence slaves? Is mankind creating a new form of life to enthrall to its bidding? Is it immoral to subjugate an entity of one's own creation? (Modern jurisprudence, in contrast to ancient Roman law, does not permit a parent to enslave his child.) Questions of "machine rights" or "robot liberation" undoubtedly will arise in the future. And if the intelligence or sensitivity of robots ever exceeds that of humankind, ought we grant them "civil rights" superior to our own? Many ethical philosophers, particularly those who support the contemporary "animal liberation" movement, might answer in the affirmative.

Could a self-reproducing, evolving machine have a concept of God? It must understand the concept of creation, since it itself creates other machines during the processes of self-replication and growth. Thus, it should recognize the role of creator. If it was aware that mankind had created it, would it view its creator as a transcendent active moral entity simply because of our role as creator? Or would it tend to view humanity much as we view lemurs and chimpanzees — ancient species that served as an important link in an evolutionary chain, but which is now merely another "lower order" of life? Would humankind be seen as nothing more than an evolutionary precursor?

Perhaps not. *Homo sapiens* evolved from more primitive mammals, not by conscious design but rather by evolution acting through differential reproduction in response to arbitrary environments. It would be silly for people to revere mammals as their gods — these animals did nothing to actively *cause* the emergence of the human race. On the other hand, humans *may* purposely engender the creation of intelligent reproducing machines whose emergent philosophy we are considering. Our role is clearly much more than that of passive precursor; rather, it is one of active creator — conceiving, planning, designing, developing, building, programming, and deploying the SRS. It seems plausible that, for this reason, mankind might also expect to play a more active role in any "machine theology" that might ultimately develop.

Related theological issues include: Could conscious, intelligent machines have a soul? Or, what is for many purposes equivalent, will they *think* they have a soul? How will human religions respond to the prospect of an intelligent machine capable of self-replication? Are there any Scriptural prohibitions or pronouncements applicable in this

matter? Is it possible to view the machine as possessing a "soul"?

What of man's view of himself? He now takes pride in his uniqueness. How will he adjust to being just an example of the generic class "intelligent creature"? On the other hand, the concept of "God" may take as much a beating as the notion of "man." After all, He is special now because He created us. If we create another race of beings, then are we not ourselves, in some similar sense, gods?

Is ethics as a concept of moral behavior a purely human or purely biological construct, or is the notion tied to evolutionary universals and environmental/developmental imperatives which will prove equally applicable to advanced intelligent machines? If machines are capable of developing their own systems of "ethics," it would probably appear as alien to human eyes as does the behavior of other animal species (e.g., the apparent "cruelty" of many insect species).

Will advanced machines have any artistic urges, a sense of humor, curiosity, or a sense of irony, or are these kinds of responses confined exclusively to biological creatures capable of displaying emotion? It is unknown whether machines even need emotionality — we are only beginning to understand the functions of these responses in mammals and humans.

Will a vast industrialized lunar complex of interacting systems be vulnerable to catastrophic accidents and breakdowns, or to attack, subversion, or disruption, either by unexpected machine responses generated out of the complexity of their interactions, or by the interference of one or more unfriendly powers on Earth? Are there subtle ways in which the lunar complex could be subverted? SRS systems, to the extent they are highly sophisticated machines and autonomous, may be subject to some forms of attack and subversion not hitherto realized. Spurious signals may be injected, or foreign machines may enter the works, for example. How might subversive signals and invading software "viruses" be detected and resisted? What identification of friendly and unfriendly machines should be employed? Which is most reliable? What means of information and control message security should be adopted? These questions will take on greater urgency as SRS come to represent ever-increasing shares of the global industrial economy.

Finally, might replicated robot warriors, war machines, or other SRS-derived combat systems make war "too horrible to contemplate"? Perhaps machine wars will still be fought, but will be exported into space to preserve the Earth. Maybe all conflicts will be fought only in computer simulations as "war games"? Or, the availability of sophisticated autonomous fighting machines might lead instead to an *increase* at least in small-scale wars, because of the low cost of such devices, the unlikelihood of human injury in autonomously waged conflicts, and because of possible increasing human boredom in a society of extreme physical

safety and material hyperabundance due to superautomation.

5.5.4 Cosmological Implications

According to Valdes and Freitas (1980), any sentient extraterrestrial civilization desiring to explore the Galaxy beyond 100 light-years from its home star should find it more efficient and economical to use self-replicating star-probes because of the benefits of exponentiation. This will secure the largest quantity of data about extrasolar systems by the end of an exploration program of some fixed duration. The entire Galaxy can be explored in times on the order of 10^6 years assuming interstellar cruising speeds on the order of $0.1c$, now considered feasible using foreseeable human technology (Martin, 1978). Many who have written on the subject of theoretical galactic demographics have suggested that most extraterrestrial races probably will be found 100 to 1000 light-years from Earth and beyond. Hence it may be concluded that the most likely interstellar messenger probe we may expect to receive will be of the reproducing variety.

One of the tremendous advantages of interstellar probes over interstellar beacons in the Search for Extraterrestrial Intelligence (SETI) is that probes may serve as cosmic "safety deposit boxes" for the cultural treasures of a long-perished civilization (Freitas, 1980d). The gold-anodized Voyager records are a primitive attempt to achieve just this sort of cultural immortality (Sagan, 1978). Starfaring self-replicating machines should be especially capable of maintaining themselves against the disordering effects of long periods of time, hence SRS will be preferentially selected for survival over nonreproducing systems. This fact, together with the aforementioned preference for using SRS for very long-term, large-distance galactic exploration implies that any alien machine we might find in our own solar system (as part of a dedicated SETI effort; see Freitas and Valdes, 1980) still in adequate working order will most probably be a replicating system.

A number of fundamental but far-reaching ethical issues are raised by the possible existence of replicating machines in the Galaxy. For instance, is it morally right, or equitable, for a self-reproducing machine to enter a foreign solar system and convert part of that system's mass and energy to its own purposes? Does an intelligent race legally "own" its home sun, planets, asteroidal materials, moons, solar wind, and comets? Does it make a difference if the planets are inhabited by intelligent beings, and if so, is there some lower threshold of intellect below which a system may ethically be "invaded" or expropriated? If the sentient inhabitants lack advanced technology, or if they have it, should this make any difference in our ethical judgment of the situation?

Oliver (1975) has pointed out that the number of intelligent races that have existed in the past may be significantly

greater than those presently in existence. Specifically, at this time there may exist perhaps only 10% of the alien civilizations that have ever lived in the Galaxy -- the remaining 90% having become extinct. If this is true, then 9 of every 10 replicating machines we might find in the Solar System could be emissaries from long-dead cultures (fig. 5.27).

If we do in fact find such machines and are able to interrogate them successfully, we may become privy to the doings of incredibly old alien societies long since perished. These societies may lead to many others, so we may be treated, not just to a marvelous description of the entire biology and history of a single intelligent race, but also to an encyclopedic travelogue describing thousands or millions of other extraterrestrial civilizations known to the creators of the probe we are examining. Probes will likely contain at least an edited version of the sending race's proverbial "Encyclopedia Galactica," because this information is essential if the probe is to make the most informed and intelligent autonomous decisions during its explorations.

Further, if the probe we find has been waiting near our Sun for long enough, it may have observed such Solar System phenomena as the capture of Phobos, the upthrusting of the Rocky Mountains or the breakup of Pangaea, the formation of the Saturnian rings, the possible ejection of Pluto from Neptunian orbit, the possible destruction of a planet in what is now the Asteroid Belt, the origin of the Moon, or even the formation of our own planetary system. Perhaps it could provide actual visual images of Earth during the Jurassic or Carboniferous eras, or data on the genomes of long extinct reptiles (e.g., dinosaurs) or mammals, possibly based on actual samples taken at the time. There are countless uses we could make of an "intelligent eye" that has been watching our planet for thousands or millions of years, meticulously recording everything it sees.

SRS probes can be sent to other star systems to reproduce their own kind and spread. Each machine thus created may be immortal (limitlessly self-repairing) or mortal. If mortal, then the machines may be further used as follows. As a replicating system degrades below the point where it is capable of reproducing itself, it can sink to a more simple processing mode. In this mode (useful perhaps as a prelude to human colonization) the system merely processes materials, maybe also parts and subassemblies of machines, as best it can and stockpiles them for the day when human beings or new machines will arrive to take charge -- and make use of the processed matter which will then be available. As the original machine system falls below even this level of automation competence, its function might then be redirected to serve merely as a link in an expanding interstellar repeater network useful for navigation or communications. Thus, at every point in its lifespan, the SRS probe can serve its creators in some profitable capacity. A machine which degrades to below the ability to self-reproduce need not simply "die."

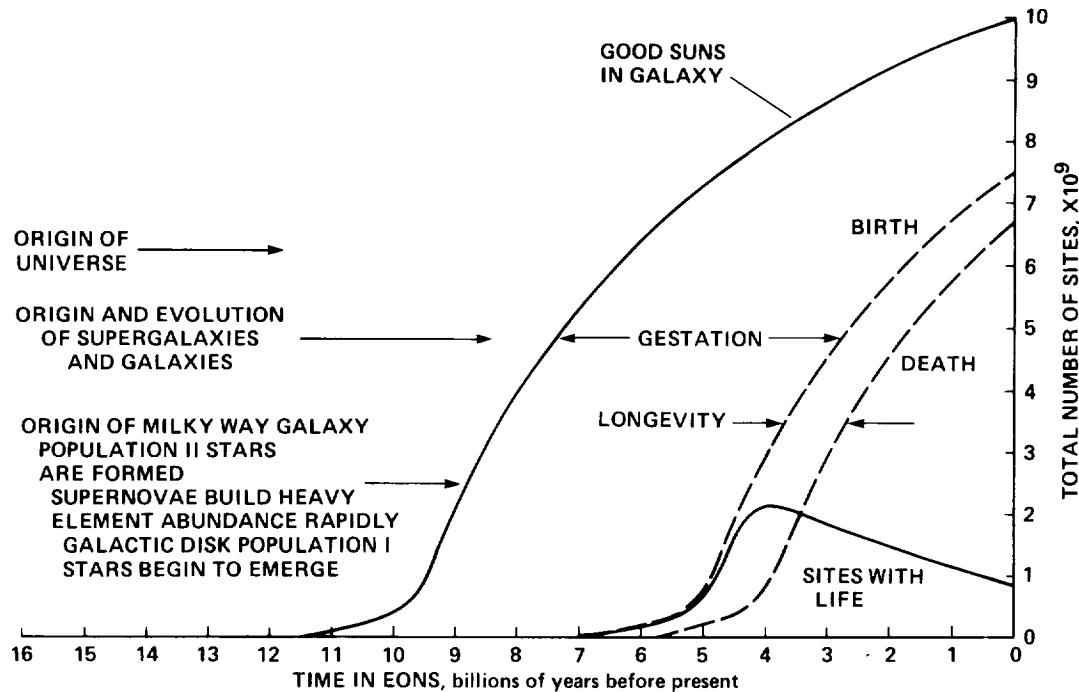


Figure 5.27. Population of extraterrestrial civilizations as a function of galactic time.

The SRS is so powerful a tool that it could have implications on a cosmological scale. With the SRS humanity could set in motion a chain reaction of organization sweeping across the Universe at near the speed of light. This organized part of the Universe could itself be viewed as a higher level “organism.” Instead of merely following the laws of mechanics and thermodynamics, something unique in our knowledge would occur. The degree of cosmic organization would increase. Life could become commonplace, whereas now it seems quite rare. New rules, the rules of life, would spread far and wide.

5.6 Realization

John von Neumann, and a large number of other researchers in theoretical computer science following him, have shown that there are numerous alternative strategies by which a machine system can duplicate itself. There is a large repertoire of theoretical computer science results showing how machine systems may simulate, construct, inspect, and repair machine systems including, to some extent, themselves. This repertoire may be useful in the design of actual replicating machine systems.

The basic concept of physical machines capable of useful self-replication is credible both from a theoretical and a practical engineering standpoint. It is reasonable to begin designing replicating systems based on current knowledge and state-of-the-art technology, though final design resolu-

tion will require significant additional research. Complete systems closure is achievable in principle, though partial closure may be more feasible from an economic and pragmatic engineering standpoint in the near term. It also appears feasible to begin immediate work on the development of a simple demonstration SRS on a laboratory scale, with phased steps to more sophisticated levels as the technology is proven and matures.

Self-replicating systems appear potentially useful in an economic or commercial sense. The main advantage in using SRS over other methods of space exploration and industrialization is that a very large capability for performing any desired task can be rapidly achieved at almost any remote location, starting with a relatively small investment of time, money, energy, and mass in the original “seed” mechanism. SRS will have many applications on Earth, in near-Earth and lunar space, throughout the Solar System, and in the interstellar realm, for future exploration and utilization, suggesting a number of significant social, cultural and economic impacts on American and human society.

In this section the Replicating Systems Concepts Team sets forth in some detail how NASA may take action at once toward the achievement of the ultimate goal of establishing a replicating manufacturing facility. A suggested statement of work (SOW) and a list of institutions which might undertake the tasks outlined in the work statement are included.

5.6.1 Prologue to Realization

The space program of the United States is at a critical point in its evolution. The easy missions, for the most part, have been accomplished. These have been limited to what could be done within the lift capacity of one or two launch vehicles. The capabilities of the payloads which have been delivered to space have been limited by (1) the rudimentary nature of payload automation (either preprogrammed or teleoperated), (2) the high penalty for life support systems and of man-rating manned payloads, and (3) the high cost of the Earth-based mission operations.

The industry of the U.S. is also at a critical juncture in its evolution. If it is to compete adequately in the world marketplace, significant increases in productivity are required. Present production methods have reached a level of maturity such that sufficiently large gains in productivity through further refinement of present-day technologies are unlikely to be realized. The only known solution is massive automation such as is now being applied in other industrialized countries, notably Japan and Germany.

Massive automation would dramatically increase the capabilities and effectiveness of the space program. Use of the emerging techniques of machine intelligence would make it possible to perform missions which previously would have required men *in situ*, thus prohibitively expensive. Highly automated programmable manufacturing by robots would permit the economical production of small numbers of spacecraft for exploratory missions. Missions which require the manipulation of large amounts of mass off-Earth (e.g., lunar/orbital bases or solar power satellites) are especially amenable to massive automation. These missions can be accomplished by employing large numbers of cheap freight rockets mass-produced by robots in automated factories and launched by robots at automated launch facilities (Cliff, Summer Study Document, 1980). These missions might also be accomplished by extraterrestrial automated manufacturing of the required hardware. In any case, the key is massive automation.

One of the most significant characteristics of massive automation is the possible regenerative or "bootstrapping" effect. Using robots to make robots will decrease costs dramatically, thus expanding the economically viable uses of robots. This in turn increases demand, leading to yet further automation, which leads to lower-cost robots, and so on. The end result is "superautomation" (Albus, 1976). A similar effect has already been noted in the computer industry where dramatic increases in performance/price have continued unabated over three decades. The use of robots to help manufacture robots, analogous to the use of computers to help make computers, should produce a similar effect. Extensive innovation should continue unabated for quite some time in such a young field.

Work is now in progress in Computer-Aided Design and Manufacture (CAD/CAM) in the United States. A partial

bibliography of recent work in this area and a list of manufacturers, equipment directory, and supplier addresses have been published (Gettleman, 1979; "Numeric Control Equipment," 1980). Several bills designed to promote automation are presently before the U.S. Congress. The Department of Commerce is beginning a program to promote industrial automation in this country. The National Science Foundation also is funding work in automation. The Department of Defense has initiated a large effort in Integrated Computer-Aided Manufacturing (ICAM) (*Business Week*, 1980). ICAM combines both CAD and CAM (see sec. 5.4.1).

Within NASA, related work is in progress or is proposed at several Program Centers. An exhaustive search of such activities has not been possible in the limited time available, but several programs are especially noteworthy. The Jet Propulsion Laboratory has an active Advanced Development Laboratory (Bejczy, 1980). The Goddard Space Flight Center (GSFC) has proposed an effort to adapt existing CAD/CAM facilities at the Center to the control of robot manipulators for complete assembly (Purves, personal communication, 1980). Self-replicating systems have been studied at Marshall Space Flight Center (von Tiesenhausen and Darbro, 1980).

NASA unique benefits and requirements. NASA is in a unique position to benefit from massive automation – particularly self-replicating systems. The minimum possible size for a totally autonomous SRS is not presently known. However, feasibility studies performed to date (Freitas, 1980a; von Tiesenhausen and Darbro, 1980) have described systems which were quite large. Although autonomous self-replicating systems have been proposed for terrestrial use (Moore, 1956), sociocultural and ecological considerations seem to make them less practical, possibly even undesirable, on the Earth itself. This planet already supports several very large symbiotic man-machine replicating systems – the industrial societies.

In contrast to the terrestrial case, autonomous or symbiotic SRS are ideally suited to space applications. In space there is room for such systems to multiply and grow. In fact the exponentially expanding, self-replicating factory is the most promising option for economically viable exploration and utilization of space beyond the near-Earth environment. The bootstrapping effect of self-replication permits the utilization of vast quantities of extraterrestrial materials with only a modest initial investment of terrestrial materials.

SRS for space use must contend with an alien environment – vacuum or unusual atmospheres, zero to many gs of acceleration, radiation, temperature extremes, and so forth. Total autonomy will be more useful in space than on Earth. For symbiotic man-machine systems, man-rated life support systems are required, but because of the expense of

man-rated systems it is worthwhile pursuing totally autonomous systems for early exploratory ventures. Because humans need for many reasons, to go into space in person it will ultimately be necessary to develop the required life support systems.

Possible approaches to realization of SRS. The Replicating Systems Team envisions a three-pronged approach to achieving working self-replicating systems. First, NASA should inaugurate a “top-down” program, starting with a strawman mission and defining the hierarchy of required steps for achieving that mission. Second, NASA should initiate in-house and sponsored research on enabling technologies, a “bottom-up” approach. Participation in research will keep the agency involved at the leading edge of automation technology and allow new developments to be fed into the mission design of the top-down and other NASA programs in a timely manner. The third recommended line of attack is a “middle-out” near-term hardware feasibility demonstration which will provide a focus for NASA involvement in self-replicating systems. The recommended feasibility demonstration is at the threshold of present-day technology, is extendable in a bottom-up manner to systems of greater capability and complexity, and can be decomposed in a top-down fashion to proceed from a feasibility demonstration to the fully self-replicating systems.

The top-down approach suffers from the fundamental impossibility of conceptualizing at the outset, in such an alien field of endeavor, just what the final system should be like. The bottom-up approach suffers from a lack of focus for driving it toward useful, realizable goals. Both approaches have merit and should be pursued, especially in the long run. But in the near term NASA should follow the middle-out approach and perform a feasibility demonstration which will strain the present state-of-the-art in robotics, gain NASA experience, and establish a NASA presence in state-of-the-art machine intelligence and robotics technology.

The feasibility demonstration has been conceived, however, to have three other benefits. First, when successful, it may have regenerative impact on U.S. productivity by, for example, helping to decrease the cost of robot manipulators. Second, the insights gained in performing the feasibility demonstration will be valuable in formulating a top-down mission plan for achieving extraterrestrial SRS, and in identifying valuable areas for future fundamental research and development. Third, NASA can start at the demonstration level and begin to work progressively upward toward a generalized autonomous replicating factory.

5.6.2 Top-Down Approach

The top-down approach consists first of carefully defining the overall problem, then decomposing that problem

into simpler subproblems. These subproblems are, in turn, decomposed into sub-subproblems, and so on. The process continues, forming a lattice structure whose lowest tier nodes are low-level problems which are readily soluble.

Advantages and limitations. In established fields of endeavor, a top-down approach to mission and system design usually provides the most manageable solution, especially in exceedingly complex situations. Top-down structured programming in computer science is one example where this approach is beneficial. Computer software systems contain literally millions of instructions. They are, to date, mankind’s most complex artifacts. Self-replicating systems will contain very complex software, in addition to being the most complex autonomous mechanical systems ever devised. For this reason, it is recommended that NASA adopt a top-down approach to the design of actual missions which employ SRS.

The top-down approach works best when there is a well-established goal and a mature technology. At present it is not clear what mission employing SRS will be undertaken first. Neither is the technology mature. The mission ultimately chosen probably will depend to some extent on the outcome of basic research which has not yet been done.

Scenario for replicating systems development. To promote the achievement of self-replicating systems, NASA should identify one or more strawman missions which take advantage of self-replication. Then one of these missions should be thoroughly studied in a top-down manner.

It is recommended that the first mission to be extensively studied be a mission executed relatively close to Earth. This will minimize cost and permit human intervention if necessary. An orbiting self-replicating system or a lunar-based self-replicating system are obvious candidates. The lunar site is recommended because manufacturing engineers presently have more experience in designing industrial facilities for a planetary surface than for orbit. Traditional designs assume a surface for structural support, gravity, and maintenance of atmosphere. On the Moon only the atmosphere is absent; in orbit all three are absent.

It is recommended that the strawman mission be a Generalized Lunar Autonomous Replicating Manufacturing Facility (GLARMF). Preliminary feasibility studies of such a system have already been done (Freitas, 1980a; Freitas and Zachary, 1981; von Tiesenhausen and Darbro, 1980). The statement of work presented below is suggested for investigation of the feasibility of the strawman GLARMF mission, and is divided into five parts. All parts could be performed by one contractor; however, it would likely be beneficial to split up the work. Parts 1 and 2 probably could best be performed by university researchers, while parts 3 through 5 might be better accomplished by one of the major aerospace companies.

Part 1: Prepare a tutorial state-of-the-art technology assessment report on autonomous manufacturing. Consider computer-aided manufacturing (CAM), computer-aided design (CAD), robotics, machine intelligence, computer vision, "telepresence" (Minsky, 1979, 1980), and other relevant fields. Separately evaluate the state-of-the-art as it exists in laboratories and in industrial practice. Determine how the state-of-the-art has progressed over time in both laboratories and in industry. Extrapolate the past and the current state-of-the-art into the future to predict when it will be feasible to construct a Generalized Lunar Autonomous Replicating Manufacturing Facility similar to that described in recent publications (Freitas, 1980a; Freitas and Zachary, 1981; von Tiesenhausen and Darbro, 1980).

Part 2: Prepare a tutorial state-of-the-art technology assessment report on nonterrestrial manufacturing. Determine how the state-of-the-art has progressed over time, both in theory and in experiment. Extrapolate the past and current state-of-the-art into the future to predict when it will be feasible to construct a Generalized Lunar Autonomous Replicating Manufacturing Facility such as that described in recent publications (Freitas, 1980a; Freitas and Zachary, 1981; von Tiesenhausen and Darbro, 1980).

Part 3: Combine the results of the technology assessment reports resulting from Part 1 on autonomous manufacturing and Part 2 on nonterrestrial manufacturing. Perform a top-down mission design for a Generalized Lunar Autonomous Replicating Manufacturing Facility. Identify those elements of the Work Breakdown Structure (WBS) which are being pursued outside NASA, but which will require additional NASA support and direction in order to achieve NASA goals. Make recommendations on how NASA should interface with the ongoing work. Identify those elements of the WBS which are unique to NASA. Make recommendations on how NASA should approach these elements.

Part 4: Perform a feasibility study for a terrestrial technology verification demonstration of a Generalized Autonomous Replicating Manufacturing Facility. Recommend one or more suitable demonstration sites. Determine what NASA in particular and the United States in general could use the facility for after the demonstration is completed. Include schedule and cost estimates (in constant dollars and real year dollars).

Part 5: Perform a feasibility study for a Generalized Lunar Automated Replicating Manufacturing Facility. Recommend one or more candidate lunar sites. Consider the construction of habitation modules and agricultural modules as output products. Compare the cost and schedule of achieving a lunar base by the use

of (a) terrestrial manufacturing, (b) lunar manufacturing without replication of production facilities, and (c) lunar manufacturing with replication of production facilities. Cost estimates should be in constant dollars and real year dollars.

A few suggested sources for obtaining studies of the GLARMF are listed in table 5.6.

5.6.3 Bottom-Up Approach

The bottom-up approach consists of supporting basic and applied fields related to the desired goal. Science and technology normally advance in a bottom-up fashion.

TABLE 5.6. SUGGESTED SOURCES FOR GLARMF DEVELOPMENT STUDIES

Sources	Activity
University	
Stanford University	established AI lab
Carnegie-Mellon University	established AI lab
Massachusetts Institute of Technology	established AI lab
University of Michigan	interest in robotics because of proximity to Detroit
University of Maryland	AI researchers, near NASA HQ and GSFC
Industrial	
SRI International	established AI facility
General Electric	engaged in industrial automation, space-craft manufacturer
Westinghouse	engaged in industrial automation
Hughes Aircraft	spacecraft manufacturer
Ford Aerospace	CAD/CAM facilities
RCA	spacecraft manufacturer
Martin Marietta	spacecraft manufacturer
Fairchild	spacecraft manufacturer recently hired several SRI expatriates
TRW	spacecraft manufacturer
Texas Instruments	rumored to have extensive internal automation work
IBM	rumored to have extensive internal automation work
RAND	operations research
A. D. Little	operations research
Bolt, Beranek & Newman	operations research

Researchers build on the work of their predecessors. At any given time the problems which are soluble and present research prospects are defined by previous research which has been done and by the supporting technology which is currently available. Inventions and breakthroughs are notoriously hard to schedule in advance. It is worthwhile noting that *Homo sapiens*, an example of an autonomous replicating manufacturing facility, was developed in a bottom-up fashion by the process of evolution.

Advantages and limitations. Occasionally, difficult goals are achieved by a concerted, directed effort. One example was sending a man to the Moon and returning him safely to Earth. Another was the Manhattan Project which produced the first atomic bomb. This approach works when the goal is clearly identified and one can determine how to achieve it. However, significant progress in science and technology is frequently made on the basis of research performed on an *ad hoc* speculative basis because someone is actively interested in doing that research. One of the greatest assets a nation has is the creativity and intuition of people who have devoted their lives to developing those qualities.

The top-down approach works well only when the relevant bottom-up "homework" has been done in advance. Rocketry and nuclear physics research existed long before the United States committed itself to sending a man to the Moon or developing the atomic bomb. Two good examples of how advancing technology (which was not planned to be available when the mission was designed) enhanced a mission are the high-quality TV system and the lunar rover used toward the end of the Apollo program. When people have good ideas, there should be resources available to bring those ideas to fruition.

The bottom-up approach suffers from several deficiencies. Since it is somewhat speculative in nature, some of the research will turn out to be of little use to the sponsor, though spinoffs to other fields may occur. Since bottom-up research is proposed on an *ad hoc* basis, careful selection is required to ensure a clear sense of direction toward the desired goal. Also, there can be some duplication of effort.

Scenario for research and development. Limitations notwithstanding, bottom-up basic and applied research is necessary to the achievement of vital and imaginative programs. Accordingly, it is recommended that NASA support moderate amounts of basic and applied research showing promise in helping to achieve NASA's goals. The mechanism that has worked fairly well (though known to have some flaws) is the publication of an Announcement of Opportunity (AO) soliciting proposals for research. These proposals are subjected to peer review, and competent ones which show some promise of payoff for NASA are funded. It is recommended that a similar mechanism be used to ensure that new ideas are factored into the mission of achieving autonomous replicative manufacturing. Other-

wise, as pointed out in a recent study, unequivocal early commitment to a particular mission scenario and technology during top-down mission design will result in a mission which is using obsolete technology when it finally becomes operational.

A sample Announcement of Opportunity (AO) for SRS-related basic and applied research supportive of the development of SRS technology is presented in table 5.7. It is recommended that the AO be given wide dissemination. This will allow NASA to ferret out those organizations and individuals of various persuasions, backgrounds, and in different locations who have done related research or are seriously interested in doing new research in these areas. The NASA personnel who evaluate the proposals will develop an excellent in-depth perception of the current state-of-the-art in the areas covered by the AO. This knowledge will prove invaluable when fed back to the top-down and middle-out programs.

It is recommended that the AO be distributed nationwide to the departments of industrial engineering, electrical engineering, mechanical engineering, computer science, mathematics, physics, astronomy, business, philosophy, law, and economics in colleges and universities. It is further recommended that the AO be announced in professional publications such as *IEEE Spectrum*; *IEEE Computer*; *IEEE Transactions on Systems, Cybernetics, and Society*; Communications of the ACM; AAAI (American Association for Artificial Intelligence) publications; SME (Society of Manufacturing Engineers) publications; *Robotics Age*; *Industrial Robots International*; *Science*; *Science News*; *Byte*, etc.

5.6.4 Middle-Out Approach

The recommended middle-out approach consists of three stages. Briefly, in stage 1 a technology feasibility demonstration of a rudimentary self-replicating system is performed. In stage 2, stage 1 is further refined in a top-down manner to produce a less rudimentary system which operates in a less structured environment. Stage 3 consists of starting at stage 1 and doing a bottom-up synthesis of a more complex SRS.

The self-replicating system envisioned for stage 1 is a computer connected to one or more manipulators. Under control of the computer, the manipulator(s) will assemble another computer and another set of manipulator(s) from well-defined subassemblies. Examples of these subassemblies are printed circuit cards for the computer and individual joints or limb sections for the manipulator(s). This approach to self-replication is inspired by the von Neumann "kinematic model" as described in section 5.2.

In stage 2, the subassemblies would begin to be assembled from still smaller sub-subassemblies such as integrated circuits, resistors, motors, bearings, shafts, and gears. This stage can proceed for quite some time as the techniques for

TABLE 5.7.— A SAMPLE ANNOUNCEMENT OF OPPORTUNITY FOR SRS-RELATED BASIC AND APPLIED RESEARCH

<p>NASA is interested in creating a Generalized Lunar Autonomous Replicating Manufacturing Facility along the lines described in "Advanced Automation for Space Missions: The Report of a 1980 NASA-ASEE Summer Faculty Workshop." Accordingly, proposals are solicited for basic and applied research in the following or related areas:</p> <ol style="list-style-type: none"> 1. Computer-aided design (CAD) 2. Computer-aided manufacturing (CAM) 3. Robotics 4. Machine intelligence 5. Telepresence/telefactors 6. Man-machine systems and interfaces 7. Computer vision 8. Robust systems (mechanical, electrical and organizational fault tolerance) 9. Organization of large-scale systems 10. Analysis of large-scale systems 11. Command and control of large-scale systems 12. Consciousness, goal-directed behavior, and free will in large-scale systems 13. Extraterrestrial resource extraction 14. Extraterrestrial resource utilization 15. Extraterrestrial materials processing 16. Extraterrestrial manufacturing processes 17. Industrial automation 18. Social, philosophical, and legal implications of a Generalized Lunar Autonomous Replicating Manufacturing Facility 19. Space industrialization 20. Orbiting and lunar settlements and colonies 21. Space and Earth science from a lunar base 22. Applications of a Generalized Lunar Autonomous Replicating Manufacturing Facility 23. Interplanetary and interstellar exploration from a lunar base 24. Exports from the Moon and to the Earth <p>Research may be proposed by academic groups, industrial groups, NASA groups, or private individuals. Research may be performed at universities, at industrial facilities, at NASA facilities, or at private facilities. Proposals should describe the qualifications of the institution and/or individuals who propose to perform the work. They should describe other work, especially that directly related to this AO.</p> <p>Proposals should describe the work to be done, the schedule envisioned, intermediate milestones, planned reports and publications, and research required. Proposals should stress the relevance of the research to construct a Generalized Lunar Autonomous Replicating Manufacturing Facility. Priority will be given to those proposals which fulfill unique NASA requirements and propose work which is not likely to be funded by other entities, such as the Department of Commerce, the Department of Defense, the National Science Foundation, or private industry.</p>

assembling each subassembly from sub-subassemblies are developed and implemented one by one. By the time stage 2 is complete, there will be extensive cross-fertilization taking place between industry and the feasibility demonstration. Indeed, accomplishment of stage 2 will mean that robots can be assembled from parts by other robots. As discussed in sections 5.4 and 5.5, this will have a profound impact on U.S. industry.

Stage 3 is the final link in achieving an autonomous self-replicating manufacturing facility. In stage 3 the manipulators, which have, in stages 1 and 2, been assembling more robots, are used to build the machines which make the

parts. For example, the manipulators could assemble a printed circuit board manufacturing machine or a gear manufacturing machine. The problem of closure, discussed at length in section 5.3.6, becomes a major practical issue at this point. One must be careful that as one adds more and more machines the total number of different parts required is eventually produced by the total population of machines.

Advantages. The middle-out approach has a number of important advantages. In the long run it will replace neither the top-down nor the bottom-up methodologies. It does, however, provide a place to start on the practical realization of SRS.

The middle-out approach begins with the feasibility demonstration and then proceeds in a top-down and a bottom-up fashion. The feasibility demonstration alone will produce useful output — the automated production of robots. The expenditure required for the feasibility demonstration is tiny compared to the expenditure required before either the top-down or the bottom-up approach begins to show useful output. The middle-out approach can then be continued at whatever level of support seems appropriate and will produce useful spinoffs for industry as it progresses.

One of the chief advantages to NASA of the feasibility demonstration is that it can begin immediately. Working on the feasibility demonstration will provide NASA with valuable insights into practical problems associated with self-replicating systems. These insights will greatly increase the efficiency with which NASA can pursue both the top-down and the bottom-up approaches. The feasibility demonstration will be a valuable learning tool for both NASA and the industrial community.

As has been previously stated in this report, achievement of robot production of robots will decrease the cost of robots. This will directly benefit U.S. productivity and indirectly benefit NASA by lowering the cost of manufactured goods. Another valuable characteristic of the feasibility demonstration is that it will produce a visible output — a functioning autonomous self-replicating system (albeit a rudimentary one). In a field which is as foreign to most people as autonomous SRS, this will lend valuable credibility to the plans to produce more complex autonomous systems in space.

Limitations. The chief limitation of the middle-out approach is that it will not, of itself, produce an autonomous self-replicating system suitable for NASA's needs in space. The direction provided by the top-down approach is also needed. Also the creativity of the bottom-up approach is necessary to provide the needed adaptations to the space environment, such as designs and processes optimized for the use of extraterrestrial materials. Another disadvantage of the middle-out approach is that it will consume resources which could otherwise be devoted to the top-down and bottom-up methodologies. However, the overall efficiency should be greatest if a balance is maintained among all three approaches.

As simple as it sounds, the team estimates, on the basis of its discussions with industry and research community representatives, that it would require about 5 years and \$5-50 million (1980 dollars) to accomplish the feasibility demonstration proposed below. The major difficulties include the following:

- Assembly by robot is a difficult task at present, and final assembly is one of the more difficult forms of assembly.

- Present-day robot manipulators are built using hand labor. They are not designed for easy automated assembly. American Robot Corporation is reported to be planning on the automated assembly of robots beginning in 1981 (Industrial Robots International, 1980). However, these robot manipulators are quite small (5 lb load capacity), and "Gallaher's forecasts of small robot acceptance seem highly optimistic as do his own production plans and pricing." The Japanese have been far more aggressive in this area (IAF Conference, 1980).

- Present-day robot manipulators are rather weak for their weight. Care must be exercised to ensure that the subassemblies are light enough for the robot manipulators to be able to manipulate them — or, alternatively, to ensure that the robot manipulator is strong enough to be able to manipulate the subassemblies.

These problems are by no means insurmountable. However, considerable re-engineering of robot manipulators will be required to facilitate their assembly by similar robot manipulators. Likewise, the packaging of the computer will require some re-engineering for easy assembly by a robot manipulator.

Scenario for replicating systems demonstration. We now present a more detailed description of the proposed demonstration scenario for SRS. The demonstration begins with a parts depot stocked with enough subassemblies for the production of two robot manipulators and their associated computer systems. One complete, operating robot, Robot 1, is also present. It will construct Robot 2 which will, in turn, construct Robot 3, thus passing the "Fertility Test" (sec. 5.3.3). This arrangement is shown schematically in figure 5.28.

Robot 1 begins its labors by obtaining, one at a time, the subassemblies for the base (which doubles as the electronics card cage assembly) of Robot 2 from the parts depot. Robot 1 assembles the base, computer, and servo controls for Robot 2. Then, one at a time, Robot 1 obtains the subassemblies for the manipulator arms of Robot 2 and constructs the arms of Robot 2 from them.

When Robot 2 has been completely assembled, Robot 1 plugs in the power cord of Robot 2. Robot 1 then obtains a blank diskette (a removable mass memory device for computers) from the parts depot, inserts the diskette into its own computer, copies its software onto the diskette, and then removes the diskette from its own computer. Reproduction is complete when Robot 1 turns on the power to Robot 2, inserts the diskette (which now has a copy of the operating software on it) into Robot 2's computer, and then pushes the start button on the computer. From then on, Robot 2 is autonomous.

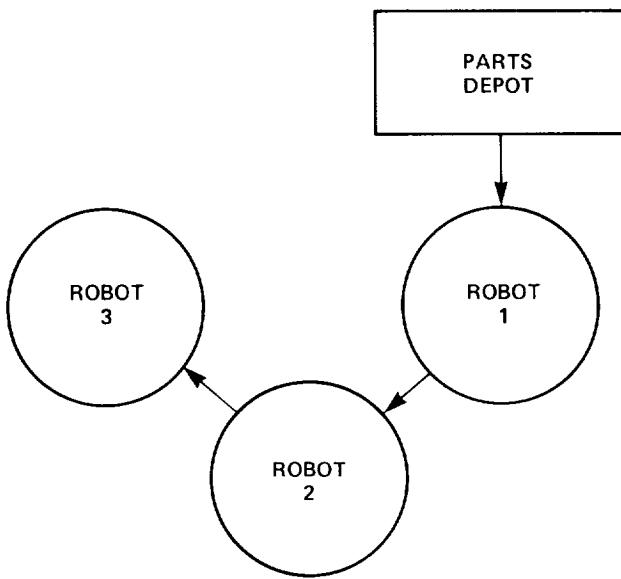


Figure 5.28. – Schematic of simple robot self-replication.

It should be noted that some additional complexity has been introduced into the demonstration by explicitly transferring the instructions from one generation of robot to the next by physical movement of a recording medium. This strategy was employed to make it clear that the generations are truly autonomous.

One of the ground rules of a demonstration such as this should be that all interaction between the robots be explicit and visible to a human observer. If the computers of the various robots were electrically interconnected the psychological impact on the observer would be more along the lines of a single system which was expanding itself, rather than producing distinct offspring. In addition, the demonstration as described should have an especially significant impact on anyone who has ever inserted a diskette into an inert computer and activated it by "booting it up."

The demonstration then proceeds by having Robot 2 construct and activate Robot 3. Robot 2 obtains the parts from Robot 1, who obtains them in turn from the parts depot and passes them along bucket-brigade style, according to its stored post-replication instructions. After Robot 3 is operational, the utility of the three robots can be shown by having each one of them construct some useful end product. Figure 5.29 is an artist's conception of the demonstration. Having accomplished the demonstration described above, it would be relatively easy to make it even more impressive by having each robot build two offspring as shown schematically in figure 5.30. At the end of the second generation (counting Robot 1 as 0th generation) there would be seven robots instead of only three.

As a culmination of the initial feasibility demonstration, each of these seven robots should then begin assembling useful output products. The demonstration should have a

much more profound impact on people who witness it directly than on those who see only a videotape or movie. It is expected, therefore, that the demonstration will be run repeatedly. To facilitate multiple demonstrations the robots can be ordered to disassemble one another and return the parts to the parts depot, by following their coded instructions in reverse. Care should be taken to ensure that each of the assembly operations is reversible so that disassembly is possible. Bolts should be used in preference to glue, welding, or rivets. Mechanical and electrical connections should be engineered to stand up under repeated connection and disconnection.

The demonstration, as thus envisioned, requires advances in state-of-the-art robot programming (Donata and Camera, 1980), as well as re-engineering of the mechanical and electrical subassemblies for easy assembly and disassembly. Appendix 5J gives a brief description of the complexity of the programming required in relation to the capabilities of a commercially available robot manipulator, the PUMA 500.

A Statement of Work for accomplishing the demonstration described above should include the following:

- (1) Design or select an autonomous robot system consisting of a computer and manipulator(s). Design or selection should be based on ease of assembly of the robot from subassemblies and the ability of the robot to do assembly work.
- (2) Partition the robot itself into subassemblies which it is capable of assembling into a complete robot.
- (3) If (2) cannot readily be done, use the knowledge gained in attempting (2) to redesign the robot to permit easier assembly by a similar robot.
- (4) Cycle through (2) and (3) until a satisfactory design has been achieved.
- (5) Produce or procure enough subassemblies for one complete robot.
- (6) Construct and test one robot.
- (7) While (2) through (6) are in progress, produce or procure a computer simulation of the robot.
- (8) Use the computer simulation to develop and test a general purpose assembly software compiler for the robot. The compiler should accept descriptions of subassemblies and generate detailed assembly instructions for the robot.
- (9) Input the subassembly descriptions from (4) to the compiler of (8) and verify with the simulation in (7) that the subassembly descriptions and robot are compatible.
- (10) Produce or procure enough subassemblies for six more complete robots plus spares of each subassembly.
- (11) Design the parts depot.
- (12) Produce or procure the parts depot.
- (13) Produce or procure the final software for the robots.
- (14) Perform Phase 1 of the demonstration, wherein Robot 1 constructs Robot 2 which in turn constructs Robot 3.

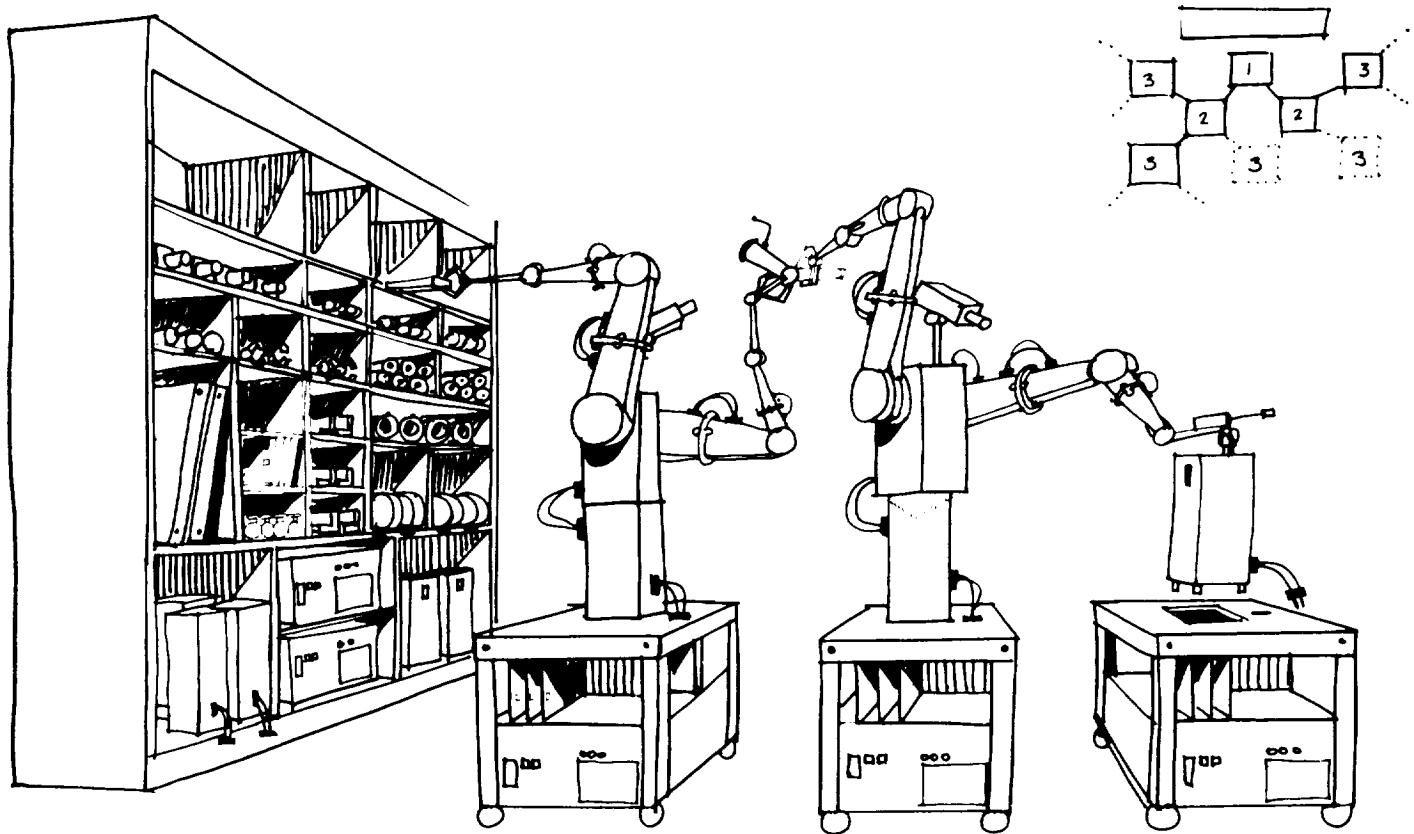


Figure 5.29.— Proposed demonstration of simple robot self-replication.

(15) Augment Phase 1 of the demonstration to produce Phase 2 wherein the three robots begin to produce useful products after they are assembled. Perform Phase 2.

(16) Augment Phase 1 of the demonstration to produce Phase 3 wherein each robot builds two copies of itself to demonstrate exponential replication. Perform Phase 3.

(17) Combine Phases 2 and 3 to produce a demonstration wherein the seven robots begin to make useful output products after they have been constructed.

It should be explicitly noted that the Statement of Work given above describes a *research* effort – the team discovered no fundamental problems which would prevent its successful accomplishment. There will be, however, many practical problems encountered along the way.

It is well known that in a research environment it is impossible to simultaneously constrain objectives, expenditures, and schedule. Therefore, it is recommended that funding be on a cost-plus-award-fee (CPAF) or cost-plus-incentive-fee (CPIF) basis with meaningful incentives for staying on schedule. Although firm fixed price (FFP) is in vogue in the government at the present time, this is definitely not desirable for an activity with the innate uncertainties of SRS research.

To the extent possible within budgetary constraints, the schedule should not be compromised if a capability for autonomous self-replicative manufacturing in space is to be achieved.

Suggested sources. The university or industrial sources listed as having AI or industrial automation capabilities could perform the demonstration. The National Bureau of Standards also has a robotics laboratory which could undertake the demonstration. The team, however, recommends that NASA give serious consideration to performing the demonstration in-house. This would allow the agency to breed a new generation of engineers, computer scientists, and managers with expertise in robotics. Competent people in this field are exceedingly difficult to recruit. According to industry and research community experts consulted during the study, the most limiting factor preventing faster development of robotics in the United States today is the inadequate supply of qualified practitioners. NASA will need a cadre of such people in the future to manage the implementation in space. The demonstration, performed in-house, will provide an interesting, challenging, educational environment which should permit NASA to attract and retain the kinds of people it will need for the space program of the 21st century.

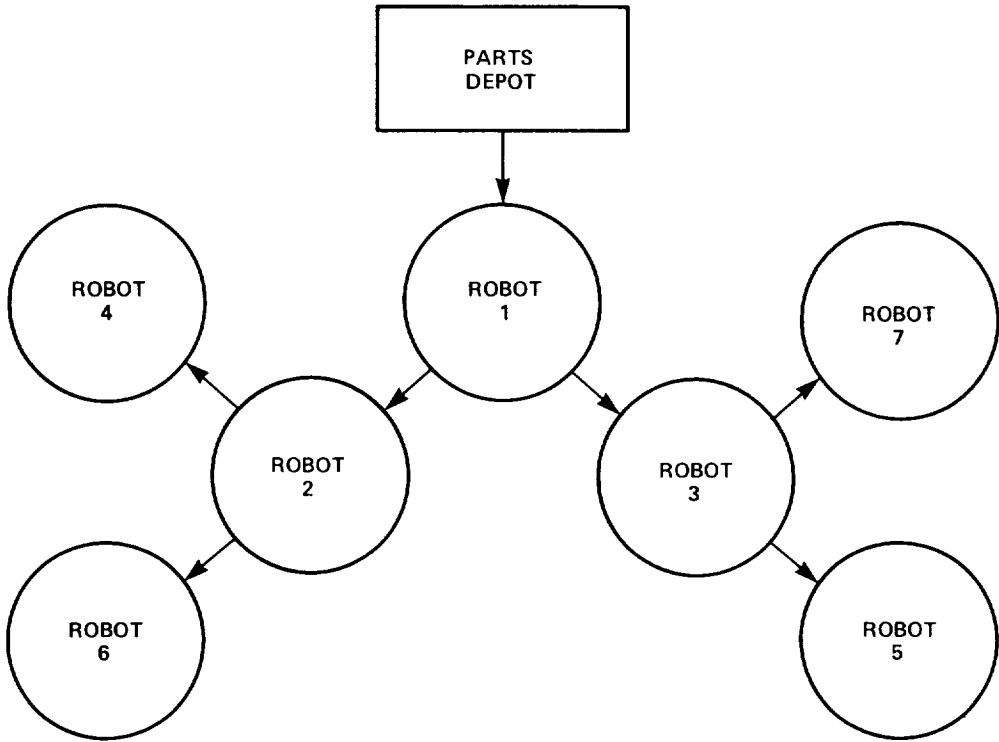


Figure 5.30.— Schematic of simple robot replication exponentiation.

5.6.5 Initiation of the Three Approaches

Section 5.6.1 proposed a Generalized Lunar Autonomous Replicating Manufacturing Facility as a strawman mission to bring NASA up to speed in advanced automation technology, in particular the technologies relevant to SRS. Succeeding discussions dealt with the top-down, bottom-up, and middle-out approaches to achieving such a facility. Figure 5.31 shows how the three approaches relate in achieving the overall program goals.

The various approaches, stages, and phases fit together in coordinating industry, university, and NASA expertise in the fields of self-replicating systems and extraterrestrial utilization of materials. Once the benefit of this expertise has been obtained, the mission design and realization of a GLARMF can begin. A proposed timeline for the development and demonstration of replicating systems may be found in figure 5.32.

It would be most advantageous for NASA to begin activities in advanced automation research and development at the present time. By beginning now, NASA will be in a strong position to seriously demonstrate and deploy advanced autonomous systems after the Shuttle becomes operational. At that time the Shuttle will not be making such large demands on the NASA budget and a means of

transporting the systems will exist. This will also be an opportune moment to begin attracting a cadre of bright, enthusiastic robotics practitioners by offering them both the chance to enter robotics as it begins to take off as a well defined field and the opportunity to contribute to the development of what may be one of mankind's most far-reaching achievements.

The 1980s may be the "Decade of the Robot." Many lay magazines such as *Newsweek* and *Business Week* have run major articles on robotics. Professional journals such as *Science* and *IEEE Computer* have also published prominent articles on robotics. In 1980 a new professional organization, the American Association for Artificial Intelligence, was created. Its first "Annual National Conference on Artificial Intelligence" was held at Stanford University during the Summer Study which produced this report. Momentum is gathering as the robotics and automation wave begins to take form. The team suggests that NASA ride the crest of this wave rather than stand back and be engulfed when it breaks.

NASA will be able to use the results of programs in automation sponsored by other government organizations; however, the space agency has some unique requirements which are unlikely to be met unless NASA takes an active role in automation research and development. Failing this,

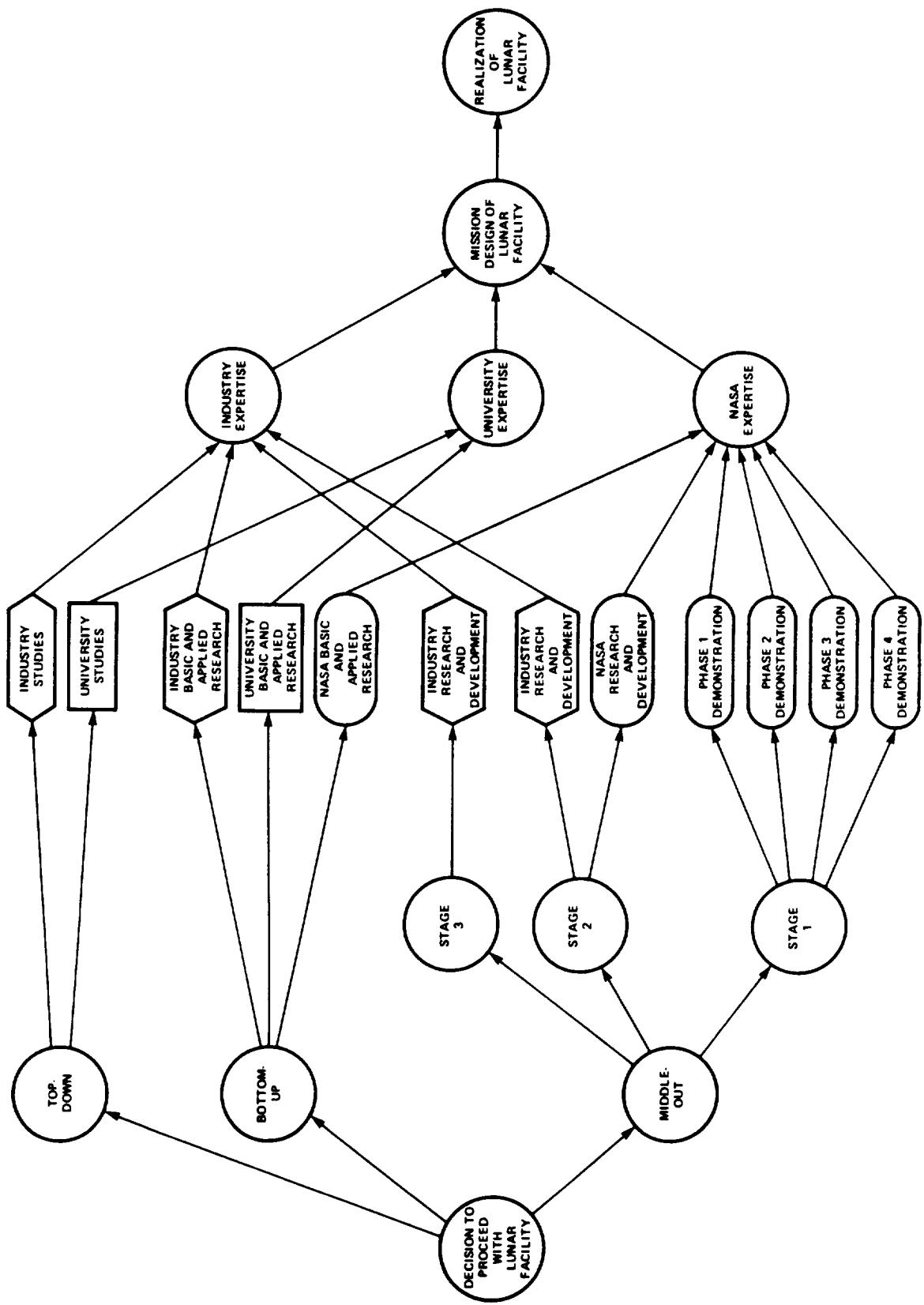


Figure 5.31.—Relationship of three R&D approaches to SRS development and demonstration.

YEAR →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. THEORY, CONCEPTS, PRELIMINARY DESIGN																				
2. SYSTEMS ANALYSES																				
3. TECHNOLOGY DEVELOPMENTS																				
4. TECHNOLOGY VERIFICATIONS																				
5. SUBSYSTEMS DEMONSTRATIONS																				
6. INDIVIDUAL DEMONSTRATION OF SYSTEMS																				
(a) UNIVERSAL CONSTRUCTION																				
(b) PRODUCTION FACILITY																				
(c) PARTS PRODUCTION																				
(d) MATERIALS PROCESSING																				
7. INCREMENTAL BUILD-UP DEMONSTRATIONS																				
8. OPERATIONAL READINESS																				
DEVELOPMENT AND CONSTRUCTION																				
INITIAL																				
FIRST DEMO																				

Figure 5.32.— Suggested timeline for development and demonstration of replicating systems technologies.

the infusion of relevant new technology into NASA programs can be expected to be slow. A recent report concluded that "NASA is 5 to 15 years behind the leading edge in computer science and technology" (Sagan, 1980). It was recommended that NASA take a more active role in computer science research and development to remedy the problem. The same phenomenon can be expected to occur with automated manufacturing. Unless NASA performs in-house R&D and sponsors university and industry R&D, significant infusion of the automation technology NASA needs in its future programs is unlikely. Many of NASA's unique needs cannot be satisfied unless the agency takes an active role in the development of automated manufacturing.

In particular, NASA is more concerned with total automation — the use of either teleoperated or completely autonomous systems than are most government agencies or industry. NASA should, therefore, perform or sponsor significant amounts of research and development in total automation. Special emphasis should be placed on the nonterrestrial environment, where such factors as vacuum or unusual atmospheres, nonterrestrial raw materials, and various gravity fields down to zero-g might be used to advantage (and must be dealt with in any case).

Replicative automation — the automation of automation — wherein robots are used to produce robots will happen in the terrestrial environment for economic reasons. There is, however, a synergism between replicative automation and total automation which has special relevance for

NASA. For operations such as lunar manufacturing or planetary terraforming exceedingly large amounts of mass must be manipulated in the extraterrestrial realm. Because of high transportation costs due to the Earth's gravitational influence, the most desirable method of achieving these missions is to send a "seed" — a replicative manufacturing facility with the minimum necessary closure for remote autonomous replication and repair — to distant operational sites. The seed can then produce, from *in situ* materials, and perhaps through several generations, the required machines to perform desired tasks. If the seed can manufacture propulsion systems and other seeds, then significant interstellar exploration becomes a very real possibility (Freitas, 1980a).

5.7 Conclusions and Recommendations

The Replicating Systems Concepts Team reached the following technical conclusions:

- The theoretical concept of machine duplication is well developed. There are several alternative strategies by which machine self-replication can be carried out in a practical engineering setting.
- There is also available a body of theoretical automation concepts in the realm of machine construction by machine, in machine inspection of machines, and machine repair of machines, which can be drawn upon to engineer practical machine systems capable of replication.

- An engineering demonstration project can be initiated immediately, to begin with simple replication of robot assembler by robot assembler from supplied parts, and proceeding in phased steps to full reproduction of a complete machine processing or factory system by another machine processing system, supplied, ultimately, only with raw materials.
- The raw materials of the lunar surface, and the materials processing techniques available in a lunar environment, are probably sufficient to support an automated lunar manufacturing facility capable of self-replication and growth.
- Tentative design of a lunar manufacturing facility capable of self-replication can begin, when current knowledge and state-of-the-art technologies are employed, but final design awaits the initial results of the demonstration-development program. Significant further research in lunar materials processing and in the design and operation of automated factories, should be conducted at once.

In addition, the team considers that the replicating systems concept, if implemented, can have the following important consequences:

- It will accelerate the design and development of sophisticated automated assembly techniques useful in carrying out future NASA missions.
- It will accelerate the design and development of improved automated assembly and processing techniques applicable to the problems of achieving increased Earth-based manufacturing productivity.
- By establishing an automated, growing, self-replicating, multipurpose, multiproduct lunar manufacturing facility, NASA capacity for space exploration and research can be enormously expanded and permanently enhanced with only modest continuing expenditures.

- The virtually cost-free expansion of mining, processing, and manufacturing capacity, once an initial investment is made in an autonomous SRS, makes possible the commercial utilization of the abundant energy and mineral resources of the Moon for the benefit of all mankind.
- The establishment of a replicating lunar manufacturing facility can be a stepping stone to the design and construction of replicating manufacturing complexes on the surfaces of other planets. These new complexes themselves may be products of automated, self-replicating manufacturing facilities located elsewhere.

Finally, the team offers the following general recommendations to NASA in furtherance of the basic objective of achieving practical self-replicating, growing machine systems in the shortest reasonable time:

- (1) NASA should begin immediately the development of a simple demonstration replicating system on a laboratory scale, with teleoperated to fully automated phased steps to higher levels of sophistication as the technology is proven and matures.
- (2) The space agency should support significant further research in lunar materials processing, lunar resource exploration, and the design and operation of automated manufacturing facilities.
- (3) NASA should implement the design, development, and construction of an automated, multiproduct, remotely reprogrammable lunar factory system to begin operation on the lunar surface early in the next century.
- (4) Studies should be conducted of scenarios in which a succession of replicating, multipurpose, multiproduct, automated, remotely reprogrammable factories could be placed in orbit or on other planets, these systems perhaps themselves products of earlier established nonterrestrial replicating facilities.
- (5) NASA should initiate additional studies of the social, political, military, and economic consequences of the proposed work, and of various other as yet unresolved issues and concepts (see app. 5K).

5.8 References

- Albus, James A.: *People's Capitalism: The Economics of the Robot Revolution*. New World Books, College Park, Maryland, 1976.
- Alvarez, L. W.; Alvarez, W.; Asaro, F.; and Michel, H. V.: Extraterrestrial Cause for the Cretaceous-Tertiary Extinction. *Science*, vol. 208, 6 June 1980, pp. 1095-1108.
- Arbib, Michael A.: Simple Self-Reproducing Universal Automata. *Information and Control*, vol. 9, 1966, pp. 177-189.
- Arnold, James R.; Criswell, David R.; and Waldron, Robert D.: Progress Report on Experimental Program to Develop HF Acid Leach Process for Refining Lunar Soils. Paper presented at the 5th Princeton/AIAA/SSI Conference on Space Manufacturing, 18-21 May 1981, Princeton, NJ.
- Asimov, Isaac: *I, Robot*. Doubleday and Company, N.Y., 1950.
- Banks, E. R.: Universality in Cellular Automata. Proc. of the 11th S.W.A.T., 1970, pp. 194-215.
- Bejczy, Antal K.: Sensors, Controls, and Man-Machine Interface for Advanced Teleoperation. *Science*, vol. 208, 20 June 1980, pp. 1327-1335.
- Bekey, Ivan; Naugle, John E.: Just Over the Horizon in Space. *Astronautics and Aeronautics*, vol. 18, May 1980, pp. 64-76.
- Bock, Edward: Space-Based Manufacturing from Lunar-Derived Materials. Convair/General Dynamics Contractor Report, 1979.
- Burger, J.; Brill, D.; and Machi, F.: Self-Reproducing Programs. *Byte*, vol. 5, June 1980.
- Burhoe, Ralph Wendell: What Specifies the Values of the Man-Made Man? *Zygon*, vol. 6, September 1971, pp. 224-246.
- Burks, Arthur W.: Von Neumann's Self-Reproducing Automata. In *Essays on Cellular Automata*, A. W. Burks, ed., University of Illinois Press, Urbana, Illinois, 1970, pp. 3-64.
- Business Week, 9 June 1980.
- Calder, Nigel: *Spaceships of the Mind*. Viking Press, New York, 1978.
- Carson, Gordon B.: *Production Handbook*. Ronald Press Company, New York, 1959.
- Cliff, Rodger A.: An Hierarchical System Architecture for Automated Design, Fabrication, and Repair. Paper presented at the 5th Princeton/AIAA/SSI Conference on Space Manufacturing, 18-21 May 1981, Princeton, NJ. (See also app. 5H.)
- Codd, E. F.: *Cellular Automata*. Academic Press, N.Y., 1968.
- Donata, G.; and Camera, A.: A High Level Programming Language for a New Multi-Arm Assembly Robot. First International Conference on Assembly Automation,
- Brighton, England, D.E.A. SPA, CSO, Torino, 21-27 March 1980.
- Dunning, Jeremy D.; and Snyder, Robert S.: Electrophoretic Separation of Lunar Soils in a Space Manufacturing Facility. Paper presented at the 5th Princeton/AIAA/SSI Conference on Space Manufacturing, 18-21 May 1981, Princeton, NJ.
- Dyson, Freeman J.: Search for Artificial Stellar Sources of Infrared Radiation. *Science*, vol. 131, 1959, pp. 1667-1668.
- Dyson, Freeman J.: *Disturbing the Universe*. Harper and Row Publishers, N.Y., 1979.
- Freitas, Robert A., Jr.: A Self-Reproducing Interstellar Probe. *J. of the British Interplanetary Soc.*, vol. 33, July 1980a, pp. 251-264.
- Freitas, Robert A., Jr.: *Xenology: An Introduction to the Scientific Study of Life, Intelligence, and Civilization in the Universe*. Xenology Research Institute, 1980b, 1400 pp.
- Freitas, Robert A. Jr.: Alien Sex. In *Xenology*, chapter 12; 1980c, 33 pp. Also *Analog*, vol. 102, May 1982.
- Freitas, Robert A., Jr.: Interstellar Probes: A New Approach to SETI. *J. of the British Interplanetary Soc.*, vol. 33, March 1980d, pp. 95-100.
- Freitas, Robert A., Jr.: Xenobiology, *Analog*, vol. 101, 30 March 1981, pp. 30-41.
- Freitas, Robert A., Jr.; Healy, T. J.; and Long, J. E.: Advanced Automation for Space Missions. Proc. of the 7th International Joint Conference on Artificial Intelligence, 24-28 August 1981, Vancouver, Canada.
- Freitas, Robert A., Jr.; and Valdes, Francisco: Search for Natural or Artificial Objects Located at the Earth-Moon Libration Points. *Icarus*, vol. 42, June 1980, pp. 442-447.
- Freitas, Robert A., Jr.; and Zachary, William B.: A Self-Replicating, Growing Lunar Factory. Paper presented at the 5th Princeton/AIAA/SSI Conference on Space Manufacturing, 18-21 May 1981, Princeton, NJ.
- Gardner, M.: Mathematical Games. *Scientific American*, vol. 224, February 1971, pp. 112-117.
- George, Frank H.: *Machine Takeover*. Pergamon Press, New York, 1977.
- Gettlemen, K. M., ed.: *Modern Machine Shop 1979 NC/CAM Guidebook*, 1979.
- Goeller, H. E., and Weinberg, A. M.: The Age of Substitutability. *Science*, vol. 191, 20 February 1976, pp. 683-689.
- Good, Irving John: Speculations Concerning the First Ultraintelligent Machine. *Advances in Computers*, vol. 6, 1965, pp. 31-88.
- Hamilton, W. D.: The Genetical Theory of Social Behavior (I and II). *J. Theoret. Biol.*, vol. 7, 1964, pp. 1-16, 17-52.
- Hay, Louise: Self-Reproducing Programs. *Creative Computing*, vol. 6, July 1980, pp. 134-136.

- Heer, Ewald: Prospects for Robots in Space. *Robotics Age*, vol. 1, Winter 1979, pp. 20-28.
- Heiserman, David L.: *Build Your Own Working Robot*. Tab Books, Blue Ridge Summit, PA, 1976.
- Hogan, James P.: *The Two Faces of Tomorrow*. Ballantine Books, N.Y., 1979.
- IAF Conference: *Robots To Produce Robots At Fujitsu*. From an English-language newspaper handout at the Conference, Tokyo, 21-28 September 1980.
- Ingel, L. Kh.: *Gravitational Focussing*. Soviet Astronomy, vol. 17, 1974, p. 836.
- Industrial Robots International, vol. 1, no. 5, August 1980, p. 2.
- Jacobson, Homer: *On Models of Reproduction*. *American Scientist*, vol. 46, 1958, pp. 255-284.
- Johnsen, Edwin G.: *Man, Teleoperators, and Robots: An Optimum Team for Space Exploration*. *J. Spacecraft and Rockets*, vol. 9, July 1972, pp. 554-556.
- Johnson, Richard D.; and Holbrow, Charles, eds.: *Space Settlements: A Design Study*. NASA SP-413, 1977. 185 pp.
- Kemeny, John G.: *Man Viewed as a Machine*. *Scientific American*, vol. 192, 1955, pp. 58-67.
- Laing, Richard A.: *Some Alternative Reproductive Strategies in Artificial Molecular Machines*. *J. Theoret. Biol.*, vol. 54, 1975, pp. 63-84.
- Laing, Richard A.: *Automation Introspection*. *J. Comp. System Sci.*, vol. 13, 1976, pp. 172-183.
- Laing, Richard A.: *Automaton Models of Reproduction by Self-Inspection*. *J. Theoret. Biol.*, vol. 66, 1977, pp. 437-456.
- Laing, Richard A.: *Automaton Self-Reference*. Dissertation, State Univ. of New York, Binghampton, Univ. Microfilms, Ann Arbor, MI, Publication No. TSZ 78-05, 542, 1978.
- Laing, Richard A.: *Machines as Organisms: An Exploration of the Relevance of Recent Results*. *BioSystems*, vol. 11, 1979, pp. 201-215.
- Lee, C. Y.: *A Turing Machine Which Prints Its Own Code Script*. *Mathematical Theory of Automata*. Polytechnic Press, Booklyn, N.Y., 1963, pp. 155-164.
- Martin, A. R., ed.: *Project Daedalus – The Final Report on the BIS Starship Study*, JBIS Supplement. Unwin Brothers Ltd., London, 1978.
- Matloff, Gregory L.: *Utilization of O'Neill's Model I Lagrange Point Colony as an Interstellar Ark*. *J. of the British Interplanetary Soc.*, vol. 29, December 1976, pp. 775-785.
- McCulloch, W. S.; and Pitts, W.: *A Logical Calculus of the Ideas Immanent in Nervous Activity*. *Bulletin of Mathematical Biophysics*, vol. 5, 1943, pp. 115-133.
- Miller, James G.: *Living Systems*. McGraw-Hill Book Company, N.Y., 1978.
- Miller, Rene H.; and Smith, David B. S.: *Extraterrestrial Processing and Manufacturing of Large Space Systems*, NASA CR-161293, vols. 1-3, September 1979.
- Minsky, M.: *Toward a Remotely-Manned Energy and Production Economy*. MIT Artificial Intelligence Laboratory, A.I. Memo No. 544, Boston, MA, September 1979. 19 pp.
- Minsky, M.: *Telepresence*. *Omni*, vol. 2, June 1980, pp. 44-52.
- Moore, Edward F.: *Artificial Living Plants*. *Scientific American*, vol. 195, October 1956, pp. 118-126.
- Morowitz, Harold J.: *A Model of Reproduction*. *American Scientist*, vol. 47, 1959, pp. 261-263.
- Myhill, J.: *The Abstract Theory of Self-Reproduction*. In *Essays on Cellular Automata*, A. W. Burks, ed., Univ. of Illinois Press, Urbana, Illinois, 1970, pp. 206-218.
- Numeric Control Equipment, *Aviation Week and Space Technology*, vol. 113, 29 December 1980, p. 249.
- Oliver, B. N.: *Proximity of Galactic Civilizations*. *Icarus*, vol. 25, June 1975, pp. 360-367.
- O'Neill, Gerard K.: *The Colonization of Space*. *Physics Today*, vol. 27, September 1974, pp. 32-40.
- O'Neill, Gerard K.: *The High Frontier: Human Colonies in Space*. William Morrow and Co., New York, 1976.
- O'Neill, Gerard K.; Driggers, G.; and O'Leary, B.: *New Routes to Manufacturing in Space*. *Astronautics and Aeronautics*, vol. 18, October 1980, pp. 46-51.
- Parkinson, R. C.: *Project Daedalus: Propellant Acquisition Techniques*. In *Project Daedalus – The Final Report on the BIS Starship Study*, JBIS Supplement, A. R. Martin, ed., Unwin Brothers Ltd., London, 1978, pp. S83-S89.
- Penrose, L. S.: *Self-Reproducing Machines*. *Scientific American*, vol. 200, June 1959, pp. 105-114.
- Phinney, W. C.; Criswell, D. R.; Drexler, E.; and Garmirian, J.: *Lunar Resources and Their Utilization*. In *Space-Based Manufacturing from Extraterrestrial Materials*, Gerard K. O'Neill, ed., AIAA Progress in Astronautics and Aeronautics Series, vol. 57, AIAA, New York, 1977, pp. 97-123.
- Preston, K., Jr.; Duff, M. J. B.; Levialdi, S.; Norgren, Philip E.; and Toriwaki, Jun-Ichiro: *Basics of Cellular Logic with Some Applications in Medical Image Processing*. Proc. of the IEEE, vol. 67, May 1979, pp. 826-856.
- Rao, D. Bhogeswara; Choudary, U. V.; Erstfeld, T. E.; Williams, R. J.; and Chang, Y. A.: *Extraction Processes for the Production of Aluminum, Titanium, Iron, Magnesium, and Oxygen from Nonterrestrial Sources*. In *Space Resources and Space Settlements*, J. Billingham, W. P. Gilbreath, B. O'Leary, and B. Gossett, eds., NASA SP-428, 1979, pp. 257-274.
- Sagan, Carl, Chmn.: *Machine Intelligence and Robotics: Report of the NASA Study Group*, NASA TM-82329, March 1980.

- Sagan, Carl; Drake, F. D.; Druyan, A.; Ferris, T.; Lomberg, J.; and Sagan, L. S.: *Murmurs of Earth: The Voyager Interstellar Record*. Ballantine Books, N.Y., 1978.
- Sandell, N. R., Jr.; Varaiya, P.; Athans, M.; and Safonov, M. G.: Survey of Decentralized Control Methods for Large Scale Systems. *IEEE Transactions on Automatic Control*, vol. AC-23, April 1978, pp. 108-128.
- Smith A. R., III: Cellular Automata and Formal Languages. *IEEE Conference Record of 11th Annual Symposium on Switching and Automata Theory*, Santa Monica, 28-30 Oct. 1970. N.Y., IEEE, 1970, pp. 216-224.
- Smith, J. Maynard: *The Evolution of Sex*. Cambridge Univ. Press, Cambridge, 1978.
- Stakem, Patrick H.: Life Versus Computer Capacity. *Byte*, vol. 4, February 1979, p. 58.
- Taneja, Vidya S.; and Walsh, Peter J.: Growth Considerations of a Self-Replicating Field Dependent on External Factors. *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-11, Aug. 1981, pp. 558-563.
- Thatcher, James: The Construction of a Self-Describing Turing Machine. *Mathematical Theory of Automata*. Polytechnic Press, Brooklyn, N.Y., 1963, pp. 165-171.
- Thatcher, James: Universality in the von Neumann Cellular Model. In *Essays on Cellular Automata*, A. W. Burks, ed., Univ. of Illinois Press, Urbana, Illinois, 1970, pp. 132-186.
- Trivers, R. L.: The Evolution of Reciprocal Altruism. *Quart. Rev. of Biology*, vol. 46, 1971, pp. 35-57.
- Turing, Alan: On Computable Numbers with an Application to the Entscheidungsproblem. *Proc. London Math. Soc.*, vol. 42, 1936, pp. 230-265.
- U.S. Bureau of the Census: Statistical Abstract of the United States, 98th Annual Ed., Washington, D.C., 1977.
- U.S. Bureau of the Census: Statistical Abstract of the United States, 99th Annual Ed., Washington, D.C., 1978.
- U.S. Bureau of the Mines: Metals, Minerals, and Fuels. In *Minerals Yearbook 1976*, vol. 1, GPO, Washington, D.C., 1978.
- Valdes, Francisco; and Freitas, Robert A., Jr.: Comparison of Reproducing and Nonreproducing Starprobe Strategies for Galactic Exploration. *J. British Interplanetary Soc.*, vol. 33, November 1980, pp. 402-408.
- Vondrak, Richard R.: Creation of an Artificial Atmosphere on the Moon. In *Advances in Engineering Science*, vol. 3, NASA CP-2001, 13th Annual Meeting of the Society of Engineering Science, sponsored by JIAFS, Hampton, Virginia, 1-3 November 1976, pp. 1215-1223.
- Vondrak, Richard R.: Environmental Impact of Space Manufacturing. In *Proceedings of the Third Princeton/AIAA Conference on Space Manufacturing Facilities*, Princeton, N.J., 9-12 May 1977, Paper no. 77-539.
- Von Neumann, John: The General and Logical Theory of Automata. In *Cerebral Mechanisms in Behavior — The Hixon Symposium*, L. A. Jeffress, ed., John Wiley & Sons, N.Y., 1951, pp. 1-31. Also in *Collected Works*, vol. 5, pp. 288-328.
- Von Neumann, John: Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components. In *Automata Studies*, C. E. Shannon, J. McCarthy, eds., Princeton Univ. Press, Princeton, N.J., 1956, pp. 43-98. Also in *Collected Works*, vol. 5, pp. 329-378.
- Von Neumann, John: Theory of Self-Reproducing Automata. Edited and completed by A. W. Burks. University of Illinois Press, Urbana, Illinois, 1966.
- Von Tiesenhausen, Georg; and Darbro, Wesley A.: Self-Replicating Systems — A Systems Engineering Approach. *NASA TM-78304*, July 1980.
- Waldron, R. D.; Erstfeld, T. E.; and Criswell, D. R.: The Role of Chemical Engineering in Space Manufacturing. *Chemical Engineering*, vol. 86, 12 February 1979, pp. 80-94.
- Wesley, J. P.: Ecophysics: The Application of Physics to Biology. Charles C. Thomas, Publ., Springfield, Illinois, 1974.
- Williams, R. J.; McKay, D. S.; Giles, D.; and Bunch, T. E.: Mining and Beneficiation of Lunar Ores. In *Space Resources and Space Settlements*, J. Billingham, W. Gilbreath, B. O'Leary, and B. Gossett, eds., NASA SP-428, 1979, pp. 275-288.
- Zachary, William B.: A Feasibility Study on the Fabrication of Integrated Circuits and Other Electronic Components in Space Utilizing Lunar Materials. Paper presented at the 5th Princeton/AIAA/SSI Conference on Space Manufacturing, 18-21 May 1981, Princeton, NJ.

APPENDIX 5A

FIRST ATTEMPT TO DEFINE A SELF-REPLICATING SYSTEM (A personal note contributed by W. E. Bradley, June 1980)

At a recent meeting a member of the NASA Advisory Council expressed excitement at the positive conclusions reached by the June 1979 Woods Hole Symposium concerning self-replicating mechanical systems. He said that he could not understand why a subject of such interest and importance to the exploration and utilization of space should be approached so timidly. Earnestly, he added:

"After all, a lathe can produce a lathe, properly operated; nowadays numerically controlled lathes are available; so why not program one to reproduce itself?"

My reaction was the following:

- A lathe cannot produce another lathe without many added subsystems (e.g., driving motor, tool grinder, tool bit production, etc.).
- Some contemplation of the self-replicating system problems at the practical engineering level has been undertaken by a few individuals in the past few months. This work is incomplete as yet, but is aimed at practical, demonstrable systems with only a few critical parts supplied from outside the system, including energy and raw materials for device fabrication. Energy and raw materials appear here in the role of "nutrient," the supply necessarily increasing as the system grows.
- The self-replicating system is indeed of great interest on fundamental grounds.
- The subject is appropriate to and important to NASA.

The work of the past few months (prior to the present study) relevant to self-replicating systems (SRS) is incomplete but has brought to light some principles and ideas of interest.

5A.1 Preliminary Investigation of the Self-Replicating Machine Shop

The town of Muncy is located in a somewhat remote part of central Pennsylvania. It is remarkable because of a nearly self-sufficient machine manufacturing capability in the Sprout-Waldron Company (now a division of another corporation, and therefore subject to change without notice). This company for many years has manufactured agricultural and food-processing equipment as well as heavy

machinery for the paper industry, especially pulp grinders. I became acquainted with them while searching for machines able to produce dense pellets for use as solid fuel from agricultural cellulosic wastes.

In the course of my visit, I was shown an excellent machine shop, a foundry, a woodworking shop, and a factory assembly space in which their machines were put together, painted and tested. They also had complete drafting and design engineering facilities. Of special interest was their toolmaking and repair shop, with which all of the milling machines, lathes, jig borers, punch presses, and so forth were kept in fine working order.

This complex, with the possible exception of the foundry, seemed to be a system which, with human assistance, could duplicate itself. In retrospect, it seems worthwhile to explore the possibility that the human operators might be replaced by general purpose automata, manufactured almost completely by the complex itself. The result would then be a major component of a self-replicating system. To complete the system would require manufacture of a prime power source which could be expanded as the complex grows, manufacture of a shelter system (sheds with roofs, walls, windows, and doors) similarly expandable, and possibly a casting and/or forging subsystem, and electronic and computer components of the automata. The foundry with its requirement for refractory furnace linings and high temperatures is a special problem and in some versions of the system may be bypassed.

Present machine shops. Each machine in a machine shop has a functional domain or "scope," assuming unlimited operator attention and guidance. Thus, a lathe (with no attachments) is able to produce objects with cylindrical symmetry having axial length and maximum diameter determined by the "bed length" and the "swing" of the machine. It can also make threads (helical structures), and, to a limited extent, can also make straight-line cuts or grooves which are more properly the work of a milling machine. Lathes can drill holes most readily on the axis of a workpiece of cylindrical symmetry and can achieve a high degree of accuracy of concentricity for this one type of drilling. Most drilling, however, is best accomplished on a jig borer.

The second major machine type in a shop is some form of drill press, or, better, a jig borer. The workpiece is held

firmly in an accurately translatable and rotatable fixture, remaining stationary while holes are drilled by a drill or boring tool held in a chuck rotating about the principal axis of the machine. Such a device can produce clusters of accurately located holes with parallel axes.

The third important shop component is the milling machine. The workpiece is clamped firmly to an accurately controlled table. The workpiece moves continuously, slowly, during operations while the rotating milling cutter shaves or saws the surface being worked. The milling machine is usually used to make rectilinear cuts to form accurately related plane surfaces or grooves.

A well-equipped machine shop usually also includes a power hacksaw, a powerful press with forming dies for forming sheet metal and for punching holes with "punch and die" sets, a bending brake, tool grinders, and possibly a surface grinder to be used like a milling machine to produce flat surfaces.

Self-replicating shop and universal machines. Each machine or subsystem of such a shop can be separated into parts from which it can be reassembled. Each machine therefore has a "parts list," and each part either can or cannot be fabricated by the set of machines and subsystems comprising the shop. The criterion for replication thus may be stated as follows:

If all parts of all machines and subsystems can be fabricated within the shop, then if properly operated the entire shop can be replicated.

"Proper operation" in this context includes supplying raw materials, energy, and manipulatory instructions or actions necessary to carry out the large number of machine operations, parts storage, and parts assembly required. Human labor is now used for these functions.

It is not necessary that the shop be able to produce anything except a replica of itself which is in turn capable of producing another. Therefore, some simplifications appear possible, such as standardization and limitation of scope where feasible. For example, a universal machine can be imagined with a wider cross feed table than a conventional lathe and with a standardized vise and tool holder so that it can be used for milling. All three dimensions of translation and one axis of rotation could be provided on the table. The head stock could be arranged to hold workpieces, milling cutters or drills. Hardened tools for the necessary cutting operations could be fabricated by the machine from carbon steel in the annealed condition, then tempered, drawn, and sharpened by a separate simpler machine including a small furnace and a tool grinding wheel equipped with tool-holder and feeds. By careful standardization of parts, tools, and fixtures, it is conceivable that such a "one-machine shop" could succeed in reproducing itself.

"Factons." After a shop had been tested with human operators and proven capable of self-replication, it would

be possible to explore the replacement of the human operators by mobile computer-controlled manipulators, or "factons." Hopefully, all of the "numerical control" features could be contained in these general-purpose programmable devices which could handle the machines like a human operator. The factons would transfer work from operation to operation, adjust the machine, perform each operation, then transfer the work to a parts storage array. Finally, the parts would be assembled by the factons and the entire shop set up in a selected location and floor-plan. The facton itself has a parts list, most designed to be manufacturable by the shop. Here it is practically inevitable that computer chips plus enormous memories will be needed which would fall outside the scope of the shop thus far envisioned. In other words most, but not all, of facton components could be fabricated by them in the shop. Still, given these extra components provided from outside, the factons could probably fully assemble themselves. The shop itself would require some exogenous elements, as noted above. Prime power, shaft power transmission such as belting or electric motors, abrasives, furnace heating arrangements for tool heat treatment, raw material such as basic feedstock including steel rods, strips, and plates are among the most obvious.

Using the same facton design, it should be possible to implement extensions of the shop, including an optical shop, a pneumatic and/or hydraulic equipment manufacturing shop, and ultimately even an integrated circuit shop. Note, however, that only the original shop with its factons and their programs would have to possess the capability for self-replication.

Computer components, probably provided from outside the system, might be furnished in an unprogrammed condition. Thus, factons would program the tapes, discs, or read-only memories by replication (and verification) of their existing programs. This procedure allows for the possibilities of "heritable" changes of program embodying "devolution" (simplification) or "evolution" (capability augmentation) by orderly program amendment.

5A.2 Program Extension Beyond Self-Replication

The "scope" of a self-replicating shop is much larger than is required for self-replication. Apparently the ability to replicate utilizes only a vanishingly small fraction of total capabilities (to produce various sizes and shapes of parts and to assemble them into machines and structures). The essential characteristic for self-replication is that the scope must be adequate to produce every part of every machine in the shop by means of a feasible program. This "closure condition" can be satisfied using only a small part of the shop's full capabilities.

A generic self-replicating shop can therefore, by means of a simple addition to its program, manufacture other machines and structures and, by means of them, interact

with its environment. For example, it can construct and operate foraging systems to procure fuel or materials, waste disposal systems, or transporters to carry replica shops to other locations.

Obviously, self-replication of such an extended system requires replication of the program-memory. This memory can be partitioned into two parts: (1) The self-replication process memory, and (2) the external process (manufacturing) memory. The distinction between these two memories is that the first is required to reproduce the basic unit (shop machines plus factons) while the second memory contains the program to produce process equipment not essential to the self-replicating nucleus.

At this point it is clear that the effect of a self-replicating system on its environment may take many forms dependent on the external process program. Using such a program, the scope of the system can be extended by construction of machines and structures capable of producing complex subsystems including mineral processing plants, solar energy power supplies, etc.

All of these extended self-replicating systems would embody the same basic nucleus of machines, factons and self-replication programming. They would differ only by addition of the external process program segment peculiar to each type.

Reliability and redundancy. Reliability is a primary concern, especially in the case of self-replicating processes. Two ideas are most important here.

First, the self-replicating program accuracy can be verified by comparison with other replicas of the same program. If a discrepancy is found between two self-replicating programs, a third or fourth replica can be consulted and the error pinpointed and corrected. The test of correctness is the ability to self-replicate.

Second, machines tend to wear, and ultimately to fail, from excessive use. On the other hand, if the system can replicate itself it can make spare parts and install them itself. A special program segment, the "maintenance program," should be devised to check machine wear and perform repairs as needed. This segment would be part of the self-replication program, although another somewhat similar maintenance program should probably be used to care for machines and structures of the external process. This external maintenance program would be specialized for each extended system and is properly part of the second memory.

Speciation. Any self-replicating system is actually another species of SRS, the species being dependent upon the contents of the second memory.

A group of interacting extended self-replicating systems may form a still larger self-reproducing system with yet more complex capabilities. It is not immediately apparent what factors limit the possibilities of such systems. Separable subsystems manufactured by a self-replicating shop may be machines of considerable complexity, themselves incapable of self-replication. Their supply is therefore dependent on the self-replication shop and its program.

It is interesting to note that a facton equipped with an aberrant program may function like a virus, visiting a self-replicating shop and using its machines for reproduction of its own type without constructing any other machine. It could then replicate its program for installation in the new "virus facton" and reproduce this way, using materials and energy from a host self-replication shop. This possibility opens up a large field of problems related to the security of self-replication systems from facton defect or infection.

APPENDIX 5B

LMF POSITIONAL TRANSPONDER SYSTEM

According to the baseline mission for a growing, self-replicating Lunar Manufacturing Facility (LMF) presented in section 5.3.4, a 100-ton seed is dropped to the lunar surface and thereupon unpacks itself, sets up the initial factory complex, and then proceeds to produce more of itself (or any other desired output). Clearly, the level of automation and machine intelligence required lies beyond current state-of-the-art, though not beyond the projected state-of-the-art two or three decades hence. Because of the already challenging design problem, it is highly desirable to keep all seed systems as simple as possible in both structure and function. This should help reduce the risk of partial or total system failure and make closure less difficult to achieve at all levels.

One of the more complicated pieces of hardware from the AI standpoint is the “camera eyes” and pattern recognition routines (visual sensing) that may be needed. Although it is possible that standardized robot camera eyes may be developed, it is more likely that each particular application will demand its own unique set of requirements, thus greatly reducing or eliminating any gains in simplicity of camera design. The pragmatic industrial approach (Kincaid et al., 1980) and design philosophy in these cases, especially in the area of computer vision, is to: (1) simplify, (2) use unconventional solutions, and (3) “cheat” (i.e., solve another problem). It may be that the best way to handle the problem of computer vision is to find a way to largely avoid it altogether.

When the seed unpacks itself it opens into a rather wild environment full of hills, bumps, ledges, crevasses, boulders, craters, and rocks. Surface navigation by mobile robots will be a serious challenge to AI technology. How will a machine know where it is, what the terrain ahead may be like, or how to get home? Laser tracking is one possibility, but probably too complicated when out of line of sight. Pattern recognition of geological and geographical landmarks is another possibility, but there are at least three serious deficiencies associated with this solution. First, the pattern recognition routines must be extremely sophisticated and the sensor very high in resolution and in the ranges of illumination that may be accommodated. Second, to recall how to get home after a lengthy perambulation across the lunar surface may require vast amounts of onboard computer memory. Every turn, every detour, every move the robot makes must be recorded, analyzed for spatial displacement geometry, and the present-position

pointer augmented against the stored features maps and correlated with the geographic images received through the vision sensors to plot the shortest route home to avoid the inefficiency of retracing the original physical path. Third, since exploration, development, and construction operations are always in progress around the site, each robot would need a memory capacity sufficient to recall in detail all changes in the landscape between the last series of explorations and the present one — the view is always changing. It may not be practical to design this much AI into each mobile robot, nor to require the central computer to exercise full teleoperator control of a large fleet of nonautonomous mobile robots.

SB.1 The Transponder Network

One way to achieve accurate positioning of all mobile robots while retaining their navigational autonomy is to employ a transponder system operating in the gigahertz frequency range. Much like the LORAN and NAVSTAR systems on Earth, these radar beacons would permit the accurate determination of position by simple triangulation for mobile robot devices located anywhere in the vicinity of the seed. A frequency of perhaps 30 GHz, easily within the range of current technology, would be required for 1-cm positioning accuracy. The transponder system could be orbital-based, but for the present design a ground-based system has been assumed with at most a single satellite for purposes of initial calibration.

When the seed unpacks, its first task is to unfurl the “home base” transponder. Power consumption has not been examined in detail but should not exceed 100 W, the amount supplied by a 1 m² solar panel. The next step is to establish an accurate navigational baseline between the home transponder and a reference transponder some distance away, perhaps using a relatively simple nonlaser surveyor’s transit. A second baseline is similarly established in some other direction, and the whole system then calibrated and synchronized to coherence. Thus deployed, a local radio navigation grid exists which can fix the position of any appropriately equipped receiver to within 1-cm accuracy, horizontally or vertically, anywhere near the seed.

Since the transponder operates on line-of-sight, each transmitter must be placed a certain distance above the ground in order to “see” the entire area for which it is responsible. The general horizon distance formula is

$X = (h^2 + 2hR)^{1/2}$, where X is the distance to the horizon, R is lunar radius, and h is height of the observer/transmitter above ground. Horizon distances for the Moon are given in table 5.8, neglecting surface irregularities.

TABLE 5.8.— HORIZON DISTANCES FOR THE MOON

Observer height h , m	Horizon distance X , m
1.0	1.9
2.0	2.6
3.0	3.2
4.0	3.7
5.0	4.2
10.0	5.9
15.0	7.2
20.0	8.3

As the original facility grows the transponder network also must be expanded. At the very minimum, a mobile robot should remain in communication with at least three noncollinear beacons to accurately fix its location. (The problems of feature shadowing and unit downtime may require the use of four or five stations. The exact number and layout can only be determined after the specific land-

ing site has been selected and mapped from orbit. One possible deployment geometry is a grid of equilateral triangles with sides roughly equal to the desired horizon distance, with transmitters at the vertices. For example, the triangle pattern edges should be roughly 2.6 km if 2-m high antennas are used. This ensures that the range circle of any mobile robot receiver always will encompass at least three transponder units, thus permitting high-accuracy triangulation. (See fig. 5.33.) Depending on the maximum size of the mature LMF and the maximum feasible height for transponder antennae, the number of transmitters necessary to support the growing seed may range from the tens up into the thousands.

In any case, the main seed computer may be presumed to carry lunar topographical maps of the landing locale, assembled prior to landing and accurate to 1-m resolution, in hard memory. This knowledge, plus the accurate positional information provided by the transponder network, should help to eliminate surprises at the expanding LMF site and lessen the need for a highly sophisticated "intelligent" vision-based surface navigation capability.

5B.2 References

Kincaid, William et al.: Summer Study Background Briefing on Computer Vision, Fault-Tolerant Systems, Large Space Structures and Antennas. Lockheed Missiles & Space Company, 7 July 1980.

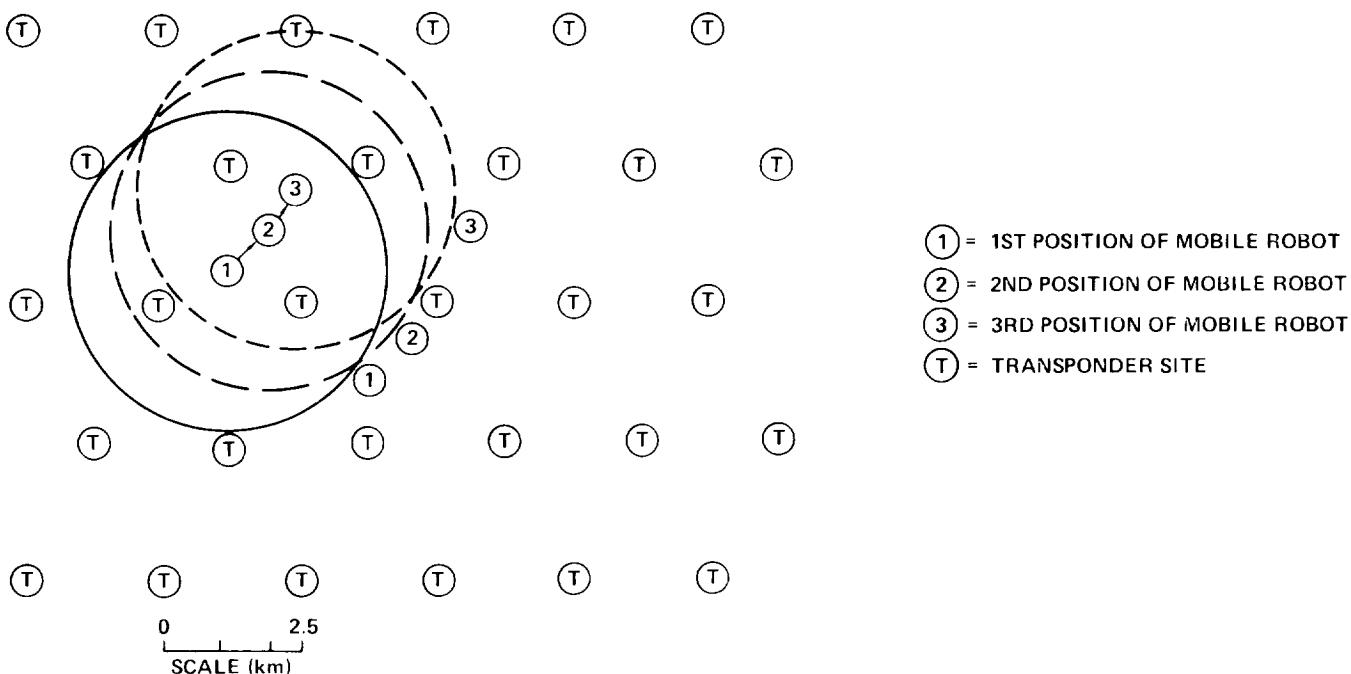


Figure 5.33.— Range circles for mobile robots using LMF transponder network for navigation.

APPENDIX 5C

LMF PAVING ROBOT SUBSYSTEM

The Platform of the Lunar Manufacturing Facility (LMF) described in section 5.3.4 serves as the physical foundation for both the original deployed Seed and the growing and mature LMF manufacturing complexes. According to Nichols (1976), "pavement is a surfacing for traveled areas, which is intended to provide a long-lasting, smooth, clean, supporting surface; to spread loads sufficiently so that base material can support them; and to protect the base against damage by traffic...." These factors are almost as important on the Moon as in terrestrial applications -- a simple graded surface would require frequent maintenance, lack cleanliness, and provide no firm foundation base to anchor SRS factory machines. A small crew of platform-building or paving robots is probably necessary for any fully automated lunar factory.

5C.1 Basic LMF Platform Design

The best material for construction of the platform ideally should be plentiful, easy to work, and most suitable for the job in terms of structural strength. Native lunar basalt appears to satisfy all three requirements adequately (Rowley and Neudecker, 1980).

Green (1980a, unpublished Summer Study document) has discussed the properties of lunar basalt at length. Raw lunar soil may be fused at about 1550 K, then allowed to cool and solidify into a very hard, exceptionally strong material. If cooling is virtually immediate -- minutes or tens of minutes -- the liquid basalt is quickly quenched and becomes a polymeric glassy substance. The material is very strong but also moderately brittle, permitting cracks to propagate rather easily. Using this option, it is necessary to divide the platform into small square-meter-size slabs to help isolate fracture failures and to permit relatively easy maintenance and repair. If the liquid basalt is permitted to cool more slowly -- allowing perhaps several hours for the melt to pass from full liquidity at 1570 K to hard solid below about 1370 K -- the material anneals into a crystalline form. This method of platform construction takes much longer and requires more energy, but would produce a far less brittle foundation. Such a basalt crystal platform could be prepared as one continuous surface, whereas the glassy basalt platform must be made in slab-sized sections.

Green has also pointed out that Moon soil has characteristics necessary to make an excellent basalt casting due to

the uncontaminated, unweathered nature of the lunar material and an extraordinarily low viscosity which is necessary for superior basalt castings. Dunning (1980, unpublished Summer Study document) considered the mechanical properties of cast basalt and found them comparable to those of cast iron and many fine steels, and superior to aluminum, brass, bronze, and copper both in compression and shear strengths. Compression strength is important in many construction applications, and shear strength is a necessary requirement for all foundation materials (U.S. Department of the Interior, 1952). A list of the properties of cast basalt is collected and modified from Anderson (1977), Baumeister and Marks (1967), and several other sources in table 5.9.

Having chosen the foundation material, the team next considered the physical configuration. According to Nichols, concrete pavements for highways are generally about 15-25-cm thick, 30 cm and higher for airport runways. Adjusting for the 0.17-g lunar gravity and the attendant reduced forces to be sustained, the equivalent load bearing strength on the Moon would require a thickness of perhaps 2.6-4.3 cm for highways. Both highways and airport runways encounter heavier use than the LMF platform is expected to receive in normal use, so a choice near the lower end of this range appears justified especially since basalt appears to be stronger than concrete in compression and shear (Baumeister and Marks, 1967; Zwikker, 1954). Consequently, a thickness of 3 cm (Green, 1980b, private communication) was tentatively selected. The square-meter size of individual slabs represents a compromise between limiting possible structural damage caused by fracture propagation and the minimum reasonable size from a practical construction standpoint.

Individual slabs comprising the platform should be formed with a 5-cm margin around the edge (slab separation 0.1 m). Rather than a second sintering pass by the paving robots, slabs are placed close enough so that overheating beyond the nominal square-meter target area for a brief period during each production cycle is sufficient to sinter neighboring blocks. (Some backfilling may be required as about 1-cm horizontal shrinkage is anticipated upon cooling.) A simple diagram of the slab pattern is shown in figure 5.34. Calculations suggest that the baseline design for paving robots should permit each device to prepare about six slabs per day in continuous operation.

TABLE 5.9.— PROPERTIES OF CAST BASALT

Physical properties	Average numerical value, MKS units
Density of magma @ 1473 K	2600-2700 kg/m ³
Density of solid	2900-2960 kg/m ³
Hygroscopicity	0.1%
Tensile strength	3.5×10 ⁷ N/m ²
Compressive strength	5.4×10 ⁸ N/m ²
Bending strength	4.5×10 ⁷ N/m ²
Modulus of elasticity (Young's modulus)	1.1×10 ¹¹ N/m ²
Moh's hardness	8.5
Grinding hardness	2.2×10 ⁵ m ² /m ³
Specific heat	840 J/kg K
Melting point	1400-1600 K
Heat of fusion	4.2×10 ⁵ J/kg ($\pm 30\%$)
Thermal conductivity	0.8 W/m K
Linear thermal expansion coefficient 273-373 K	7.7×10 ⁻⁶ m/m K
273-473 K	8.6×10 ⁻⁶ m/m K
Thermal shock resistance	150 K
Surface resistivity	1.0×10 ¹⁰ ohm-m
Internal resistivity	1.0×10 ⁹ ohm-m
Basalt magma viscosity	10 ² -10 ⁵ N·sec/m ²
Magma surface tension	0.27-0.35 N/m
Velocity of sound, in melt @ 1500 K	2300 m/sec (compression wave)
Velocity of sound, solid @ 1000 K	5700 m/sec (compression wave)
Resistivity of melt @ 1500 K	1.0×10 ⁻⁴ ohm-m
Thermal conductivity, melt @ 1500 K	0.4-1.3 W/m K
solid @ STP	1.7-2.5 W/m K
Magnetic susceptibility	0.1-4.0×10 ⁻⁸ V/kg
Crystal growth rate	0.02-6×10 ⁻⁹ m/sec
Shear strength	~10 ⁸ N/m ²

5C.2 Power Requirements for Paving Robots

To obtain a baseline design for LMF paving robots a rough estimate of the power required to fuse the basalt slabs required (in a reasonable amount of time) must be made. For this crude model, basalt platform slabs were taken as square plates with horizontal dimension x and vertical dimension y , with a sintering margin of width s (2s between slabs). A platform of radius R must be constructed within a time τ , so a total of $\pi R^2/(x + s)^2$ slabs must be produced in τ sec, a rate of $t^{-1} = \pi R^2/\tau(x + s)^2$ slab/sec.

The total input power to each square meter of lunar regolith for slab production is given by:

$$P = P_h + P_m + P_r + P_c$$

where P is total power required, P_h is the power needed to heat the basalt material to its melting point, P_m is the power necessary to melt the slab at the melting point, P_r is the rate at which energy is lost due to radiation from the top surface of the slab, and P_c is the rate of energy loss by conduction into the lunar subsurface (modified from Davies and Simpson, 1979). Radiation losses through the thin slab side walls are ignored.

To a first approximation it is sufficient to simply calculate the total energy which must be supplied and divide this by the length of time spent on each slab, hence:

$$P_h = H_s(T_m - T_L)x^2y\rho/t$$

$$P_m = H_f x^2 y \rho / t$$

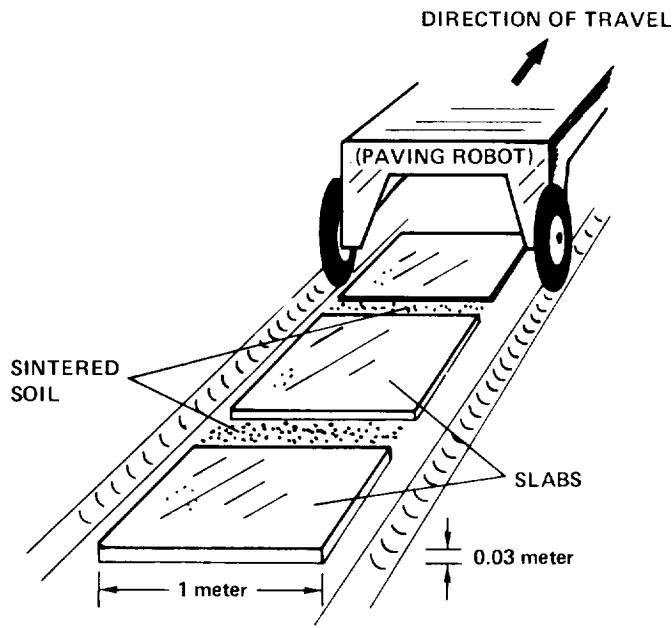


Figure 5.34. – Slab pattern of LMF cast basalt platform.

where H_s and H_f are the specific heat and heat of fusion of lunar regolith, respectively, T_m is the melting point of lunar basalt, T_L is the mean daylight temperature of the lunar surface under direct sunlight at the LMF site, and ρ is the mean density of lunar basalt.

Assuming that heating time is long compared to melting time so that the latter may be neglected, the mean radiative power loss through the exposed face of the slab is given by:

$$P_r \sim \left(\epsilon_L \sigma x^2 \int_0^t T^4 dt' \right) / t,$$

where ϵ_L is the emissivity of lunar regolith, σ is the Stephan-Boltzmann constant, and T is temperature at elapsed time t' . If heat is applied such that temperature rises at a linear rate, then:

$$P_r \sim (1/5) \epsilon_L \sigma x^2 (T_m - T_L)^4$$

$$P_c \sim (1/2)(T_m + T_L)Cx^2 / \lambda$$

where C is thermal conductivity of lunar soil and λ is the depth at which regolith temperature returns approximately to T_L .

Taking the parameters as listed in table 5.10 as typical, then for a team of five paving robots each capable of processing two slabs at once:

$$t = 10\tau(x + s)^2 / \pi R^2 = 30,600 \text{ sec}$$

and

$$P = 20,530 \text{ W}$$

TABLE 5.10. – TYPICAL VALUES FOR LMF
PAVING ROBOT PARAMETERS

$R = 60 \text{ m}$	$y = 0.03 \text{ m}$
$\tau = 1 \text{ yr} = 3.14 \times 10^7 \text{ sec}$	$T_m = 1573 \text{ K}$
$x = 1 \text{ m}$	$T_L = 503 \text{ K}$
$s = 0.05 \text{ m}$	$\rho = 2700 \text{ kg/m}^3$ (for melt)
$H_s = 840 \text{ J/kg}$	$H_f = 4 \times 10^5 \text{ J/kg}$
$\epsilon_L = 0.80$ (typical for silica brick and fire brick)	$\sigma = 5.67 \times 10^{-8} \text{ W/m}^2 \cdot \text{K}^4$
	$C = 1 \text{ W/m K}$
	$\lambda \sim 0.2 \text{ m}$

5C.3 Paving Robot Design

For the given platform layout there are many possible different modes of operation for paving robots. For instance, each robot might scoop out a hole of the appropriate dimensions, “ingest” the soil and melt it in an internal furnace, then drain the basalt magma back into the hole, neatly filling the depression. Alternative heating techniques may be readily imagined – resistance heating, controlled oxyhydrogen combustion torch with hydrogen recovery, arc furnaces (molten basalt is surprisingly electrically conductive), or induction/dielectric heating using vertical-parallel plates, finger electrodes or “stray field heating” (Cable, 1954; Curtis, 1950; Davies and Simpson, 1979). However, from a pragmatic standpoint, direct solar energy is preferred both for practical convenience and to reduce total external demand on the main LMF power grid.

The solar option for paving robots also has many degrees of design freedom, but for illustrative purposes a comparatively simple model was selected. The basic paving power module consists of a large, spherical polished aluminum mirror, constructed with easily manufactured small planar segments and affixed to a single-axis equatorial-drive turntable with a 90° sweep. This large dish is mounted on the north side of paving robots working in the lunar northern hemisphere. The robots travel east-west to maintain near-constant directional orientation at all times (except when beginning or completing a row of slabs). A planar rectangular mirror is mounted low in front of the dish, leaning forward at about 45° to direct the focus of the solar rays downward onto the carefully graded lunar surface. This second mirror may require three degrees of freedom for tracking and to permit it to project a proper square beam. Assuming accurate dish and plate mirror servo gearing, mirror positions are at all times accurately known. If the position of the robot vehicle is precisely fixed by the transponder network (see app. 5B), and an updated monthly lunar solar ephemeris is provided each robot by the seed central computer when work begins each lunar dawn, then the entire mirror pointing task can be fully automated and sun-tracking sensor apparatus eliminated. The basic optical geometry is shown in figure 5.35.

Main dish size is given by:

$$D = 2(P/\pi k^2 \alpha I \cos \delta)^{1/2}$$

where D is mirror diameter, k is the reflectivity of either of the two polished mirror surfaces (which may range up to 0.86 for aluminized glass, Weast, 1969), α is the coefficient of absorption of solar radiation for lunar basalt (taken as 0.93 for lunar albedo of 7%), I is solar insolation (1400 W/m^2), and δ is the angle between the mirror pointing axis and the Sun. In a worst case of $\delta = 20^\circ$ error, $D = 5.4 \text{ m}$.

The planar mirror is roughly rectangular, long end pointing downward, of approximate dimensions $2\text{m} \times 4\text{m}$. The heat absorbed by this mirror is at most $P(1 - k)/8 + I$ or 1710 W/m^2 , corresponding to a blackbody radiation temperature of 417 K which seems manageable. Mirrors should require resurfacing only rarely, since oxidation and meteorite pitting are not expected to be major problems.

The tentative design for the LMF paving robot is shown in figure 5.36. Each machine has a pair of dish and rectangular mirrors. Two small navigational receivers are at either end of the flatbed, permitting the onboard computer to calculate its rotational orientation with respect to the transponder network as well as its position, and a two-axis

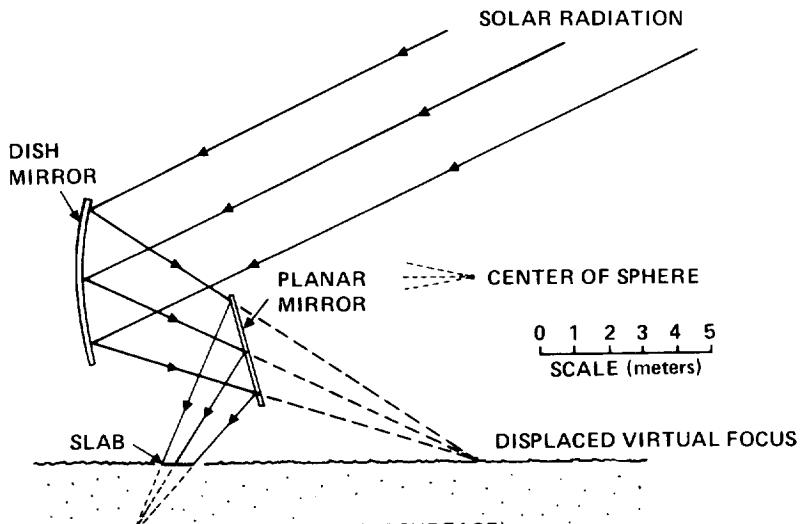


Figure 5.35.—LMF paving robot optical geometry.

level sensor measures tipping angle. Simple retractable IR sensors extend down near the slab working area to monitor energy flux and temperature, and a steerable low-resolution camera with two degrees of freedom (vertical and horizontal rotation) is installed between the two main dish mirrors to check slab placement as construction proceeds. Tires are made of soft woven basalt fibers (see sec. 4.2.2), and the vehicle is driven by four low-power electric motors fore and aft geared to steerable front and rear wheel pairs. Energy requirements for mobility and onboard computing are expected to be modest, so a few square meters of exterior solar cell paneling augmented by a rechargeable fuel cell should suffice.

SC6.4 Mass and Information Estimates

A 5.4 m spherical dish made of aluminum 1-cm thick will have a mass of about 620 kg, or 1240 kg for a pair. Similarly, the total for both planar mirrors is 440 kg. Assuming 2×10^{-2} kg computer/kg serviced (Freitas, 1980), each robot computer is about 50 kg. Camera, sensors, and navigational equipment add another estimated 50 kg. Solar panels and fuel cells may total 100 kg. The aluminum vehicle frame should be able to support its own weight (1700 kg) on Earth, so in low lunar gravity only 280 kg are required to obtain equivalent support. Each tire and drive assembly is about 40 kg, a total of 240 kg for all six wheels.

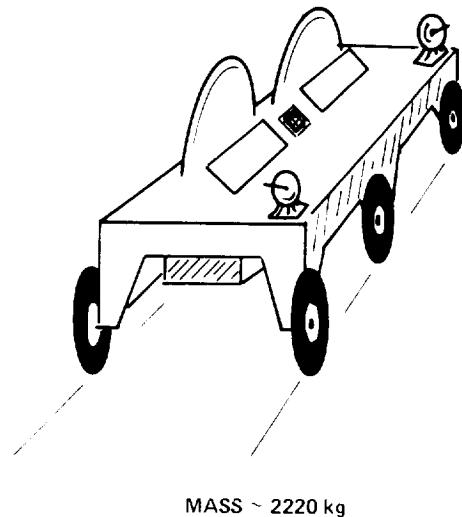
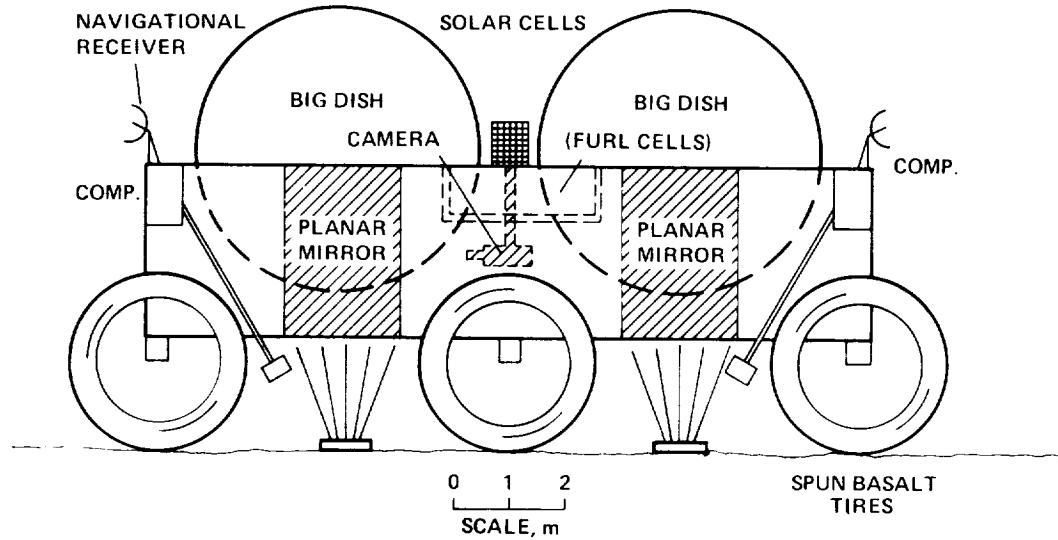


Figure 5.36.— Tentative LMF paving robot design.

Hence, the mass of each paving robot is about 2400 kg. The fleet of five included with the original seed totals 12,000 kg.

Paving robot computers must serve a number of functions, including autopilot, dish mirror guidance and control, planar mirror guidance and control, executive operating program execution, operational "timesheet" memory for the run in progress, traffic pattern coordination with other robots, neighbor machine avoidance, self-diagnostic routines for simple malfunctions, pattern recognition for slab

working area imaging, sensor control and data processing, energy system maintenance, lunar solar ephemeris memory and calculation of solar pointing angles, navigation and drive wheel control, and various routines for recognition and verification of task completion. The computation capacity needed to handle these functions probably is in the range 10^6 - 10^7 bits (about 64 K-512 K bytes). The information necessary to completely describe the machine for purposes of replication is probably about an order of magnitude greater, roughly 10^7 - 10^8 bits.

5C.5 References

- Anderson, Alfred T., Jr.: Basalt. In Encyclopedia of Science and Technology, vol. 2, McGraw-Hill, New York, 1977, pp. 110-110D.
- Baumeister, Theodore; and Marks, Lionel S., eds.: Standard Handbook for Mechanical Engineering. 7th Ed. McGraw-Hill, New York, 1967.
- Cable, J. Wesley: Induction and Dielectric Heating. Reinhold Publishing Corp., New York, 1954.
- Curtis, Frank W.: High-Frequency Induction Heating. 2nd ed., McGraw-Hill Book Company, Inc., New York, 1950.
- Davies, John; and Simpson, Peter: Induction Heating Handbook. McGraw-Hill Book Company, Ltd., London, 1979.
- Freitas, Robert A., Jr.: A Self-Reproducing Interstellar Probe. J. of the British Interplanetary Soc., vol. 33, July 1980, pp. 251-264.
- Nichols, Herbert L., Jr.: Moving the Earth: The Workbook of Excavation. 3rd ed. North Castle Books, Greenwich, Conn., 1976.
- Rowley, J. C.; and Neudecker, J. W.: Melted In-Place Lunar Soil for Construction of Primary Lunar Surface Structures. In Extraterrestrial Materials Processing and Construction, David R. Criswell, ed. Final Report, NSR 09-051-001 Mod. No. 24, Lunar and Planetary Institute, Houston, Texas, 31 January 1980, pp. 215-219. (Also published as NASA CR-158870.)
- U.S. Department of the Interior: Earth Manual: A Manual on the Use of Earth Materials for Foundation and Construction Purposes. Reprinted by Bureau of Reclamation, with revisions, February 1952.
- Weast, Robert C.: Handbook of Chemistry and Physics. 49th ed. CRC Company, Cleveland, Ohio, 1968.
- Zwikker, C.: Physical Properties of Solid Materials. Interscience, New York, 1954.

APPENDIX 5D

LMF MINING ROBOTS

5D.1 Mining Robot Functions

The requirements of seed mining robots which are components of the proposed growing, replicating Lunar Manufacturing Facility (LMF) include six basic functions:

- Strip mining – Mining robots must be able to strip mine the lunar surface without the need for overburden removal down to a depth of at least 2-3 m.
- Hauling – Having “mined” a certain amount of unbeneficiated lunar regolith, mining robots must haul their loads back to the central LMF complex for further processing. It is possible that greater efficiency might be achieved by separating the hauling and excavating functions. Also, Carrier (1979) has pointed out that it may be more efficient to beneficiate raw materials at the pit site so that only useful soil components need be transported some distance to the LMF. This would reduce the mass of mining robots required, but would increase the mass of equipment located a distance from the base site.
- Landfilling – On the return leg of each journey from the strip mining pit to the growing seed factory, each robot carries a load of unused slag or waste materials back to the pit where it is packed in as landfill. The pit might perhaps be excavated in a spiral pattern, with the fill site lagging the dig site by a gradually increasing amount.
- Grading – Mining robots must be sufficiently general-purpose to be capable of rough leveling of hilly terrain and then precision centimeter-level grading preparatory to paving robot activities.
- Cellar digging – It is conceivable that the LMF computer at the center of the circular factory complex, and perhaps certain other LMF components as well, will need to be buried under a few meters of lunar topsoil for reasons of temperature control, radiation shielding, and so forth.
- Towing – Miners are the mobile workhorses of the growing LMF complex beyond the confines of the factory platform. When mining robots or other machines break down somewhere outside, miners must go to them and tow them back if they are immobilized. For example, if a robot has become trapped because of pit wall collapse, from a fall in

loose lunar soil, or has become jammed into the surface or under fallen rocks, mining robots are ideal rescuers because they are also the LMF excavation machines and are smart and mobile enough to handle such tasks with ease.

Many additional mining-related functions conceivably could be performed by a robot system capable of the most general classes of excavation and mining activites. Indeed, such capacity might be absolutely essential if seed packages are dispatched to other planets than the Moon (e.g., Mars, Titan, Mercury, or Earth). These added functions include drilling, tunneling, blasting, and many others. But the basic six capabilities described above appear both necessary and sufficient for system survival and growth on the lunar surface.

5D.2 Design Alternatives

There exists a bewildering variety of mining and excavation machine technologies from which to draw in conceiving an autonomous vehicle (Nichols, 1976). The final design is a variant of the system devised by Carrier (1979) during a 1978 NASA-sponsored study on extra-terrestrial materials processing and construction (Criswell, 1978).

In Carrier’s system, strip mining proceeds in an annular sector Ω radians wide as shown in figure 5.37. The total system is designed for gradual expansion, based on Earth supply or lunar colony supply, over a 30-year period. In the first few years of operation, all stripping and hauling to the central processing plant is performed by front-end loaders (also called, variously, the “shovel dozer,” “dozer shovel,” “tractor loader,” “end loader,” “front loader,” or “loader”). These machines are used on Earth for digging, loading, rough grading, and limited hauling. In the lunar case, according to Carrier, the loader should be used at the outset for long hauling as the easiest way to start ore flowing into the central plant. After a few years the loaders may be augmented by a system of haulers, essentially large-volume ore trucks carrying lunar topsoil back to the central plant. This permits the loaders to strip-mine full time.

While useful as a starting point in the present study, the Carrier system cannot perform all required LMF functions. Figure 5.38 shows the basic design for the LMF mining robot. This machine is a modified loader with a rollback bucket; has a dozer blade formed on the lower face of

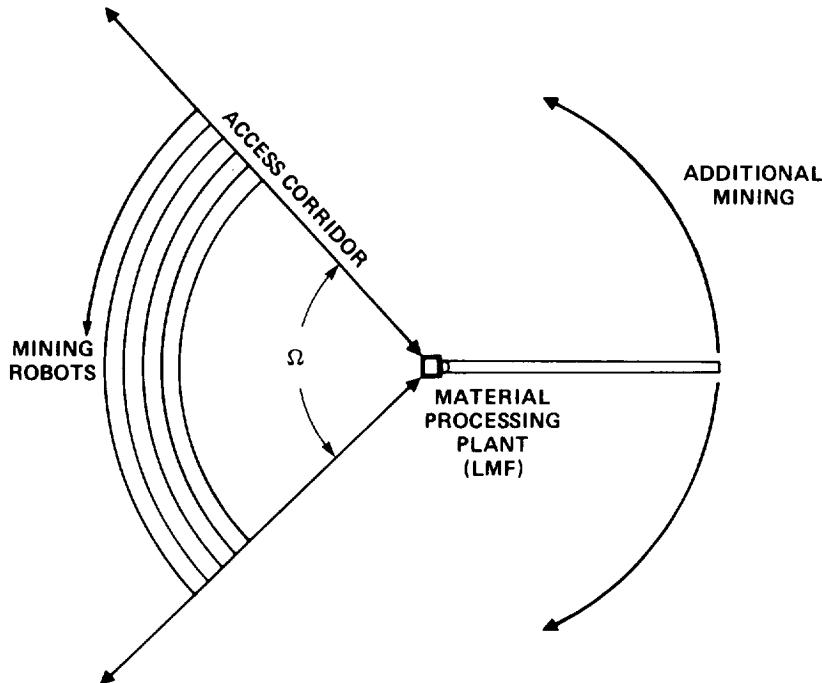


Figure 5.37.—Lunar surface strip mining.

the loader bucket, reinforced so that the bucket can be placed in a locked, elevated position and the robot driven as a dozer; and has three attachments aft which are removed during normal work, including a precision grading blade with surface contour sensors, a simple tow bar, and a somewhat more versatile towing platform.

A loader equipped in this fashion should be able to perform all six basic LMF functions enumerated above. According to Nichols, in a pinch the mining robots should also be able to act as a primitive crane, as a more versatile variable blade pitch bulldozer, as a "reach down" dozer able to cut below the depth accessible to most dozers, and as a backdragger to smooth loose dirt. Finally, it should also be possible for two loaders to join face to face to lift large boulders which neither could conveniently lift alone.

5D.3 Mining Robot Design Specifics

The team considered various specific aspects of LMF mining robot design, including machine mass, power consumption, sensor configuration, and computational and information requirements. The results and conclusions are presented below.

Robot mass and power estimates. According to Carrier (1979), haulers may be much less massive on the Moon than on Earth since the lower gravity enables the same

physical structure to carry more payload mass because the force per unit mass is less. In loaders, the vehicle mass is used as a counterbalance to prevent the machine from tipping over when fully loaded, so the mass relations for these machines change little from Earth in the lunar environment. Usual terrestrial practice is to multiply the bucket load mass by a factor of 2.0 to determine a safe tipping mass (the mass of the vehicle used as a counterweight). However, lunar equipment might incorporate automatic sensing systems to prevent tipping over so a safety factor of 1.2 should be sufficient (Carrier, 1979).

If the hauling mass per trip for all mining robots is M_h , m is the rate at which lunar materials must be mined to support the LMF replication schedule, and t is the time required for a robot to complete one cycle of operation (scoop up soil, deliver to LMF, return to pit), then $M_h = mt$. Using a factor of 1.2, the mass of mining robots is approximately $M = 1.2 M_h = 1.2 mt$.

Conservatively estimating an average of 40 km travel distance per round trip to the LMF per robot (from a 20 km radius annular pit surrounding the growing seed), an average transport speed of 10 km/hr, and a typical duty cycle of 50% for actual mining work (to leave time for repairs and nonmining labors such as grading, towing, or cellaring), then the mean cycle time

$$t = (40 \text{ km})(3600 \text{ sec/hr})/(50\%)(10 \text{ km/hr}) = 28,800 \text{ sec}$$

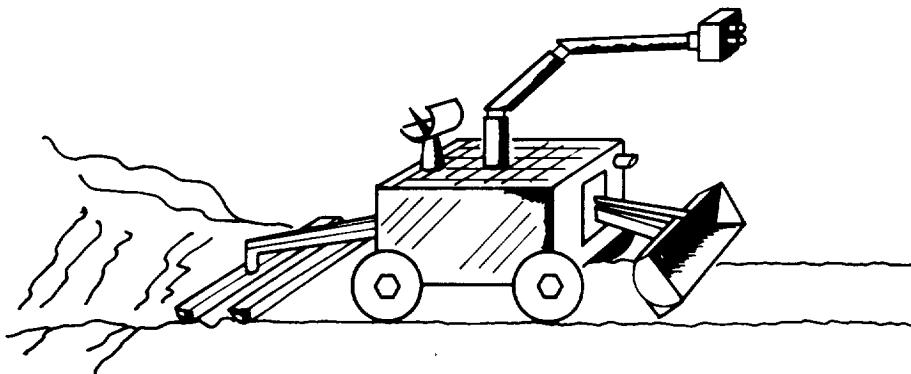
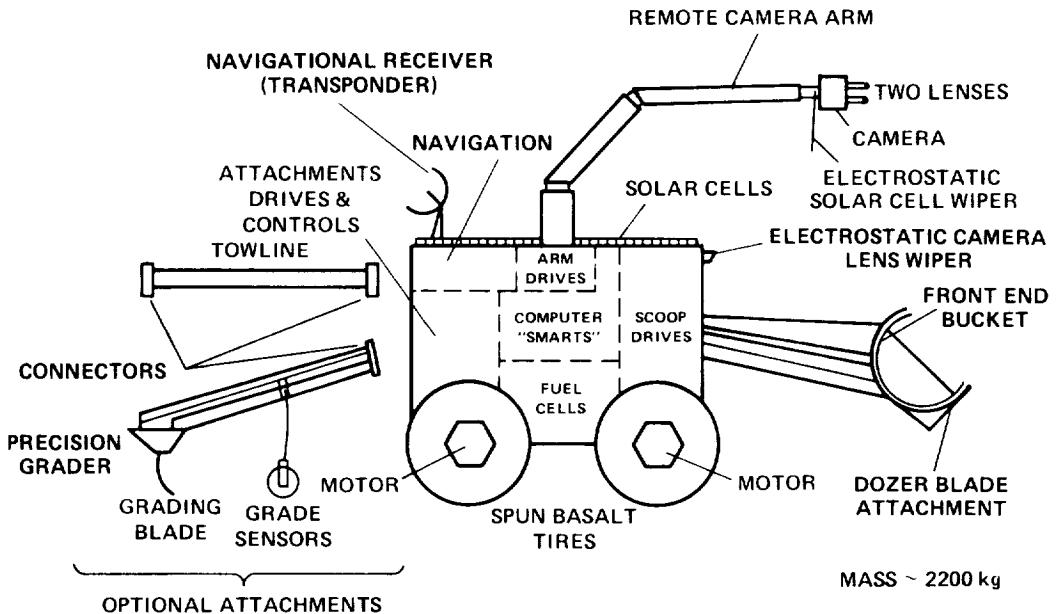


Figure 5.38.—LMF mining robot design.

The annual lunar soil hauling requirement is approximately 4×10^6 kg (see app. 5E) to replicate a new 100-ton seed each working year, so,

$$m = (4 \times 10^6 \text{ kg}) / (3.14 \times 10^7 \text{ sec}) = 0.127 \text{ kg/sec}$$

Hence, mining robot mass is

$$M = (1.2)(0.127 \text{ kg/sec})(28,800 \text{ sec}) = 4400 \text{ kg}$$

(Approximately $4400 \text{ kg}/1.2 = 3700 \text{ kg}$ of lunar material are transported each cycle.) Note that M is the total mass of robots required, not necessarily the mass per robot. In fact, it is essential that the seed carry at least two such machines so that strip mining can proceed almost continuously given a 50% duty cycle and so that a “spare” is always available

in emergency situations. Assuming linear downscaling the mass of each robot is 2200 kg.

In Carrier's strip-mining system the machines require an average of 0.3 W/kg. Mostly this is due to the hauling function, the most energy-intensive operation performed. Hence each mining robot requires about 660 W which may be drawn from 4 m^2 of photovoltaic solar cell panels mounted on every available surface. A fuel cell module (Fickett, 1978) is included in the robot design, for buffer storage and peak load coverage when power consumption may rise as high as 10 kW (as during rescue operations). This module may be recharged at any time from the LMF power grid, but this should not be necessary as the robots should be fully self-sufficient in this regard. Finally, an electrostatic lunar dust wiper is provided to maintain solar cells and camera lenses at maximum efficiency.

Sensor configuration. Sensing equipment on board includes the usual navigational receiver which ties into the high-accuracy transponder network; a two-axis level sensor so the robot knows its tipping angle with respect to the local gravity field; a detachable grading sensor which rolls along the ground just in front of the precision grading blade and provides immediate real-time feedback to permit exact control of grading angle, pitch, and slew.

The most complex sensor system is the remote camera arm. (See discussion of state-of-the-art techniques by Agin, 1979.) The camera is binocular to allow ranging and depth perception, and to provide a spare in case one camera "eye" fails. This is mounted on a long robot arm which can be directed to observe any part of itself or to survey the landscape during roving activity. The camera arm will need at least seven degrees of freedom — rotation of the arm shaft, flexure of the two intermediate joints, bending at the wrist, camera rotation, lens rotation for focus, and telephoto capability for close scrutiny of interesting features in the environment.

The mining robot camera arm is absolutely essential if the vehicle is to function in the versatile manner envisioned for it. It is not enough simply to know position in space, because the environment in which the system must operate is highly complex. It might be possible for the seed computer to give the robot a "road map" to 1 m accuracy, but this would not allow for proper navigation once the miners begin to physically alter their surroundings by digging, hauling, dozing, etc. Also, there may be objects smaller than 1 m that could cause major difficulties such as crevasses and boulders. Hence, it seems necessary to give the mining robots a true generalized "intelligent" roving capability.

Automation and AI requirements. The camera arm will require some high-level AI that lies beyond state-of-the-art. The onboard computer must keep track of the position of the moving arm in order to know where the camera is at all times. There must be routines for avoiding obstacles — for instance, the system should avoid hitting the camera with the loading bucket. Complex pattern recognition routines must be available to permit image focusing, telephoto operation, interpretation of shadows and shapes, differentiation between protrusions and depressions in the surface, and intelligent evaluation of potential risks and hazards of various pathways and courses of action. The onboard computer must have an accurate representation of its own structure stored in memory, so that the camera may quickly be directed to any desired location to inspect for damage, perform troubleshooting functions, or monitor tasks in progress. Finally, the computer must have diagnostic routines for the camera system, in case something simple goes wrong that can easily be corrected *in situ* without calling for outside assistance.

According to Carrier (1979) the automatic haulers can easily be designed to operate in an automatic mode, requir-

ing only occasional reprogramming but substantially more advanced AI pattern recognition systems. (In 1980 a child's toy was marketed which can be programmed to follow simple paths (Ciarcia, 1981; "Toy Robots," 1980).) Carrier suggests that since there are so many variables associated with excavation "it is doubtful that the front-end loader could operate automatically," though the team disputes this conclusion. In addition to sophisticated pattern recognition and vision systems (Williams et al., 1979), the robot miners need a "bulldozer operator" expert system of the kind under development at SRI for other applications (Hart, 1975, and personal communication, 1980). Such an expert system would embody the knowledge and skills of a human excavator and could substitute for human control in most circumstances. In addition, expert systems might be executed remotely by a process called "autonomous tele-operation." In this mode of operation, mining robots can be remote-controlled via transponder network links by the master LMF computer, thus reducing onboard computer complexity.

Additionally, the onboard computer must handle such comparatively mundane chores as clocking, operating drive trains on the wheels, turning controls, blade angle control and configuration, task completion testing and verification, guidance and navigation, and internal diagnostics. An executive program is also required, capable of accepting new orders from the central LMF computer (e.g., "rescue machine X at position Y") and semiautomously calculating how best to execute them (Sacerdoti, 1980).

Computation and information requirements. A first-cut estimate of the computational capacity required on board reveals that three major computer subsystems are involved: (1) robot camera arm (seven degrees of freedom, binocular vision, rangefinding, sophisticated AI such as pattern recognition and inference); (2) excavator expert system (controls physical operations, understands a world model, has expectations about outcomes, and can troubleshoot simple problems); and (3) high-level executive system (reprogrammability, interpretation, and "common sense" reasoning). Each of these subsystems represents a different problem and must be separately analyzed.

The robot system with mobile camera studied by Agin (1979) engaged in very primitive pattern recognition. This included insertion of bolts into holes, positioning a movable table relative to a fixed camera, velocity tracking (a Unimate PUMA arm, camera in hand, follows an object moving past on a conveyor belt), spot welding on a moving assembly line, and following a curved path in three dimensions at constant velocity (simulating industrial activities such as gluing, sealing, and seam-following). Again's visual recognition routines ran on a PDP-11/40 minicomputer, a 28K application, and the PUMA robot arm was controlled by the usual LSA-11 microcomputer which has a 16K capacity using 32-bit words. The visual system for the proposed

mining robot will be at least 1-2 orders of magnitude more complicated than Agin's system, so we would estimate a control requirement of 10^6 - 10^7 bytes, or about 10^7 - 10^8 bits of computer capacity.

The SRI expert system "PROSPECTOR" runs on a DEC-10 computer with a 150K operating program and a 1M database, a total of about 3.2×10^7 bits (Hart, personal communication, 1980). PROSPECTOR "knows" about 1000 different factors related to prospecting. It is difficult to imagine a general excavation expert system requiring more than ten times this, or 10,000 factors, to achieve adequate autonomous operation with troubleshooting capability — the PROSPECTOR expert has generated some impressively accurate results in searches for ore-bearing bodies. If the "EXCAVATOR" expert system is thus about one order of magnitude larger than PROSPECTOR, the basic computational requirement is 10M or 3.2×10^8 bits.

Mining robot executive computer requirements are more difficult to estimate, as there are few previous directly applicable models. A simple passenger aircraft autopilot probably will run on a 32K microprocessor, and a "smart rover" vision-equipped wheeled mobile robot with a 6-degree-of-freedom arm developed in the 1970s at JPL used state-of-the-art microprocessors. Remarks by Sacerdoti (1979, 1980) on the subject of autonomous planning and execution in robotics suggest that the system required for

robot miners is perhaps 1 to 2 orders of magnitude beyond current technology; thus the executive system may require a memory capacity of about 1 to 10M, or 3 to 30×10^7 bits.

Summing the requirements for the three major computer subsystems gives an "information bandwidth budget" of 3.6 - 7.2×10^8 bits, centering on about 500 Mb. The information necessary to completely describe the system for purposes of self-replication is probably on the order of 10^9 bits.

5D.4 LMF Approach and Access Geometry

In the baseline LMF scenario, mining robots must assume all hauling duties beyond the factory platform. Thus, it becomes necessary to specify how these mobile machines, normally bearing loads of strip-mined soil to be processed, will approach the factory and deposit their cargoes at an appropriate input location. A related query is how and where robots will accept waste products for transport to the pit for use as landfill. These questions are of some importance, because as the seed expands to full maturity it may become physically more difficult to exchange raw materials and wastes with interior LMF processing systems unless the access geometry has been designed to accommodate growth.

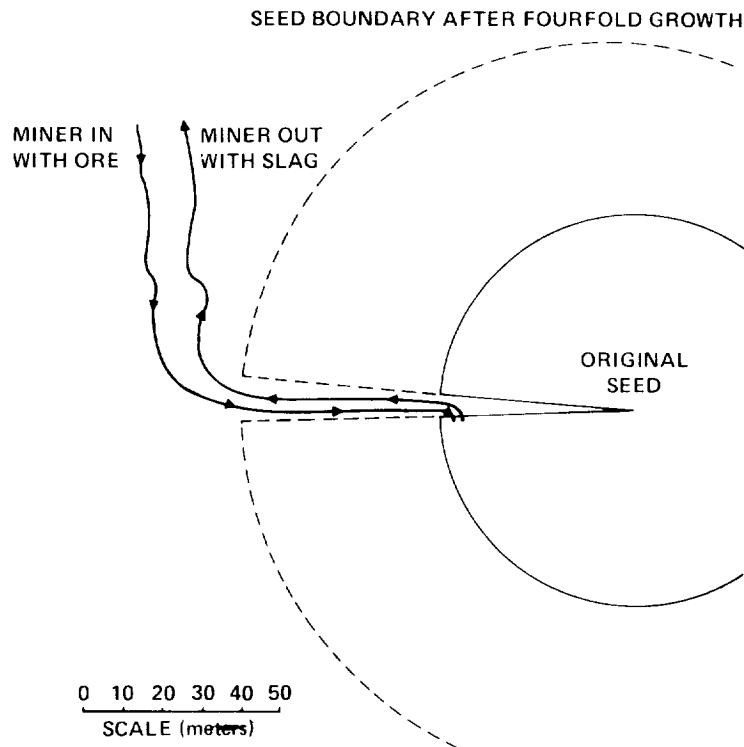


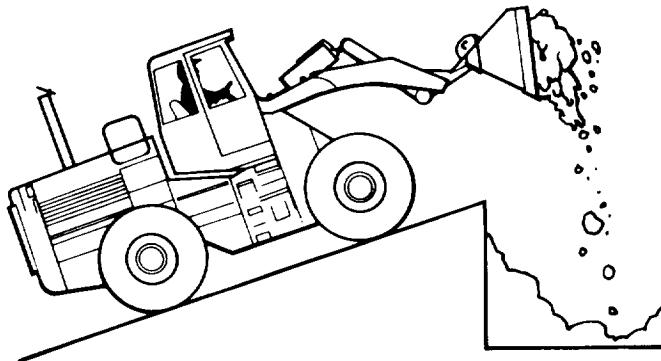
Figure 5.39. — LMF constant-angle wedge corridor access route.

The solution adopted by the team is to earmark a constant-angle wedge corridor for permanent use as a mining robot access road. A 5° angle provides a corridor width of 5 m at the perimeter of the initial 60-m radius seed — comfortably enough room for a mining robot to enter, drop off its cargo, pick up a load of waste materials, and then withdraw. The area of the constant-angle corridor increases as R^2 . This is the same dependence on radius exhibited by the area of the growing “seed,” hence LMF mass, raw materials requirements, and waste production will increase at the same rate as the access corridor which supports interior factory systems. In other words, the expanding corridor prevents internal LMF systems from becoming “landlocked” as seed mass and radius grow exponentially in time.

The wedge corridor geometry is shown in figure 5.39. Note that as the LMF grows larger, mining robots (or any other external transport vehicle) must traverse ever greater distances, on average, to reach the entry corridor. For this

reason a minimum of two such corridors should be provided, with the factory organized as two identical halves as suggested in figure 5.19. Further studies will be required to determine the optimum access and LMF configuration geometry from the standpoint of scheduling, efficiency, and access time.

Mining robots deliver raw materials to an input hopper located in the chemical processing sector, as shown in figure 5.40. Outshipments of waste materials are delivered to them in similar fashion. These hoppers serve as materials depots, able to help sustain LMF operations during periods when the supply of lunar topsoil is interrupted for any reason. Since each of the two initial seed robots makes one round trip about every eight hours, a hopper intended to serve as a one-week buffer must have a capacity of 42 mining robot loads or 76,900 kg of lunar regolith. A roughly cubical hopper constructed of 1 cm sheet aluminum and able to contain the weekly input volume of 42.7 m^3 has a mass of 1650 kg.



COURTESY OF INTERNATIONAL HARVESTER COMPANY

Figure 5.40.— Raw material delivery to input hopper.

5D.5 References

- Agin, Gerald J.: Real Time Control of a Robot with a Mobile Camera. SRI International Technical Note 179, February 1979.
- Carrier, W. D., III: Excavation Costs for Lunar Materials. Fourth Princeton/AIAA Conference on Space Manufacturing Facilities, Princeton, New Jersey, 14 May 1979, AIAA Paper 79-1376.
- Ciarcia, Steve: A Computer-Controlled Tank. Byte, vol. 6, February 1981, pp. 44-66.
- Criswell, David R.: Extraterrestrial Materials Processing and Construction. NASA CR-158870, 1978.
- Fickett, Arnold P.: Fuel-Cell Power Plants. Scientific American, vol. 239, December 1978, pp. 70-76.

- Hart, Peter E.: Progress on a Computer-Based Consultant. SRI International, Menlo Park, January 1975.
- Nichols, Herbert L., Jr.: Moving the Earth: The Workbook of Excavation. 3rd ed. North Castle Books, Greenwich, Conn., 1976.
- Sacerdoti, Earl D.: Problem Solving Tactics. SRI International Technical Note 189, July 1979.
- Sacerdoti, Earl D.: Plan Generation and Execution for Robotics. SRI International Technical Note 209, 15 April 1980.
- Toy Robots Gaining Intelligence. Byte, vol. 5, October 1980, p. 186. See also “Inside Big Trak,” Robotics Age, vol. 2, Spring 1980, pp. 38-39.
- Williams, D. S.; Wilf, J. M.; Cunningham, R. T.; and Eskenazi, R.: Robotic Vision. Astronautics and Aero-nautics, vol. 17, May 1979, pp. 36-41.

APPENDIX 5E

LMF CHEMICAL PROCESSING SECTOR

Mining robots deliver raw lunar soil strip-mined from the pit to large input hoppers along the edge of the entry corridors into the chemical processing sector. The primary responsibility of the materials-processing subsystems is to accept lunar regolith, extract from it the necessary elemental and chemical substances required for system growth, replication, and production, and then return any wastes, unused materials, or slag to an output hopper to be transported back to the surrounding annular pit by mining robots for use as landfill.

It is possible to achieve qualitative materials closure (see sec. 5.3.6) — complete material self-sufficiency within the Lunar Manufacturing Facility (LMF) — by making certain that chemical processing machines are able to produce all of the 84 elements commonly used in industry in the United States and the global economy (Freitas, 1980). However, such a complete processing capability implies unacceptably long replication times T (on the order of 100–1000 years), because many of the elements are so rare in the lunar or asteroidal substrate that a vast quantity of raw soil must be processed to obtain even small amounts of them. By eliminating the need for many of these exotic elements in the SRS design, replication times can be cut by as much as three orders of magnitude with current or foreseeable materials processing technologies.

Hence, it is desirable to determine the minimum number of elements and process chemicals and to fix the lowest extraction ratio R (kg input material/kg useful output material, see sec. 5.3.6) which can still maintain closure of the system, thus minimizing the replication time T .

SE.1 Minimum LMF Requirements: Elements and Process Chemicals

The elemental and chemical requirements of the expanding LMF fall into a fairly small number of broad categories summarized in table 5.11. Note that these are the minimum (or very nearly so) requirements for LMF qualitative materials closure — an “adult” LMF entering production phase may need additional chemical processing capabilities which may be programmed into the factory’s operational software. Table 5.11, however, lists only those minimum requirements necessary to achieve closure for a seed during the growth phase.

TABLE 5.11.— MINIMUM SEED ELEMENT AND PROCESS CHEMICAL REQUIREMENTS

- I. Structural metals, alloys, hard parts, tubing, containers, etc. — Fe, Al, Mg, Ti, Mn, Cr, C, Si, Ca
- II. Building materials, insulation, fabrics, glass plate, ceramics, crucibles, furnace linings, chemistry glassware, high-temperature refractories, etc. — lunar soil as found (basalt when fused), anorthite ($\text{CaAl}_2\text{Si}_2\text{O}_8$), silica (SiO_2), alumina (Al_2O_3), magnesia (MgO), feldspar
- III. High purity electronics-grade materials for the manufacture of solar cells, computer chips, etc. — Si, O₂, Al, P, B
- IV. Magnetic materials — Fe
- V. Fluorine chemistry containers — Fe, C, F₂
- VI. Process chemicals for bulk manufacturing, high-purity electronics chemical production — H₂O, HF, N₂, H₃PO₄, HNO₃, SiH₄, CF₄ (Freon for microelectronic “dry etching” processes), NaOH, Cl₂, H₂SO₄, CaCl₂, Na₂CO₃, NH₃
- VII. Process minerals, inputs to chemical processing sectors — olivines, pyroxenes, feldspars, spinels, ilmenite, apatite, anorthite, tincalconite (anhydrous borax).

Total of 18 elements, 12 minerals/mineral types, and 11 additional process chemicals.)

It will be argued that a chemical processing system capable of producing each of the above from raw lunar soil has achieved full self-sufficiency, or materials “closure.”

Demonstration of materials closure plausibility. The components in table 5.11 were obtained first by taking a very basic list of necessary elements (the first four categories) for the entire LMF and adding to these any additional substances necessary to chemically produce the original items. This resulted in an increase in the number of items, therefore, all newly added items themselves had then to be similarly checked to ensure that each of them could be produced from the materials already at hand. This procedure was iterated until closure apparently was achieved.

The list includes reagents necessary for the production of microelectronic circuitry (Oldham, 1977), even though "wet chemistry" may not be necessary for this application in space manufacturing (Zachary, 1981). The team is unaware of any significant omissions in table 5.11, which demonstrates essential qualitative closure.

SE.2 Derivation of Minimum Requirements: Qualitative Materials Closure

The lunar substrate from which the required substances are extracted or manufactured has a mean global mineral content as shown in table 5.12. Source minerals for boron do not appear in this list, nor do the sources for volatiles implanted by the solar wind. A summary of all elements found to date in the lunar regolith samples returned by Apollo and Luna missions may be found in table 4.1.

To plausibly demonstrate materials closure, it must be shown that every item on the requirements list can be derived from other items on the list and that all elements are derived from those found in the lunar regolith. To fully and rigorously demonstrate closure, a detailed element-by-element breakdown of the entire factory would be required, giving the mass of each element or process chemical required followed by a convincing demonstration that such quantities could indeed be produced using only the amounts of other substances known to be available and an input of lunar material. This latter set of conditions is called quantitative closure.

Preparation of process minerals. A comparison of the list of process chemicals in category VI in table 5.11 with the minerals found in lunar soil (table 5.12) suggests that it may be possible to use raw lunar soil as input to the materials processing extraction machines if these minerals require no beneficiation. In the event such beneficiation is needed to obtain the specific minerals in separated form, the electrophoretic separation technique described in section 4.2.2 may be used. This method involves placing finely divided powdered lunar dust in aqueous (or slag, or other solvent) suspension which has a solvent pH tuned to match the isoelectric potential of the desired mineral species. A cross voltage is applied and all minerals but the one desired migrate away, leaving behind a purified residue – in the present case, anorthite and the category VII (table 5.11) process minerals may be recovered. Preliminary testing of the electrophoretic separation concept with simulated lunar soil has been successful (Dunning and Snyder, 1981).

In addition, the electrophoretic technique may prove invaluable in separating out "trace minerals" from lunar soil, in particular apatite and possible differentiated boron-containing minerals which may exist in the lunar regolith.

Separation of iron. The magnetic properties of lunar soil are due almost entirely to the presence of metallic iron, which occurs in lunar soil as a free element in the amount of 0.5% by weight, roughly 5% of the total iron content of the lunar regolith. Since it is magnetic, metallic iron may be

TABLE 5.12.— MINERALS TYPICALLY FOUND
IN LUNAR REGOLITH
(from Williams and Judwick, 1980)

Major	Minor
Olivine $(\text{Mg},\text{Fe})_2\text{SiO}_4$	Spinel $(\text{Fe},\text{Mg},\text{Al},\text{Cr},\text{Ti})\text{O}_4$
Pyroxene $(\text{Ca},\text{Mg},\text{Fe})\text{SiO}_3$	Armalcolite $(\text{Fe}_2\text{TiO}_5)$
Plagioclase feldspars $(\text{Ca},\text{Na})\text{Al}_2\text{Si}_2\text{O}_8$	Silica (quartz, tridymite, cristobalite) SiO_2
	Iron Fe (variable amounts of Ni and Co)
	Troilite FeS
	Ilmenite FeTiO_3
	Trace
<i>Phosphates</i>	<i>Oxides</i>
Apatite ^a $\text{Ca}_5(\text{PO}_4)_3(\text{F},\text{Cl})_3$	Rutile TiO_2
Whitlockite ^a $\text{Ca}_9(\text{Mg},\text{Fe})(\text{PO}_4)_7(\text{F},\text{Cl})$	Corundum (?) Al_2O_3
	Hematite (?) Fe_2O_3
	Magnetite Fe_3O_4
	Goethite (?) $\text{FeO}(\text{OH})$
<i>Zr mineral</i>	
Zircon ^a ZrSiO_4	<i>Metals</i>
Baddeleyite ZrO_4	Copper (?) Cu
	Brass (?)
	Tin (?) Sn
<i>Silicates</i>	
Pyroxferroite $(\text{Fe},\text{Mg},\text{Ca})\text{SiO}_3$	<i>Zr-rich mineral</i>
Amphibole $(\text{Ca},\text{Mg},\text{Fe})(\text{Si},\text{Al})_8\text{O}_{22}\text{F}$	Zirkelite or zirconolite ^a
Garnet (?)	$\text{CuZrTi}_2\text{O}_7$
Tranquillityite ^a $\text{Fe}_8\text{Zr}_2\text{Ti}_3\text{Si}_3\text{O}_4$	<i>Meteoritic minerals</i>
	Schreibernite $(\text{Fe},\text{Ni})_3\text{P}$
<i>Sulfides</i>	Cohenite $(\text{Fe},\text{Ni},\text{Co}),\text{O}$
Mackinawite $(\text{Fe},\text{Ni})_9\text{S}_8$	Niningerite $(\text{Mg},\text{Fe},\text{Mn})\text{S}$
Pentlandite $(\text{Fe},\text{Ni})_9\text{S}_8$	Lawrencite (?)
Cubanite CuFe_2S_3	$(\text{Fe},\text{Ni})\text{Cl}_2$
Chalcopyrite CuFeS_2	
Sphalerite $(\text{Zn},\text{Fe})\text{S}$	

^aThese minerals are known to exhibit complex substitutions, particularly of elements as Y, Nb, Hf, U, and the rare earth elements that are concentrated in these minerals.

separated from the raw lunar substrate by straightforward electromagnetic techniques directly as the raw input material leaves the input hopper. This Fe will be fairly pure, containing only about 5% nickel and 0.2% cobalt (Phinney et al., 1977).

Structural metals and metal oxides. Of all the chemical materials processing options studied to date, the hydrofluoric (HF) acid leach technique appears to have the best potential for minimum operating mass, ease of element separations to high purity, and favorable energy and heat-rejection requirements (Arnold et al., 1981; Waldron et al., 1979). HF acid leach (Waldron et al., 1979), shown in figure 5.41 in flowsheet form, uses low-temperature hydrochemical steps to separate the silica content of the lunar raw material from metallic oxides in minerals by converting them to fluorides and fluorosilicates. The silica is then vaporized as SiF₄, leaving Ca, Al, Fe, Mg, and Ti fluoro salts to be separated by a variety of solution, precipitation, ion exchange, and electrolytic steps. These are then reduced to the pure metallic form with sodium metal, which is recycled. (HF is added as a major process chemical.)

Sodium for the reduction of metals and silicon may be obtained by a modified Castner cell process, which involves the electrolysis of molten NaOH to produce Na, O₂, and H₂. Iron electrodes can be used in this application. (NaOH must also be added to the process chemicals list.)

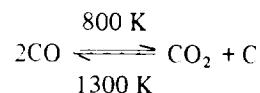
Metal oxides and silicon dioxide can be obtained, where needed as ceramics, refractories, or for glasses, by hydrolysis of the fluoride or fluorosilicate with H₂O steam (for the metal oxides), with NH₃ (for silicon dioxide), or by ion exchange methods. (Water and ammonia are thus added to the list of process chemicals.) Electronics-grade silicon may be prepared through zone-refining and other techniques with up to nine-9s purity, although these processes have not been thoroughly investigated in the present study.

In a discussion of the HF acid leach technique, Criswell (1978) points out that the process with its various options is adaptable to several of the potential lunar minerals or concentrates including feldspars, pyroxenes, olivines, and even nonsilicates such as ilmenite and spinels. Beneficiation of these minerals (the major constituents of lunar soil) seems unnecessary since the appropriate separations are performed later on the fluorides and fluorosilicates. However, if necessary, this beneficiation can be accomplished using the electrophoretic method described above.

In addition to Fe, Al, Mg, Ti, Ca, Si, and O₂, it is possible that the HF acid leach process may be used to prepare Cr and Mn. These two elements are present in pyroxene (up to 0.5% MnO, up to 1.25% Cr₂O₃), olivine and spinel (which contain Cr). CrF₂ is slightly soluble in water; MnF₂ is soluble, so the techniques described above should still be applicable although the details of this extension have not been extensively studied.

One final problem unique to the HF process is the question of containers. Process vessels and tubing normally employed in terrestrial industry are attacked by hydrofluoric acid. One solution is to use special carbon steel alloys for this purpose — these are customarily employed for storage of fluorine gas because a protective layer of iron fluoride forms which greatly impedes further chemical attack. A second alternative is to use hydrocarbon-based waxes, paraffins or plastics which are not attacked by HF, applied as a thin layer to the insides of pipes and containers. Yet a third option is to develop new structures perhaps based on sulfur and phosphorus (Allcock, 1974) and other inorganic polymers (Lee, 1979) which could be in reasonably plentiful supply in the lunar factory.

Extraction of volatiles. Lunar soil heated to 1300 K releases 0.1% by weight of the following trapped volatiles: CO, CO₂, N₂, H₂, H₂O, SO₂, H₂S, CH₄, and inert gases (He, Ar, Ne, Kr, Xe). As much as 0.5–1.5% by weight may be released upon heating to 1700 K (Phinney et al., 1977). CO may be reduced to carbon by methanation followed by decomposition of the CH₄ species over a refractory catalyst (such as MgO) to C and H₂. CO₂ may be reduced to CO by making use of the reversible reaction:



That is, CO₂ passed over elemental C above 1300 K reduces to CO, which can then be methanated and further reduced to C over hot refractory. N₂, H₂, H₂O, and SO₂ are desirable process chemicals. H₂S may be burned in O₂ to yield SO₂ and water. A sharply limited supply of O₂ results in steam and sulfur vapor. If SO₂ and H₂S are mixed at room temperature, they react to form water and elemental sulfur. Finally, oxygen bubbled through an aqueous solution of H₂S produces a precipitate of elemental sulfur.

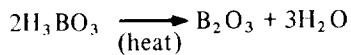
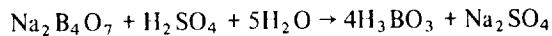
Inert gases are useful in lasers and for providing a non-reactive atmosphere, and may be separated by fractional condensation using cold traps at various temperatures.

Boron production. Historically on Earth the most important source of boron has been borax or tincal (Na₂B₄O₇·10H₂O), though today the more common source is kernite or rasorite (Na₂B₄O₇·4H₂O). Other boron minerals include colemanite (Ca₂B₆O₁₁·5H₂O), ulexite (NaCaB₅O₄·8H₂O), priceite (Ca₄B₁₀O₁₉·7H₂O), boracite (Mg₃B₇O₁₃Cl) in salt beds, and sassolite (H₃BO₃).

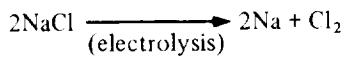
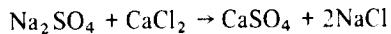
Boron minerals on the Moon are likely associated with phosphorus-bearing apatite species (Dunning, personal communication, 1980), although it is possible that local concentrations of the most common anhydrous boron mineral,

tinicalconite ($\text{Na}_2\text{B}_4\text{O}_7$), may be found in the vicinity of ancient lunar volcanic vents. In either case it should be possible to isolate the boron species using a combination of chemical and electrophoretic techniques. However, the details of this process cannot be specified until available boron resources on the Moon are more precisely characterized.

Terrestrial boron-containing minerals are either calcium or sodium borates. A calcium borate may be converted to a sodium borate by treatment with Na_2CO_3 , yielding borax and CaCO_3 which precipitates out of solution. (Calcium carbonate may be recycled by roasting to obtain CaO and CO_2 , from the latter of which elemental carbon can be recovered.) Sodium borates are reduced to boric oxide in two steps:



The sodium and sulfur may be recycled by the following steps:



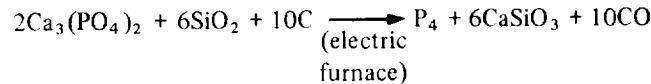
(Sulfuric acid and calcium chloride are added to the list of process chemicals.)

Low-purity boron is prepared by reduction of B_2O_3 with Mg, followed by vigorous washing with sodium alkali and HF. The impurities are a mixture of oxides and borides. Almost pure boron (up to 99.9999% is available commercially by this method) for electronics applications may be prepared by vapor phase reduction of BCl_3 (or BBr_3) with hydrogen on electrically heated filaments. BCl_3 is prepared by heating B and Cl_2 directly at 800 to 1100 K. Possible filament materials have not been investigated, but the mass requirement is probably less than 1 kg. Chlorine is added to the process chemicals list, since F_2 cannot be substituted for Cl_2 for vapor phase purification.

Phosphorus and halogens. More than 200 minerals containing up to 5% phosphorus by weight are known on Earth, but the two main species available on the Moon are fluorapatite, $\text{Ca}_5(\text{PO}_4)_3\text{F}$ and chlorapatite, $\text{Ca}_5(\text{PO}_4)_3\text{Cl}$. The other lunar phosphorus-bearing mineral, whitlockite, is generally given as $\text{Ca}_3(\text{PO}_4)_2$, but often is found associated with Mg, Fe, F, and Cl. Fluorapatite is by far the most abundant and is also the major source of fluorine on the lunar surface. (Amphibole has a trace of fluorine, but this small amount is probably not worth the trouble to extract.)

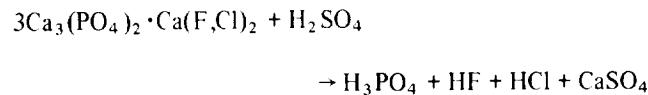
Chlorapatite, very rare by comparison, is the major source of chlorine on the Moon, except for lawrencite (a nickel-iron chloride believed derived from meteorites). Whitlockite is also very rare.

Apatite is separated from lunar soil by the electrophoretic process described above. The calcium phosphate is then reduced to P_2O_5 by heating with silica (available from the HF leach stage) yielding pure phosphorus when treated with carbon:



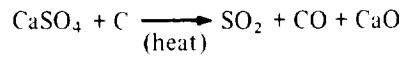
Alternatively, calcium phosphate dissolved in sulfuric acid gives phosphoric acid plus insoluble calcium sulfate (which may be recycled, see below). The acid is then reduced with carbon to obtain elemental phosphorus.

The sulfuric acid technique appears best for halogen extraction. When acted upon by sulfuric acid, a natural mixture of fluorapatite and chlorapatite undergoes the following net reaction:



This results in a solution of the three acids. If heated to above 390 K (but below 486 K), the HF and HCl boil off leaving pure orthophosphoric acid behind. The evaporate is condensed, then separated into HF and HCl by either of two methods. First, the acid solution is desiccated in vapor form over anhydrous CaCl_2 , then cooled to 273 K. HF condenses and is removed in liquid form, leaving HCl gas to be electrolyzed to obtain H_2 and Cl_2 . Or, second, after desiccation with CaCl_2 the HF/HCl solution is electrolyzed with the release of H_2 at one electrode and a mixture of F_2 and Cl_2 at the other. This mixture is cooled to 240 K which liquefies the Cl_2 (to be drained off) leaving F_2 gas, which may be combined directly with the liberated H_2 to make HF. This entire problem may also be circumvented if fluorapatite and chlorapatite can be separated using electrophoretic beneficiation.

To recover sulfur, a valuable volatile, from the above process, the calcium sulfate is recycled by roasting according to:



Supporting reagents. Reagents necessary to ensure closure of the LMF chemical processing sector include sodium hydroxide, silane, sulfuric acid, nitric acid, freon, ammonia, calcium chloride and sodium carbonate. The derivation of each is briefly reviewed below.

Lunar pyroxene contains up to 0.2% and lunar plagioclase up to 1.5% Na₂O (Williams and Jadwick, 1980). Specific pyroxene minerals containing Na are acmite or aegirite, Na₂O·Fe₂O₃·4SiO₂ and jadeite, Na₂O·Al₂O₃·4SiO₂. Among plagioclase feldspars are anorthoclase, albite, and andesine, Na₂O·Al₂O₃·6SiO₂. After these minerals are obtained by electrophoresis, roasting causes the Na₂O component to sublime above 1200 K. By 1800 K as much as 70% of the available Na₂O may have evaporated, leaving behind a still solid residue of iron, silicon, and aluminum oxides (Williams and Jadwick, 1980). The liberated sodium oxide is dissolved in water to give NaOH. The small amount of Na produced during boron reduction may be added directly to the HF leach system as metal, or hydrated to form NaOH, as required.

Silane for microelectronic wafer fabrication may be prepared in either of two ways. First, elemental silicon may be heated in the absence of air with magnesium to form the silicide, which is then hydrolyzed with sulfuric acid to silanes and MgSO₄ (which can be recycled for sulfur much like calcium sulfate). This hydrolysis gives about 25% yield of silicon hydrides, comprised of 40% SiH₄, 30% Si₂H₆, 15% Si₃H₈, 10% Si₄H₁₀, and 5% of Si₅H₁₂ and Si₆H₁₄. These may be separated by fractional distillation; or, if cooled to below 258 K, all species liquefy except SiH₄, which remains a gas and can be removed. A second process suggested by Criswell (1980a) involves hydrolysis of the Mg₂Si with HCl, with the magnesium chloride hydrolyzed by steam to recover the HCl.

Sulfuric acid is relatively simple to prepare, provided a suitable catalyst is available. In the two-step contact process, SO₂ is burned in oxygen and in the presence of catalyst to the trioxide, which is then dissolved in water to yield the acid. The usual catalyst was, traditionally, finely powdered platinum, and more recently vanadium pentoxide. If possible, the use of these substances should be avoided as Pt and V are rare in the lunar regolith. Fortunately, practically all refractory substances have some degree of catalytic activity in the contact process, provided they are immune to impurities. Alternative and plentiful viable catalyst agents include pumice (SiO₂·Al₂O₃), porcelain or powdered ceramic, and ferric oxide (Fe₂O₃), all of which are active and readily available in the LMF.

Nitric acid is more difficult to prepare, primarily because of the difficulty of “fixing” nitrogen chemically. The two most common commercial processes for acid production involve the use either of existing nitrate stocks or of platinum (for the catalytic oxidation of ammonia), neither of which is feasible at the LMF. A third method, not feasible commercially because of its low energetic efficiency, is the electric arc technique first discovered by Priestley in 1772. Elemental nitrogen and oxygen are passed through a spark discharge, producing nitric oxide with a yield of 2.5% under ideal conditions. After rapid quenching of the reaction mixture, the NO reacts rapidly below 873 K in an excess of O₂

to form NO₂, which makes nitric acid upon contact with water. Biological nitrogen fixation using *Rhizobium* and *Azotobacter* microorganisms is an interesting alternative and should be investigated further.

Freon (CF₄) is prepared by fluorination of methane with elemental fluorine. The resulting mixture of CF₄ and HF is separated by dissolution in water. There are two potentially feasible methods for producing ammonia. First is the standard Haber process, in which elemental nitrogen and hydrogen are combined directly at 800 K in the presence of iron and aluminum oxide catalysts. In the second process, magnesium is ignited at 600 K in a nitrogen atmosphere to form the nitride, which is then hydrolyzed to yield ammonia and magnesium hydroxide. Water and MgO are recycled by roasting the hydroxide.

Only very limited amounts of CaCl₂ are needed, so direct combination of the elements (both of which are already available) is the preferred production pathway. Sodium carbonate for boron production is obtained by bubbling CO₂ gas through an aqueous solution of NaOH, then gently heating to recover the solute.

5E.3 Quantitative LMF Materials Closure

The arguments presented in section 5E.2 demonstrate that a surprisingly simple system involving 18 elements and perhaps two dozen mineral species and process chemicals can probably achieve virtually 100% materials processing closure. Reagents necessary for electronics parts fabrication were included so that the lunar SRS has the materials needed to replicate its own computer and robot equipment. While the above is probably not the minimum size chemical processing plant that can retain closure, it is certainly one example of such a system. Other possibilities should be pursued in future research. Of course, once a growing seed reaches full adult size, it can install a whole new series of production equipment (say, for the recovery of platinum group metals) making possible a new range of capabilities that were unnecessary during the early growth/replication phases.

Quantitatively, in order to rigorously demonstrate complete materials closure it would be necessary to work through every chemical process described above, calculate the exact materials mass for every structure, robot, and other LMF device on an element-by-element basis, then verify that enough of each could be produced by the system. Such a detailed computation clearly lies beyond the scope of the present study. However, the team has attempted to estimate some of the most critical throughputs and analyze their anticipated effects upon total system closure. In this context, “closure” is a relationship between a given machine design and a particular substrate from which the machine’s chemical elemental constituents are to be drawn. Hence, the numerical calculation of closure requires a knowledge of the precise composition both of

the intended base substrate to be utilized as well as of the products which the SRS must manufacture from that substrate.

Following a method suggested by the work of Freitas (1980), the "extraction ratio" R (see sec. 5.3.6) is defined as the total mass of raw substrate material which must be processed (input stream) to obtain a unit mass of useful system output having the desired mass fractions of each required element (output stream). Consider the significance of the extraction ratio to the problem of materials closure. An $R = 1$ means that 1 kg of lunar regolith contains exactly the mass of all necessary LMF elements to manufacture a kilogram of desired output product. $R = 10$, on the other hand, means that 10 kg of lunar regolith must be processed to extract all of the elements required to make 1 kg of final product (see sec. 5.3.6).

For the purposes of the present study the team chose a trial value of $R = 40$ kg/kg. This choice is based on information from previous studies which suggests that 40 represents a good intermediate value between low closure and high complexity SRS materials designs.

On the one hand, for $R < 10$, the available mass fractions of certain critical but relatively rare elements such as H, C, B, and Cl fall too low to remain credible for a system requiring 100% closure. The missing materials must be imported as "vitamins" or the entire SRS must be redesigned to eliminate chemical processing and electronics using these elements. Examples of low closure models include the lunar processing factory designs proposed by Ho and Sobon (1979), $R = 1.7$; O'Neill (1976), $R = 1.7$; Phinney et al. (1977), $R = 1.2$; and Waldron et al (1979), $R = 1.1$. These systems are capable of extracting only half a dozen of the most abundant lunar elements and are not expected to achieve more than 60-90% materials closure.

On the other hand, for $R > 100$ the problem lies not in extracting rare elements but in processing them fast enough to meet a $T = 1$ year replication time deadline. For instance, Freitas (1980) gives an example of a high complexity system which could extract 84 elements from asteroidal material. For $R = 26,800$ the replication time is 500 years. It appears that $10 < R < 100$ is a plausible condition for 100% closure and 1-year replication in SRS. The maximum recoverable mass from lunar soil for each element assuming $R = 40$ is estimated in table 5.13. The question remains whether or not these quantities are adequate to achieve quantitative materials closure.

Certainly 100% closure exists for the six primary structural elements Al, Ca, Fe, Mg, O, and Si. Even if the entire 100-ton seed were comprised entirely of any one of these there is enough available of each. A similar argument may be made for Ti, since 80 tons in theory can be extracted. Steels and other alloys typically have 1% Mn, 0.2% Cr, and 0.1% C or less, which limits the total steel mass to 400 tons, 4000 tons, and 400 tons, respectively. Hence, alloy production will not be materials-limited by these three elements.

Carbon is also used in the boron and phosphorus production cycles. The mass of boron is so low that the carbon requirement is negligible in terms of mass. In the phosphorus cycle, 10 atoms of C are needed to cycle 4 atoms of P. Phosphorus is required as a dopant in silicon micro-electronic chip manufacture and in phosphoric acid which is used as a photolithography process chemical and which also appears during the halogen recovery cycle. At most, 40 kg of phosphorus are required, necessitating a carbon budget of 100 kg. This leaves more than 200 kg of carbon to account for losses and special uses such as CO₂ gas lasers.

Boron is used solely as a microelectronic silicon dopant; 4 kg of B can produce perhaps 10³-10⁴ kg of chips, more than enough for the 100-ton seed. A few kilograms of phosphorus (though high purity is required) will suffice for the same purpose, and the use of P as a process chemical should be more hydrogen-limited than phosphorus-limited because of the relative abundance of P in the lunar regolith.

According to calculations by Waldron et al. (1979), about 63 metric tons of H₂, F₂, and Na, half of which is F₂, are needed for an HF acid leach extraction facility having a total mass of about 823 tons. According to Criswell (personal communication, 1980) this model may

TABLE 5.13.— MAXIMUM MASS OF CHEMICAL ELEMENTS EXTRACTABLE FROM LUNAR SOIL, per year, for a 100-ton Seed with Extraction Ratio $R = 40$

Element	Typical global lunar abundance, kg element/kg soil	Maximum extractable mass, kg (R_{ch})(4)(100 tons)
Fe	0.10	400,000
Al	.07	280,000
Mg	.05	200,000
Ti	.02	80,000
Mn	.001	4,000
Cr	.002	8,000
C	.0001	400
Si	.2	800,000
O	.4	1,600,000
P	.0005	2,000
B	.000001	4
F	.0001	400
N	.0001	400
H	.00005	200
Na	.003	12,000
Cl	.00002	80
Ca	.07	280,000
S	.001	4,000
(Inert gases) ?	.00001	40

scale almost linearly down as low as 1 ton. The equivalents for a scaled-down 2.5-ton HF leach system are 90 kg F₂ and 100 kg of H₂ and Na. Sodium is about an order of magnitude more abundant than required, and fluorine does not appear to be a limiting factor even if recovery losses and spillages permit only 50% utilization of available stock. The supply of hydrogen, however, is crucial in achieving quantitative materials closure (see below). The 2.5-ton plant described above can output about 91 tons/year, which should be adequate to replicate a 100-ton seed once per year.

The primary use of nitrogen is in making NH₃ for the recovery of silica and as N₂ and HNO₃ for the production of microelectronic chips. The 400 kg N₂ given in table 5.13 is sufficient to prepare a maximum of 490 kg NH₃ or 1800 kg HNO₃. (These applications would require a maximum of 86 kg and 29 kg of H, respectively, hence are not seriously hydrogen-limited.) The amount of nitric acid seems more than sufficient, and the NH₃ can produce 100 to 1000 kg of silica, which should be adequate with recycling and provided losses can be held to a minimum.

Chlorine appears in the boron- and phosphorus-production cycles – in the former it is consumed and must be recycled; in the latter it is incorporated in a deliquescent compound and should not incur serious losses or require chemical recycling. The preparation of 1 mole of boron requires recycling 0.25 mole of Cl, hence (0.5)(4 kg)(35.45/10.8) = 6.6 kg of chlorine are needed to produce 4 kg of boron. As for the phosphorus cycle, 80 kg Cl produces 125 kg of deliquescent CaCl₂ which is capable of absorbing roughly its own weight in water. This should be sufficient with recycling (by simple heating) no more often than once a month on a T = 1 year schedule.

Sulfur is used primarily in the casting subsystem in the fabrication sector (about 600 kg required) and in the manufacture of sulfuric acid. This product is mass-limited about equally by the amounts of S and H available. The 4000 kg of sulfur can be used to prepare 12,000 kg H₂SO₄, and the 200 kg of hydrogen can make up 9800 kg of the acid. Since hydrogen also has many other uses, available S will be underutilized and perhaps 1 or 2 tons of H₂SO₄ reasonably can be produced. Is this enough? The main uses of sulfuric acid are in the recovery processes for B, P, F, and Cl, and in the preparation of silanes. The ratio of B:H₂SO₄ is about 4:1 moles, so to extract 4 kg B requires 9.1 kg acid. For phosphorus extraction, P:H₂SO₄ :: 2:3 moles, so (3/2)(98.1/31)(40 kg) = 190 kg H₂SO₄. For fluorine extraction, F:H₂SO₄ :: 2:1 moles, which requires (1/2)(98.1/19)(200 kg) = 516 kg acid. For chlorine extraction, Cl:H₂SO₄ :: 2:1 moles, which requires (1/2)(98.1/35.45)(80 kg) = 110 kg H₂SO₄. The quantity of silane needed for microelectronics processing is expected to be minimal, so it appears that adequate supplies of sulfuric acid can be made available with reasonable loss factors to sustain the growth of a fully autonomous LMF on a sulfur budget of about 1500 kg.

The only critical element appears to be hydrogen. This criticality is not especially peculiar to the present design, but rather stems from the relative scarcity of the element in lunar materials and the many chemical processing applications to which it may be put. Any hydrogen-chemistry-based materials processing system will encounter similar difficulties. The 200 kg of available hydrogen could make the maximum quantities of H-bearing compounds listed in table 5.14, although the available hydrogen must be spread among these applications as required with lower masses in

TABLE 5.14.— HYDROGEN-LIMITED MATERIALS
PROCESSING REAGENTS

LMF reagent	Estimated LMF requirements, kg	Maximum, limited only by hydrogen available, kg	Fraction of available hydrogen required
NH ₃	300	1,100	0.273
H ₂ SO ₄	1,000	9,800	.102
SiH ₄	100	1,600	.0625
HNO ₃	600	12,600	.0476
HF	190	4,000	.0475
H ₃ PO ₄	100	6,500	.0154
NaOH	100	8,000	.0125
H ₂ O	700	1,800	.3895
(5% losses)	---	---	.05
			1.000

each case. Expected requirements of hydrogen-bearing reagents are listed in table 5.14. Although these calculations are highly sensitive to the assumptions employed, closure may be achieved if an allowance of 5% for spillage and other losses is adequate. Obviously a major leak could seriously jeopardize a hydrogen-based LMF system.

If hydrogen supply remains a critical problem it may become necessary to: (1) redesign the processing system for greater hydrogen frugality, (2) select a slightly higher extraction ratio R to permit recovery of a greater mass of H, (3) locate and "mine" particular lunar soils extra-rich in H, such as the suggested use of ilmenite as a hydrogen "ore" (Green, personal communication, 1980), (4) accept a replication time longer than 1 year, or (5) relax the 100% closure requirement and permit resupply of small amounts of hydrogen "vitamin" from Earth.

5E.4 Sector Mass and Power Estimates

The overall functional layout of the LMF chemical processing sector is illustrated in figure 5.16. The operations flowsheet shows that there are 13 components within the sector: (1) input hopper, (2) electrophoretic separators, (3) P/F/Cl extractors, (4) boron extractors, (5) sodium extractors, (6) volatiles extractors, (7) HF acid leach system, (8) freon producer, (9) ammonia producer, (10) silane producer, (11) nitric acid producer, (12) sulfuric acid producer, and (13) output hopper.

Mass and power consumption for LMF materials processing may be estimated by comparison with other automated chemical processing designs that have been considered, and which are summarized in table 5.15. For $R = 40$, a 100-ton/year (self-replicating) output demands a 4000-ton/year raw materials input, or 0.13 kg/sec. Taking the range of values given in table 5.15, Sector mass should lie within 18,200 to 78,000 kg. Similarly, the estimated power requirements range from 455 kW up to 10.9 MW, although in this case the lower values seem more appropriate.

Dry thermal chemical processing techniques are associated with very high energy requirements, whereas lower values are found in wet chemistry processes – of which the HF acid leach selected for the present design is an example.

5E.5 Information and Control Estimates

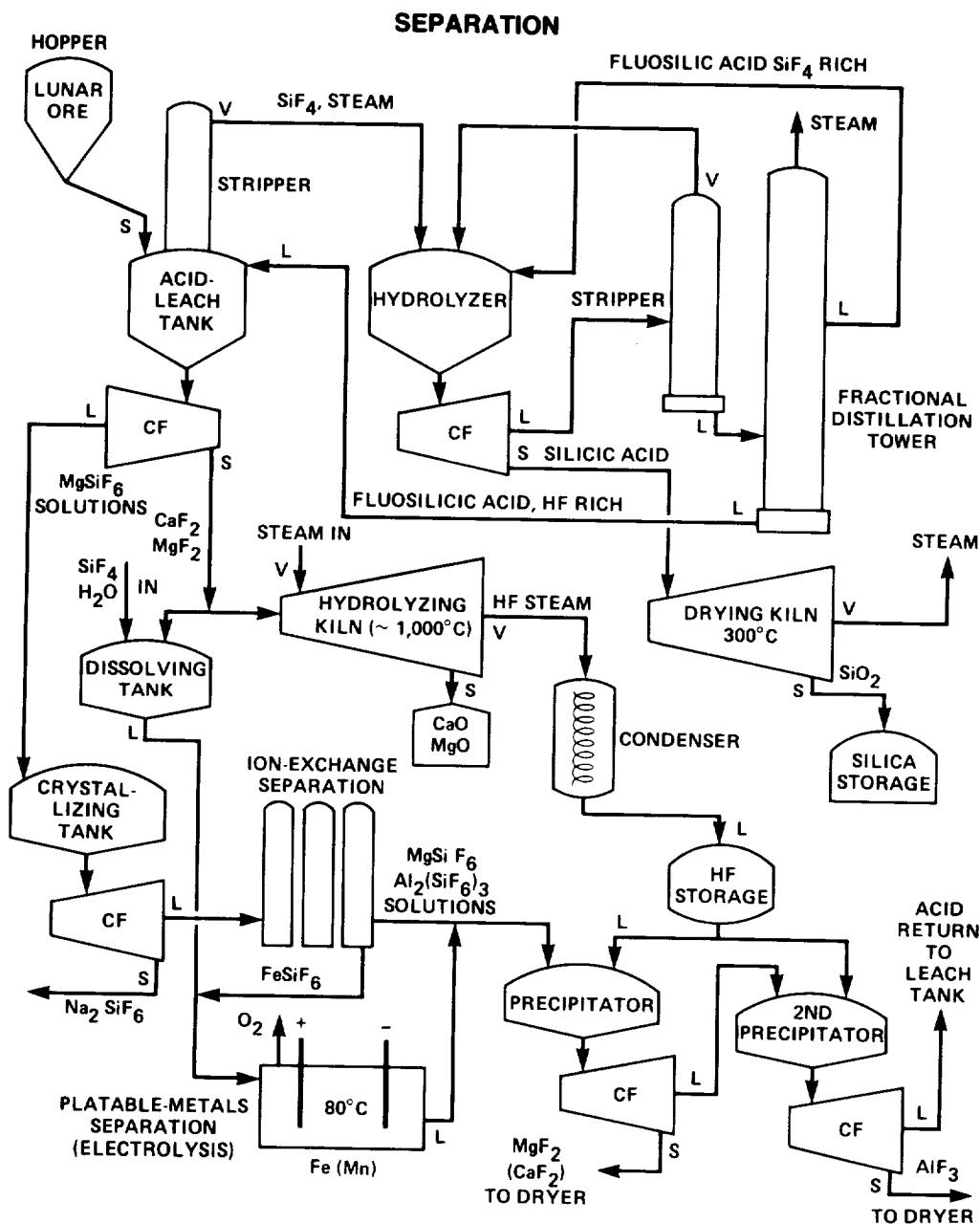
Probably the most complex of the 13 sector components which appear in figure 5.41 is the HF acid leach system. From figure 5.41 this appears to consist of 34 component subsystems such as "precipitator," "dissolving tank," "fractional distillation tower," "centrifuge/filter," "Castner cell," etc. Each subsystem performs a single well defined task. In addition, there are 111 nodes (each denoting a point of connection of a pipe or supply line to another pipe or to a subsystem) each requiring at least one valve and valve control mechanism. At each valve there must be a number of sensors indicating valve position (open, closed, fractionally open), valve malfunction and cause (if simple), and volume or velocity of flow of material through the valves. Interface with actuators and reportage to the subsystem subcomputer are additional requirements.

Assuming each valve can be automated with a 1K computer allocation, and each subsystem can be automated with a 10K memory allocation, then the total computer capability required for continuous leach system operation is $(1)(111) + (10)(34) = 451\text{K}$ which is 7.2×10^6 bits using 16-bit words. This should be sufficient to handle normal system operations and troubleshooting, although actual repair must be done by mobile repair robots. Also, any catastrophic malfunctions such as pipe ruptures, jammed fixtures, leaks, heating element burnouts or explosions must be diagnosed and corrected by the mobile repair robots.

The chemical processing sector looks not to be a place where complicated new automation techniques will be

TABLE 5.15.— COMPARISON OF CHEMICAL PROCESSING PLANT MASSES AND POWER REQUIREMENTS FROM PREVIOUS RELATED STUDIES

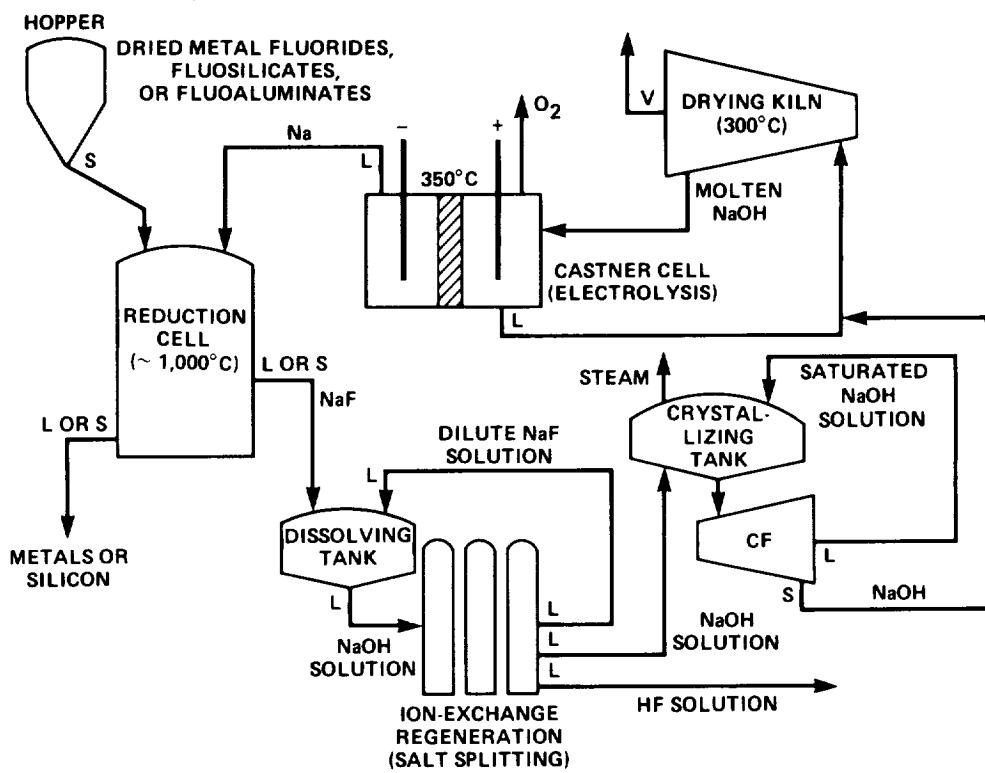
Source	kg plant kg/sec input	kg chemicals kg/sec input	kg chemicals kg plant	Power, W kg plant
Johnson & Holbrow (1977) Al-processing plant	3.1×10^5	2.6×10^4	8.5×10^{-2}	25
Phinney et al. (1977) carbo/silico-thermic plant	3.3×10^5	---	---	140
Waldron, Erstfeld, & Criswell (1979) HF acid-leach metal extraction plant	$3-6 \times 10^5$	3.5×10^5	4.1×10^{-1}	37
O'Neill, Driggers, & O'Leary (1980) reference design figure	3.6×10^5	---	---	---
O'Neill (1976) carbothermic Space Manufacturing Center	5.2×10^5	---	---	100
Criswell (1980)	3.6×10^5	---	---	100
Vajk et al. (1979) Space Manufacturing support requirements	1.4×10^5	---	---	---



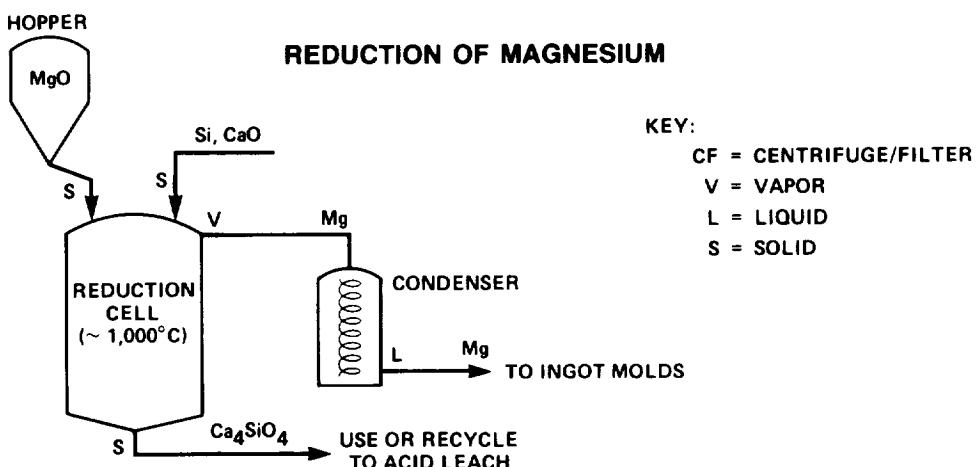
1. $x\text{MO} \cdot \text{SiO}_2 + (4 + 2x) \text{HF} = x\text{MF}_2 + \text{SiF}_4 \text{ (aq)} + (2 + x) \text{H}_2\text{O}$
- 1'. $x\text{MO} \cdot \text{SiO}_2 + (5 + 2x) \text{HF} = x\text{MF}_2 + \text{HSiF}_1 \text{ (aq)} + (2 + x) \text{H}_2\text{O}$
2. $\text{SiF}_4 \text{ (aq)} + n\text{H}_2\text{O} = \text{SiF}_4 \text{ (v)} + n\text{H}_2\text{O} \text{ (v)}$
- 2'. $\text{HSiF}_5 \text{ (aq)} + n\text{H}_2\text{O} = \text{SiF}_4 \text{ (v)} + \text{HF} \text{ (aq)} + n\text{H}_2\text{O} \text{ (v)}$
3. $(1 - y) [\text{SiF}_4 \text{ (v)} + 4\text{H}_2\text{O} = \text{Si(OH)}_4 + 4\text{HF}]$
- 3a. $(1 - y) [\text{SiF}_4 \text{ (v)} + 2\text{H}_2\text{O} = \text{SiO}_2 + 4\text{HF}]$
4. $(1 - y' - z) [x\text{MF}_2 + x\text{H}_2\text{O} = x\text{MO} + 2x\text{HF}]$
5. $y [\text{SiF}_4 + 4\text{Na} = \text{Si} + 4\text{NaF}]$
6. $y' [x\text{MF}_2 + 2x\text{Na} = x\text{M} + 2x\text{NaF}]$

Figure 5.41.— Flowsheet and process equations for the HF acid-leach process.

REDUCTION OF METALS OTHER THAN MAGNESIUM



REDUCTION OF MAGNESIUM



KEY:

CF = CENTRIFUGE/FILTER
V = VAPOR
L = LIQUID
S = SOLID

7. $z [x\text{MF}_2 + x\text{SiF}_4 \text{ (aq)} = x\text{MSiF}_6 \text{ (aq)}]$
 8. $z [x\text{MSiF}_6 \text{ (aq)} + x\text{H}_2\text{O} + \text{electrical energy} = (x/2)\text{O}_2 + x\text{M} + x\text{H}_2\text{SiF}_6]$
 - 8a. $z [x\text{MSiF}_6 \text{ (aq)} + \text{M}'\text{SO}_3\text{R}^+ = x\text{M}'\text{SiF}_6 \text{ (aq)} + x\text{MSO}_3\text{R}^+]$
 9. $m \text{NaF} + m\text{R}^+\text{OH} = m\text{NaOH} + m\text{R}^+\text{F}$
 - 9a. $m \text{NaF} + (m/2) \text{Ca(OH)}_2 = m\text{NaOH} + (m/w) \text{CaF}_2$
 10. $m \text{NaOH} + \text{electrical energy} = m\text{Na} + (m/2)\text{O}_2 + (m/2)\text{H}_2\text{O}$
 11. $(1 - y) [\text{Si(OH)}_4 = \text{SiO}_2 + 2\text{H}_2\text{O}]$
- R^+ = ion-exchange resin
 $m = 4y + 2xy'$

Figure 5.41.—Concluded.

required (Ayres, 1952; Foster, 1963; Kallen, 1961; Luke, 1972). Component processes are all state-of-the-art. There is no need for visual processing during normal operations, and procedures are standardized so no expert systems or judgmental algorithms are required beyond the simple integration of well defined sensor data. All operations will probably be hard-automated, and materials will be moved about almost entirely in sealed vessels. If there is need for additional transport within the system a materials transit network may be erected using metal or basalt tracks, electric motors and small carrier vehicles.

If each of the 13 sector components is as complex as the HF acid leach system (certainly a gross overestimate), then the total computer control capability required is about 6 megabytes or 9.4×10^7 bits using 16-bit words. The information needed to describe the sector sufficiently for purposes of self-replication must also be estimated. Assuming

that each HF leach subsystem requires 2×10^6 bits for complete description (about a 200-page printed book, or 80,000 English words), and that each valve requires about half as much (say, 100 book pages), then the total for the HF leach system is 1.8×10^8 bits. Again conservatively multiplying by 13, the total information to describe the sector components is 2.3×10^9 bits. If sector equipment is distributed across a floor space of 5000 m^2 , then to store a map with 1-cm placement resolution requires a memory capacity of 8×10^8 bits assuming one 16-bit word to describe the nominal status of each 1 cm^2 of platform space. Note that large empty areas convey useful information and must be mapped, since they may be used for traffic routes, repair routes, temporary warehousing, etc. The total information for sector replication is thus about 3.1×10^9 bits. The information control budget is 9.4×10^7 bits.

SE.6 References

- Allcock, Harry R.: Inorganic Polymers. *Scientific American*, vol. 230, March 1974, pp. 66-74.
- Arnold, J. R.; Criswell, D. R.; and Waldron, R. D.: Progress Report on Experimental Program to Develop HF Acid Leach Process for Refining Lunar Soils. Paper presented at the 5th Princeton/AIAA/SSI Conference on Space Manufacturing, 18-21 May 1981, Princeton, NJ.
- Ayres, Eugene: An Automatic Chemical Plant. *Scientific American*, vol. 187, September 1952, pp. 82-88.
- Criswell, David R.: Extraterrestrial Materials Processing and Construction. NASA CR-158870, vol. 1, 1978, and vol. 2, 1980.
- Dunning, Jeremy D.; and Snyder, Robert S.: Electrophoretic Separation of Lunar Soils in a Space Manufacturing Facility. Paper presented at the 5th Princeton/AIAA/SSI Conference on Space Manufacturing, 18-21 May 1981, Princeton, NJ.
- Foster, David: Modern Automation. Sir Isaac Pitman and Sons, Ltd., London, 1963.
- Freitas, Robert A., Jr.: A Self-Reproducing Interstellar Probe. *J. British Interplanet. Soc.*, vol. 33, July 1980, pp. 251-264.
- Ho, Darwin; and Sobon, Leon E.: Extraterrestrial Fiberglass Production Using Solar Energy. In *Space Resources and Space Settlements*, John Billingham, William Gilbreath, Brian O'Leary, and B. Gossett, eds. NASA SP-428, 1979, pp. 225-232.
- Johnson, Richard D.; and Holbrow, Charles, eds.: *Space Settlements: A Design Study*, NASA SP-413, 1977, 185 pp.
- Kallen, Howard P., ed.: *Handbook of Instrumentation and Controls*. McGraw-Hill Book Co., New York, 1961.
- Lee, Stuart M.: Lunar Building Materials - Some Considerations on the Use of Inorganic Polymers. In *Space Resources and Space Settlements*, John Billingham, William Gilbreath, Brian O'Leary, and B. Gossett, eds. NASA SP-428, 1979, pp. 233-241.
- Luke, Hugh D.: Automation for Productivity. John Wiley and Sons, New York, 1972.
- Oldham, William G.: The Fabrication of Microelectronic Circuits. *Scientific American*, vol. 237, September 1977, pp. 110-128.
- O'Neill, Gerard K.: Engineering a Space Manufacturing Center. *Astronautics and Aeronautics*, vol. 14, October 1976, pp. 20-28, 36.
- O'Neill, Gerard K.; Driggers, G.; and O'Leary, B.: New Routes to Manufacturing in Space. *Astronautics and Aeronautics*, vol. 18, October 1980, pp. 46-51.
- Phinney, W. C.; Criswell, D. R.; Drexler, E.; and Garmirian, J.: Lunar Resources and Their Utilization. In *Space-Based Manufacturing from Nonterrestrial Materials*, Gerard K. O'Neill, ed., AIAA, New York, 1977, pp. 97-123. (Progress in Astronautics and Aeronautics, vol. 57.)
- Vajk, J. P.; Engel, J. H.; and Shettler, J. A.: Habitat and Logistic Support Requirements for the Initiation of a Space Manufacturing Enterprise. In *Space Resources and Space Settlements*, John Billingham, William Gilbreath, Brian O'Leary, and B. Gossett, eds., NASA SP-428, 1979, pp. 61-83.
- Waldron, R. D.; Erstfeld, T. E.; and Criswell, D. R.: The Role of Chemical Engineering in Space Manufacturing. *Chemical Engineering*, vol. 86, 12 February 1979, pp. 80-94.
- Williams, Richard J.; and Jadwick, James J.: *Handbook of Lunar Materials*, NASA RP-1057, 1980.
- Zachary, Wm. B.: A Feasibility Study on the Fabrication of Integrated Circuits and Other Electronic Components in Space Utilizing Lunar Materials. Paper presented at the 5th Princeton/AIAA/SSI Conference on Space Manufacturing, 18-21 May 1981, Princeton, NJ.

APPENDIX 5F

LMF PARTS FABRICATION SECTOR

There are two distinct classes of fabrication production machines in any general-product self-replicating system — parts or “bulk” fabrication and electronics or microcircuit fabrication. Appendix 5F is concerned exclusively with LMF subsystems required for bulk manufacturing. Microelectronics production in space manufacturing facilities is considered in section 4.4.3 and is the subject of Zachary (1981); estimated mass of this component of the original LMF seed is 7000 kg, with a power draw of perhaps 20 kW to operate the necessary machinery (Meylink, personal communication, 1980).

SF.1 Overall Design Philosophy

The plausibility of both qualitative and quantitative materials closure has already been argued in appendix 5E. A similar line of reasoning is presented here in favor of a very simple parts fabrication system, to be automated and deployed in a self-replicating lunar manufacturing facility. To rigorously demonstrate parts closure it would be necessary to compile a comprehensive listing of every type and size of part, and the number required of each, comprising the LMF seed. This list would be a total inventory of every distinct part which would result if factory machines were all torn down to their most basic components — screws, nuts, washers, rods, springs, etc. To show 100% closure, it would then be necessary to demonstrate the ability of the proposed automated parts fabrication sector to produce every part listed, and in the quantities specified, within a replication time of $T = 1$ year, starting from raw elemental or alloy feedstocks provided from the chemical processing sectors.

Unfortunately, such a detailed breakdown and analysis probably would require tens of thousands of man-hours even for the simplest of systems. Not only is the seed not a simple system, but the present baseline design is not conveniently amenable to this sort of detailed analysis. Thus, a completely rigorous demonstration of parts closure is beyond the scope of the present study.

However, it is possible to advance a plausibility argument based upon a generalized parts list common to many complicated machines now in use in various terrestrial applications (Spotts, 1968; von Tiesenhausen, unpublished Summer Study document, 1980). Although machines designed for construction and use in space may employ radically different components than their terrestrial coun-

terparts, to a first approximation it may be assumed that they will be comprised generally of the same kinds of parts found in commonplace machines on Earth such as bolt, nut, screw, rivet, pulley, wheel, clutch, shaft, crank, rod, beam, wire, plate, disk, bushing, cable, wedge, key, spring, gasket, seal, pipe, tube, and hose. If this is valid, then a showing that all parts classes in the general parts list can be manufactured by the proposed automated fabrication system may serve as a valuable plausibility argument in favor of parts closure for that system.

The achievement of a sound design which incorporates the advantages of maximum economy in manufacture and functional requirements of a part is dependent upon the designer's ability to apply certain basic rules (Yankee, 1979). There are four recognized rules, equally applicable to terrestrial factories and lunar replicating machine systems, as follows:

(1) Design all functional and physical characteristics for greatest simplicity. As a general principle, service life of a part is greatly increased when design of that part is both simple and sturdy (“robust”). Performance is more predictable and costs (money, build time, repair time) are lower for simpler parts.

(2) Design for the most economical production method. The particular production design selected should, if possible, be optimized for the part or set of parts the system must produce. The production of scrap (input/output ratio) is one valuable index by which optimality may be compared. This factor is relatively simple to evaluate where only one part is manufactured. In multipart production lines the problem is far more complicated, since each of the many parts may be expected to have dissimilar optima. Consequently, only the production of the entire system can be truly optimum.

(3) Design for a minimum number of machining operations. All types of costs are lower when fewer operations are required to produce a part according to specifications. The greatest savings result when the number of separate processing operations necessary to complete a part is reduced. Multiple operations which can be combined into fewer operations, or functionally similar parts requiring fewer production steps, should be changed in a design. “Needless fancy or nonfunctional configurations requiring extra operations and material” should be omitted from the design (Yankee, 1979).

(4) Specify finish and accuracy no greater than are actually needed. If a part will adequately serve its intended purpose at some lower level of accuracy of machining than is technologically possible, then cheaper, simpler production processes may be used which make closure easier to attain. The specification of needlessly close tolerances and an unreasonable degree of surface finish invariably results in a low part production rate, extra operations, high tooling costs, and high rejection rates and scrap losses (Yankee, 1979).

5F.2 Selection of Basic Production Processes

A wide variety of fabrication processes is available using current technology, each of which is optimum for the production of one or more classes of parts or in certain specialized applications (see table 4.17). From inspection of table 4.10 it is reasonable to conclude that there are perhaps only 300 fundamentally distinct fabrication techniques in widespread use today. Ultimately, the LMF factory in production phase may be called upon to perform many if not all of these functions. However, most may be unnecessary for initial system growth or replication. Indeed, optimum seed design should permit maturation to adulthood in the minimum time with the fewest parts using the fewest machine operations possible.

The team concluded that four basic processes — plaster casting, vapor deposition, extrusion, and laser machining — are probably sufficiently versatile to permit self-replication and growth. These four techniques can be used to fabricate most parts to very high accuracy. Plaster casting was selected because it is the simplest casting technique for producing convoluted parts as well as flat-surface parts, to an acceptable level of accuracy. (A number of alternatives have already been reviewed in app. 4B.) The laser machining tool can then cut, weld, smooth, and polish cast parts to finer finishes as required. Vapor deposition is the least complicated, most versatile method of producing metal film sheets to be used as the manufacturing substrate for microelectronics components, mirrors or solar cells, or to be sliced into narrow strips by the laser for use as wire. The extruder is used to produce thread fibers of insulating material, presumably spun basalt drawn from a lunar soil melt as described in section 4.2.2.

5F.3 Casting Robot

The casting robot is the heart of the proposed automated fabrication system. It is responsible for producing all shaped parts or molds from raw uncut elemental materials. The moldmaking materials it works with are of two kinds. First, the casting robot receives thermosetting refractory cement with which to prepare (a) molds to make iron alloy parts, (b) molds to make iron molds to cast basalt parts (but not aluminum parts, as molten aluminum tends to

combine with ferrous metal), and (c) individual refractory parts. Second, the robot receives hydrosetting plaster of Paris with which to prepare (a) molds to cast aluminum parts and (b) substrates for the vacuum deposition of aluminum in sheets. According to Ansley (1968), small castings using nonferrous metals (aluminum, magnesium, or copper alloys) may be produced using plaster molds with a surface finish as fine as 2-3 μm and an accuracy of ± 0.1 mm over small dimensions and ± 0.02 mm/cm across larger surfaces (a drift of 2 mm over a 1 m^2 area).

Traditionally, the plaster casting technique requires a split metal pattern in the shape of the object to be cast. This pattern is used to make a hollow mold into which molten metal is poured, eventually solidifying to make the desired part. Alternatively, patterns may be manually carved directly into the soft, setting plaster, after which metal again is poured to obtain the desired casting.

The casting robot should have maximum versatility. It will have access to a template library located within its reach, containing samples of each small or medium-sized part of which the LMF is comprised. If the SRS seed is designed with proper redundancy, it will use the fewest number of different kinds of parts and there will be large numbers of each kind of part. Assuming that on average there are 1000 pieces of each type of part in the original LMF architecture, then the total template library has a mass of only 100 tons/1000 = 100 kg and there are perhaps a thousand different kinds of parts (see below).

In addition, the casting robot is equipped with shaping and carving tools which can create any desired shape in the slowly hardening plaster. (Pure gypsum plaster hardens in 6-8 min after water is added, but this setting time may be extended up to 1-2 hr by adding lime, CaO, to the emulsion. Setting time is also temperature-dependent.) The shaping tools may represent perhaps 100 specific shapes and sizes and should also include at least a dozen "universal" carving instruments.

To make a given part, the robot searches its template library to see if it has a convenient pattern already in stock. If so, it uses the pattern to form the mold; if not, it uses its many tools to carve out a mold of the appropriate size and shape. Plaster of Paris is a hydraulic cement — it sets with the addition of water. Refractory cement is thermosetting and has to be heated to 1300-1400 K in a kiln to set the mold.

Water used to make the plaster molds cannot remain liquid in the lunar vacuum. Thus, the casting robot plaster system must be pressurized, probably with nitrogen gas to permit the pouring of molten aluminum. The triple point of water (the bottom end of its liquid phase) occurs at 608 Pa, but a 1.3×10^4 Pa atmosphere (16 kg N₂ to fill a 100 m³ working volume) prevents water from boiling off up to about 323 K.

Mass requirements for plaster molding are estimated by assuming that 10% of the volume of each mold contains a

useful part (10% mold volume utilization). If the mean density of LMF parts (mostly aluminum) is taken as 3000 kg/m^3 , and the entire plaster mass is recycled once a day, then for a 100-ton seed the robot must have 2600 kg (0.91 m^3) of plaster compound (gypsum, or calcium sulfate) on hand. To hydrate (set) this much plaster requires 483 kg of water, an amount of precious hydrogen already allowed for in LMF materials estimates presented in appendix 5E. Availability of sulfur is not a concern, since 2600 kg of plaster requires only 475 kg of S. Terrestrial plasters commonly have a small amount of strengthener added, but in the lunar application this substance should be designed to be recyclable or must be eliminated altogether.

Plaster casting is not the only way to make parts in a growing, self-replicating factory, but it is definitely one of the easiest both conceptually and in common industrial practice. Plaster methods are especially well suited for producing parts with hard-to-machine surfaces such as irregularly shaped exterior surfaces and in applications where a superior as-cast surface is important (Yankee, 1979). Plaster molded products commonly include aluminum match plates, cores and coreboxes, miscellaneous parts for aircraft structures and engines, plumbing and automotive parts, household appliances, hand tools, toys, and ornaments. The technique is good for manufacturing parts requiring high dimensional accuracy with intricate details and thin walls ($\geq 0.5 \text{ mm}$). Castings of less than 0.45 kg and as massive as 11,350 kg have been made on Earth. Commercially, when compared to aluminum die casting, plaster mold casting is considered economical if 1000 parts or less are produced, although production runs up to 2000 parts may also be considered economical if the parts are especially complex.

Refractories. Refractories are materials which remain useful at very high temperatures, usually 1500–2300 K. They are employed primarily in kilns, blast furnaces, and related applications. In the lunar SRS refractories are needed as linings for drying kilns, roasting ovens, in the production of iron molds (to cast basalt parts) and iron parts, and also as material for special individual parts such as nozzles and tools which must operate at very high temperatures.

Refractories are usually, but not always, pure or mixtures of pure metal oxides. Tables in Campbell and Sherwood (1967) list the most important simple and complex refractory substances which LMF designers might choose. There are a few basic considerations, such as vapor pressure. For instance, although magnesia melts at 3070 K and has a useful operating temperature to about 2700 K in oxidizing atmospheres, it cannot be used in a vacuum at temperatures above about 1900 K because of volatilization (Johnson, 1950). Similarly, zinc oxide volatizes above 2000 K and tin oxide sublimes excessively at 1780 K even in an atmosphere.

Refractory bodies are fabricated from pure oxides by powder pressing, ramming, extruding, or slip casting. The last of these is the simplest, but requires a very fine powder. This powder is normally prepared by ball milling. Steel mills and balls are used, and the iron is later separated by chemical means. For simplicity in LMF design, the iron alloy powder inevitably mixed with the milled product can be removed by magnetic separation.

High-alumina cements and refractories may be the best option for lunar manufacturing applications. Alumina is a major product of the HF acid leach system in the chemical processing sector, and is capable of producing castable mortars and cements with high utility up to 2100 K (Kaiser, 1962; Robson, 1962). It will permit casting iron alloys, basalts, and low melting point metals such as Al and Mg. Unfortunately, it will not be possible to cast titanium alloys in this fashion, since in the liquid state Ti metal is very reactive and reduces all known refractories.

Alumina can be slip-cast from water suspensions. The oxide powder is first ball-milled as described above to 0.5–1.0 μm , then deflocculated by the addition of either acid (HCl) or base (NaOH), and finally the refractory body is developed by absorbing the liquid in a porous mold (plaster of Paris may be used with a base deflocculant). Gravity and hydrodynamic pressure of the flowing liquid produce a well compacted body of the suspended particles (Campbell and Sherwood, 1967). A fairly comprehensive review of alumina and alumina ceramics may be found in Gitzen (1966).

Metal alloys. A number of different metal alloys will be required for casting various parts and molds. Different alloys of iron may be chosen for the steel balls for ball milling, the basalt casting molds, and the individual part that might be comprised of steel or iron. Various aluminum alloys may be selected for parts, whereas pure metal is required for vapor deposition processes. Castable basalt may require fluxing but otherwise is a fairly straightforward melt.

Metallurgical duties are performed at the input terminus of the fabrication sector. Mobile chemical processing sector robot carriers dump measured quantities of metals and other substances into cold fabrication sector input hoppers (made of cast basalt and perhaps stored under a thin oxygen atmosphere to preclude vacuum welding). Mixing is accomplished by physical agitation, after which the contents are fed into a solar furnace to be melted. If net solar efficiencies are roughly the same as for the 5 kg capacity induction furnace (output 30 kg/hr) described in the MIT space manufacturing study (Miller and Smith, 1979), then about 30 kW of power are required which may be drawn most efficiently from a large collector dish roughly 6 m diam. There are at least three hopper/furnace subsystems required – a minimum of one each for iron, basalt, and aluminum alloys. Possibly another would be needed for

magnesium alloys, and several more to forestall contamination between disparate batches, but three is the absolute minimum requirement.

Parts manufacturing. The construction of a machine system as complex as a lunar SRS will require a great many individual parts which vary widely in mass, shape, function, and mode of assembly. If a complete parts list were available for the seed, then the manufacturing steps for each could be explicitly specified, precise throughput rates and materials requirements given, and closure demonstrated rigorously. Unfortunately, no such list is yet available so the team was forced to resort to the notion of the "typical part" to gain some insight into the performance which may be required of the casting robot.

Modern aircraft have about 10^5 parts and weigh up to about 100 tons, for an average of 1 kg/part (Grant, 1978). The average automobile has 3000–4500 parts, depending on its size and make, so the typical part weights perhaps 0.5 kg (Souza, personal communication, 1980). A study performed for General Motors concluded that 90% of all automotive parts weigh 2 kg or less (Spalding, personal communication, 1980). A design study by the British Interplanetary Society of a very advanced extrasolar space probe assumed a figure of 9 kg per typical part (Grant, 1978). Conservatively estimating that the typical LMF part is only 0.1 kg, then a 100-ton seed is comprised of roughly a million parts.

If most components may be made of aluminum or magnesium then the density of the typical part may be taken as about 3000 kg/m³, so the characteristic size of the typical part is $(0.1/3000)^{1/3} = 3.2$ cm. This result is consistent with Souza's (personal communication, 1980) suggestion that the average automobile part could be characterized as "roughly cylindrical in shape, an inch in length and half an inch in diameter." The casting robot must be able to cast all 10^6 parts within a replication time $T = 1$ year. If the casting bay is only 1 m² in horizontal extent, and only 10% of that area is available for useful molding, then each casting cycle can prepare molds for 0.1 m² of parts. The characteristic area of the typical part is $(0.1/3000)^{2/3} = 0.001$ m², and dividing this into the available area gives 100 parts/casting cycle as the typical production rate for the robot. To produce 10^6 parts/year the casting robot must achieve a throughput rate of 10,000 cycles/year, or about 52 min/cycle. This in turn implies that the system must be able to carve or mold at an average rate of 30 sec/part. Since most parts should be simple in form or will have patterns available, this figure appears feasible. After the casting robot makes molds for the parts, the molds are filled with molten aluminum alloy. The metal hardens, the mold is broken, and the pieces are recycled back into plaster of Paris; the aluminum parts formed in the mold are conveyed to the laser machining and finishing station.

Very thin sheets of aluminum also are required in various applications, among them solar cell manufacture, production of microelectronic components, and solar furnace mirror surfaces. Extrusion, rolling, and direct casting were considered and rejected on grounds of lack of versatility and complexity. Vapor deposition, currently used in industry to apply coatings to surfaces and to prepare thin sheets of aluminum and other substances, was tentatively selected both because of its tremendous versatility (any curved surface may be coated) and because it is state-of-the-art technology. The major problems with the process in terrestrial applications are maintenance of the vacuum and high energy consumption, neither of which are factors on the lunar surface or in an orbital environment.

Plaster molds to be surfaced are passed to a laser honing station where they are finished to any desired accuracy, after which they move to the vapor deposition station and are coated with appropriate metals or nonmetals to the requisite thickness. The process is expected to proceed much as described by Miller and Smith (1979). The plaster mold is then removed and recycled, and the fabricated aluminum sheet is passed on to the electronic fabrication system or is sliced into wires by a fine cutting laser (Miller and Smith, 1979).

Mass throughput rates for this system appear adequate. Assuming that 10^4 m² of solar cells are needed for the original seed (Freitas, 1980) and that the casting bay is about 1 m² in area, then for $T = 1$ year the required deposition rate to produce 0.3 mm thick aluminum sheet is $r_d = (10^4 \text{ m}^2 \text{ solar cells/year})(3 \times 10^{-4} \text{ m thick/sheet}) (1 \text{ sheet/m}^2)(1 \text{ year}/5.23 \times 10^5 \text{ min})(10^6 \mu\text{m/m}) = 5.7 \mu\text{m/mm}$. State-of-the-art deposition rates attained for aluminum commercially are about 50 $\mu\text{m/min}$ (Miller and Smith, 1979), nearly an order of magnitude higher than required. (The above throughput rate would also be equivalent to 1 m/sec of 0.3 mm aluminum wire production if cutting and wrapping can keep pace with deposition.) Cycling time is about 52 min/sheet. Following Johnson and Holbrow (1977), a heat of vaporization of 10^7 J/kg for 10^4 solar cells each made of 0.3 mm Al of density 3000 kg/m³ requires a continuous power draw of only 2.9 kW, which can be supplied by a small solar collector mirror 2 m diam.

A small number of LMF parts are expected to be made of cast basalt – fused as-found lunar soil perhaps with fluxing agent additives. Most parts will probably be aluminum because Al is an easily worked metal with high strength, low density (hence supporting structures need not be large), and relatively low melting point (hence is easily cast). The major advantages of basalt are its easy availability, its tolerance of machining, good compressive strength, and high density in some uses. Anticipated applications include machine support bases, furnace support walls, robot manipulator tools (to avoid vacuum welding), and other special parts where weight is not a problem. Because

plaster fuses at 1720 K – very near the melting point of basalt – and loses its water of crystallization around 475 K, it cannot be used to make basalt castings. Iron molds cast from refractory templates are required; they may be reused or recycled as necessary.

Another principal application for basalt is as an insulating fiber. Spun basalt threads can be used to wrap electrical conductors to provide insulation, woven to produce “mineral fabrics” as filler to strengthen cements, shock-absorbing resilient packing material, filters and strainers for materials processing, or as thermal insulation or to prevent cold welding of metals (Green, unpublished Summer Study document, 1980). The technology for producing spun basalt products (Kopecky and Voldan, 1965; Subramanian and Kuang-Huah, 1979), basalt wool, and drawn basalt fibers (Subramanian et al., 1975) is well established commercially and customarily involves extrusion or simple mechanical pulling from a melt (see sec. 4.2.2).

Ho and Sobon (1979) have suggested a design for a fiberglass production plant for the lunar surface using a solar furnace and materials obtained from lunar soil (anorthite, silica, alumina, magnesia, and lime). The entire production facility has a mass of 111 metric tons and a power consumption of 1.88 MW, and produces 9100 metric tons of spun fiberglass per year. Assuming linear scaling, the production for the replicating LMF of even as much as 10 tons of fiberglass thread would require a production plant of mass 122 kg and a power consumption of 2.1 kW (a 2-m solar collector dish).

A small number of LMF parts will also be made of iron (from refractory molds) and refractory cements (carved

directly from ceramic clay by the casting robot) in order to take advantage of the special properties of these substances. The total mass of such items is expected to be relatively low. Used refractory molds may be fed to the ball mill and recycled if necessary.

5F.4 Laser Machining and Finishing

The plaster casting parts manufacturing technique was chosen in part because of its ability to produce ready to use “as-cast” components. Thus, it is expected that the majority of parts will require little reworking, machining, or finishing. A small fraction, perhaps 10%, of all lunar SRS parts may require more extensive machining. A laser machining system was selected for this function in the LMF. The characteristic circumference of the typical part is $3.14(0.1/3000)^{1/3}$ or about 10 cm. If surface articulations cause an increase by a factor of ten in the total average path length that must be machined, then the mean operating speed of the laser system must be (10^6 parts/year) (10% machinables)(0.1 m/part)(10 m path/m circum.) (1 year/ 8722 hr) = 11.5 m/hr. Table 5.16 compares the performances of several different types of lasers, and table 5.17 gives specific performance parameters for high-power gas lasers used in industry for welding (butt, lap, corner, and edge) and for cutting. Inspection of these values suggests that a 5-10-kW continuous-wave (CW) carbon dioxide laser should be able to weld and cut “typical parts” with characteristic dimensions up to 3 cm at the required throughput rate.

TABLE 5.16.— CHARACTERISTICS AND PERFORMANCE OF VARIOUS LASERS COMMONLY USED FOR WELDING (Acharekar, 1974)

Laser	Operation	Pulse length, msec	Pulse energy, J	Peak power, W	Maximum weld thickness ^a		Speed of welding	
					in.	mm	in./min	mm/sec
Ruby	Pulsed	3-10	20-50	1-5k	0.005 to .020	.13 to .50	3.0	1.2
Nd:glass	Pulsed	3-10	20-50	1-5k	.005 to .020	.13 to .50	1.5	0.63
Nd:yag	Pulsed	3-10	10-100	1-10k	.005 to .025	.13 to .60	5.0	2.1
CO ₂	Pulsed	5-20	0.1-10	1-5k	.005	.13	3.0	1.2
Nd:yag	CW				1000	.150	3.81	30.0
CO ₂	CW				1000	.025	.60	30.0
Gas dynamic	CW				20 k	.750	19.0	50.0
								12.7
								12.7
								21.2

^aMaximum thickness given here is for Type 304 stainless steel.

TABLE 5.17.— TYPICAL PERFORMANCE DATA FOR CO₂ WELDING/CUTTING LASERS

Demonstration butt welds on tanker construction steels (Nagler, 1976)						
Thickness		Laser power, kW	Weld speed		Comment	
in.	mm		in./min	mm/sec		
0.375	9.5	10.8	50	21.2	Single pass	
0.375	9.5	10.8	45	19.0	Single pass	
0.5	12.7	12.0	27	11.4	Single pass	
0.5	12.7	12.0	30	12.7	Single pass	
0.625	15.9	12.0	24	10.2	Single pass	
0.75	19.1	12.0	45	19.0	Dual pass	
1.0	25.4	12.0	30	12.7	Dual pass	
1.0	25.4	12.0	30	12.7	Dual pass	
1.125	28.6	12.8	27	11.4	Dual pass	
0.375-0.5	9.5-12.7	11.0	90	38.1	Tee joint	
0.375-0.5	9.5-12.7	7.5	65	27.5	Tee joint	
1.0	25.4	12.0	27	11.4	Dual pass ^a	
1.0	25.4	12.0	25	10.6	Dual pass ^a	

^a 0.001-in. (0.03 mm) aluminum foil preplaced at weld interface.							
Material	Thickness		Weld type	Laser power, kW	Weld speed		Number of pieces
	in.	mm			in./min	mm/sec	
HY-130 steel	0.25	6.4	Butt	5.5	50	21.2	3
HY-180 steel	0.062	1.6	Butt	5.5	160	67.7	2
HY-180 steel	0.062	1.6	Lap	5.5	140	59.2	1

Typical cutting and drilling rates for a 1-kW CO ₂ laser (Yankee, 1979)						
Metal thickness, in.	Stainless steel	Cutting rates (in./min)			Titanium	
		Aluminum	Galvanized steel			
0.020	750	800	250		---	
0.032	650	---	---		---	
0.040	550	350	100		250	
0.062	450	200	50		150	
0.080	325	100	---		100	
0.125	200	---	---		---	

Drilling rate: Less than 1 msec is required to drill each of these holes:						
Material	Thickness		Hole diameter			
Tungsten Ceramic Brass	0.020 in. (0.51 mm) 0.101 in. (2.57 mm) 0.010 in. (0.25 mm)		0.020 in. (0.51 mm) 0.050 in. (1.27 mm) 0.250 in. (6.35 mm)			

Laser cutting speeds typically are as much as 30 times faster than friction sawing (Yankee, 1979). Cutting accuracy is about 0.01 mm/cm under closely controlled conditions. All metals – including high-strength, exotic, and refractory alloys such as Inconel and titanium, as well as aluminum, stainless steel, and brass – and nonmetals such as diamond, ceramics, and plastics may be vaporized by laser beams. Hence, parts of these materials may be easily machined. Burr-free laser holes may be drilled as small as 10-100 μm . Lasers can also be used for pattern cutting, gyro balancing, insulation stripping, surface hardening, trimming, photoetching, measurement of range and size to 1 μm accuracy or better, scribing 5-10 μm lines on micro-electronic wafers, flaw detection, marking or engraving parts, and impurity removal (e.g., carbon streaks in diamond). Laser beam machining is “especially adaptable and principally used for relatively small materials processing applications such as cutting, trimming, scribing, piercing, drilling, or other delicate material removal operations similar to milling or shaping” (Yankee, 1979).

Dunning (unpublished Summer Study document, 1980) has suggested a variety of space and lunar applications for laser machining, including flash trimming of cast basalt parts; engraving bar codes on parts to enable quick and accurate recognition by robot vision systems; drilling holes in workpieces an inch thick or less; internal welding of cast basalt joints, pipe, and structural members; impurity removal from lunar-produced semiconductor chips; cutting operations on gossamer structures (Brereton, 1979) in orbit; and case hardening of cast basalt or metal parts. Dunning has also suggested two potential major problems associated with the use of lasers in the context of a self-replicating, growing lunar manufacturing facility: (1) the need for gas jets, and (2) the requirements of closure.

In normal industrial usage, vaporized workpiece material is carried away by a gas jet, usually oxygen (Yankee, 1979). The gas serves three functions: (1) to oxidize the hot working surface, decreasing reflectivity, (2) to form a molten oxide (i.e., the metal “burns”) which releases a large fraction of the useful cutting energy, and (3) to remove slag and hot plasma from the path of the beam. There is no problem maintaining a moderate-pressure O₂ atmosphere around the laser work area, as the beam penetrates air easily. In this case the usual gas jet can still be used. Or, the laser could be placed outside the pressurized working area, shooting its beam through a transparent window. If pressurization must be avoided, laser machining can be done entirely in vacuum and the ionized plasma wastes removed by a magnetic coil following the cut or weld like an ion “vacuum cleaner.” However, it is estimated that up to 80% of the laser cutting energy comes from the exothermic oxidation reaction, so in this latter case laser energies would have to be on the order of five times the value for the equivalent O₂-atmosphere machining.

The problem of closure is even more critical in a replicating autonomous remote factory. The materials closure problem is solved in large measure by resorting to CO₂ gas laser technology. This gas is available in limited quantities on the Moon, whereas materials for solid state lasers such as yttrium, ruby, garnet or neodymium are generally very rare (although Dunning has suggested that spinel, which is plentiful on the Moon, might be substituted for garnet). Quantitative materials closure may be argued as follows. A typical CO₂ laser uses three gases for high-power operation – carbon dioxide to lase, nitrogen to sustain the reaction, and helium for cooling because of its excellent heat conducting properties. Since oxygen is plentiful, the three limiting elements are C, N, and He. From appendix 5E, the LMF in one year can produce 400 kg C, 400 kg N₂, and about 40 kg inert gases (at least 90% of which is He). This is sufficient to make 747 m³ (33,300 moles) of CO₂, 320 m³ (14,300 moles) of N₂, and 224 m³ (10,000 moles) of He, at STP. Even if the laser machining device requires several hundred moles of these gases (a few thousand liters at STP), still only a few percent of available LMF stocks of these elements need be diverted for this purpose, a negligible resource drain.

The problems of parts and assembly closure cannot be answered satisfactorily at the present time. However, it is often asserted that machining the laser end mirrors to high accuracy may be a major roadblock to automated manufacture of lasing devices. Nazemetz (personal communication, 1980) has pointed out that a laser is accurate enough to surface a rough-hewn mirror to the accuracy required for its own construction. In a pinch, concave mirrors could be hewn from solid metal or basalt blanks simply by sweeping the laser beam radially across the disks, applying higher power nearer the center so more material volatizes there, thus creating a perfect spherical or parabolic surface gradient. There appear to be no major unresolvable difficulties associated with the use of lasers in an autonomous lunar manufacturing facility.

After parts leave the laser machining station they may require some slight further treatment such as annealing or coating to prevent cold weld, though this latter function may be unnecessary if laser welding takes place in an oxygen atmosphere (a thin layer of metal oxide prevents the vacuum-welding effect). Once fabrication is completed each part may have one of three possible destinations: (1) assembly sector, where the part is given to a mobile robot for transport to wherever it is needed, (2) parts warehouse (which serves as a buffer supply of extra parts in the event of supply slowdowns or interruptions), where the part is taken to storage by a mobile robot, or (3) fabrication sector, when more fabrication must be performed upon an already manufactured “part” (e.g., solar cell aluminum sheets), where a mobile robot carries the part to wherever

it is needed in the fabrication sector. A general flowchart of the entire automated parts fabrication process appears in figure 5.17.

5F.5 Parts Fabrication: State-of-the-Art

In the operation of any general-purpose fabrication machine (mill, lathe, laser machining system, casting robot, there are seven distinct functions which must be performed either manually or automatically, according to Cook (1975):

- (1) Move the proper workpiece to the machine,
- (2) Load the workpiece onto the machine and affix it rigidly and accurately,
- (3) Select the proper tool and insert it into the machine,
- (4) Establish and set machine operating speeds and other conditions of operation,
- (5) Control machine motion, enabling the tool to execute the desired function,
- (6) Sequence different tools, conditions, and motions until all operations possible on that machine are complete, and
- (7) Unload the part from the machine.

Traditionally all seven operations were performed by the human operator. The development of numerical-control (N/C) machining relieved human operators of the need to manually perform step (5), and automatic tool-changing systems supplanted step (3). Although most modern computer-controlled machining systems have "a finite number of tool-storage locations — 24, 48, or 60 tools, for example — the number that could be built into a system runs into the thousands" (Gettleman, 1979). If the seed is comprised of about 1000 different kinds of parts, each requiring a template pattern for the casting robot, Gettleman's estimate for N/C machine tooling makes plausible the satisfaction of this requirement by extensions of current technology. Adaptive control of N/C machine tools, with sensors that measure workpiece and tool dimensions, tool application forces, vibration and sound, temperatures, and feed rates to optimize production have already been developed (Nitzan and Rosen, 1976) but will require further improvements to achieve the kind of generalized capability required for a lunar SRS.

The next logical developmental step is the design of a completely computer-managed integrated parts manufacturing system. Cook (1975) describes such a system developed and built by Sunstrand Corporation. One version in operation at the Ingersoll-Rand Company is used primarily for fabricating hoists and winches, while another at the Caterpillar Tractor Company is used for making heavy transmission casing parts (Barash, 1976). As of 1975 there were about ten similar systems in operation in the U.S., Japan, Germany, and the U.S.S.R. (Barash, 1975).

The Ingersoll-Rand system consists of six N/C tools — two 5-axis milling machines, two 4-axis milling machines, and two 4-axis drills — arranged around a looped transfer system as shown in figure 5.42. Machining operations include milling, turning, boring, tapping, and drilling, all under the control of an IBM 360/30 central computer. At any given time about 200 tools are in automatic tool-changing carousels, available for selection by the computer, although about 500 are generally available in the system. The computer can simultaneously direct the fabrication of as many as 16 different kinds of parts of totally different design which are either being machined, waiting in queue to be machined, or are in the transfer loop. The entire system is capable of manufacturing about 500 completely different parts. During each 12-hr shift the system is run by three human operators and one supervisor. It is calculated that to achieve the same output using manual labor would require about 30 machines and 30 operators. Finally, the circular pallets used to present parts to each control station have maximum dimensions which fit inside a 1-m cube, exactly the scale discussed earlier in connection with the casting robot.

Another major advance is the variable-mission manufacturing system developed by Cincinnati Milacron Inc. This system not only has the general character of computer-managed parts manufacture seen in other systems but also provides for the processing of low-volume parts at higher rates than those which can be achieved with more conventional N/C machines. For instance, an ingenious five-axis "manufacturing center" automatically changes clusters of tools mounted on a single head so that a number of operations can be performed simultaneously. By means of a novel scheme of handling workpieces from above, the Cincinnati Milacron system provides efficient management of coolants and chips, together with easy access for inspection and servicing (Cook, 1975).

The Japanese have been most aggressive in pursuing the "total automation" concept. During 1973 through 1976 their Ministry of International Trade and Industry (MITI) supported a survey and design study entitled "Methodology for Unmanned Manufacturing" (MUM) which forecast some rather ambitious goals. The MUM factory was to be operated by a 10-man crew, 24 hr/day, and replace a conventional factory of about 750 workers. The factory will be capable of turning out about 2000 different parts at the rate of 30 different parts (in batches of about 1-25) per day, which will be inspected and assembled to produce about 50 different complex machine components such as spindle and turret heads, gear boxes, etc. Machining cells, based on the principle of group technology, will be controlled by a hierarchy of minicomputers and microcomputers, and will receive workpieces via an automated transfer system. Each machine cell will be equipped with inspection and diagnostic systems to monitor such useful

parameters as tool wear, product quality, and the conditions of machine operation. Assembly cells, much like the machining cells, will be equipped with multiple manipulators fashioned after present industrial robots, together with an automated transfer system for movement of assemblies (Nitzan and Rosen, 1976). One ultimate program goal, explicitly stated, was to design a system "capable of self-

diagnosis and self-reproduction ... [and] capable of expansion" (Honda, 1974).

Following this initial study, MITI in 1977 initiated a 7-year national R&D program at a funding level of 12 billion yen (about \$57 million) to develop, establish, and promote technologies necessary for the design and operation of a "flexible manufacturing system complex," a prototype

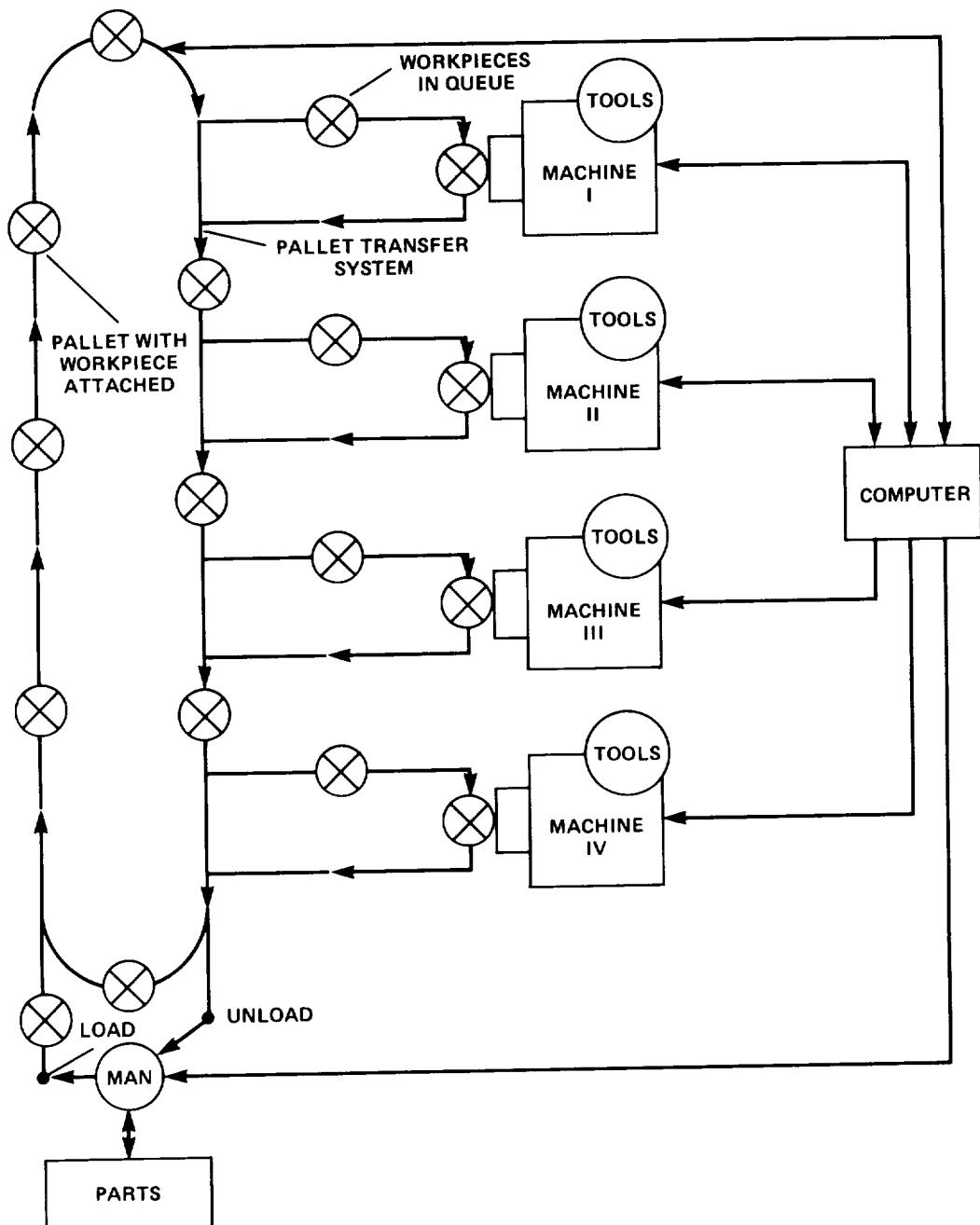


Figure 5.42. – Computer-managed parts manufacturing.

"unmanned" factory to be built sometime in the mid-1980s (Ohmi et al., 1978). The technologies currently receiving emphasis include:

- Optimum design and integrated control of manufacturing systems including blank fabrication, machining and assembly,
- Flexible machining for mechanical parts and components,
- Enlargement of the flexibility of blank fabrication,
- Enlargement of the applicable area of automatic assembly and automatic transfer,
- Application of high-power (20 kW) CO₂ lasers to metalworking,
- Automatic diagnosis of manufacturing facilities to detect malfunctions, and
- Planning and production management to optimize system operation.

MUM presently is being pursued vigorously by three government research institutes and 20 private companies, and is being managed by the Agency of Industrial Science and Technology of MITI (Honda et al., 1979).

The original forecast was that MUM technology would go into operation sometime during the 1980s. At a conference in Tokyo in September of last year, Fujitsu FANUC Ltd., a leading international manufacturer of numerical control (N/C) machining equipment, announced its plans to open a historic robot-making factory near Lake Yamanaka in Yamanashi Prefecture in late November. At the plant, then still under construction, industrial robots controlled by minicomputers would produce other industrial robots without major human intervention save minor machine operation and administrative tasks. The plant is the first "unmanned" factory in the world machinery industry, producing robots and other equipment worth about \$70 million in the first year of operation with only 100 supervisory personnel. In 5 years the plant is expected to expand, perhaps with some of the robots it itself manufactures, to a \$300 million annual output with a workforce of only 200 people, less than a tenth the number required in ordinary machine factories of equivalent output. The mainstay products are to be various kinds of industrial robots and electronic machines. A spokesman said that FANUC's fully automated system is suitable not only for mass production of a single product line but also for limited production of divergent products (IAF Conference, 1980).

An automated plant in which robots make robots is a giant first step toward the goal of a practical self-reproducing machine system. When a factory such as the FANUC plant can make all of the machines and components of which it itself is comprised, its output can be specified to be itself and thus it can self-replicate. It appears likely that the automation technology required for LMF

fabrication and assembly operations could become available within the next 10-20 years, given adequate funding and manpower support targeted specifically to the development of such a system.

5F.6 Automation of Specific LMF Systems

It is useful at this point to consider the automation potential of specific LMF systems. Most critical are the casting robot and the laser machining system, but several other subsystems will also require automation.

Casting Robot Automation

There are two potential precursor technologies to the general-purpose casting robot described in section 5F.3, in addition to established robotics devices such as the Unimate 4000 that produces lost wax ceramic molds for use in investment casting (Moegling, 1980). One of these lines of development has been in the field of precision machining, the other in the area of art and sculpturing.

Engraving and tracer milling are well established machining techniques. These machines use high-speed spindles mounted on pantograph mechanisms guided by master patterns which permit the cutting tools to be guided from an original which may be larger or smaller than the workpiece. The original pattern may be wood, plastic, or metal; the operator follows it with a guide and the machine faithfully reproduces each motion — but enlarges or reduces it as desired (Ansley, 1968).

Modern machines work in three dimensions and can be used for very intricate carving in metal from arbitrary solid originals. A contour milling machine developed by Gorton Machine Corporation uses numerical control to replace entirely the master pattern and the human operator (Ansley, 1968). A skilled technician can preprogram the complete machining cycle for any given part. The Lockheed CAD/CAM system (see below) permits still more sophisticated computerized design and parts fabrication. It seems but a few conceptually simple steps from this level of technology to that required for a "universal" contour-carving device like the casting robot. Such a system will require a vision system, excellent tactile sensing, an automatic tool-changing and pattern-changing capability, and development of an automatic feedstock handling system for metals, gases, and refractories.

Another possible precursor technology to the casting robot may be found in the area of artistic sculpting, otherwise known as "three-dimensional portraiture." An excellent summary of 19th-century attempts to construct machines able to automatically size and shape a human head for personalized sculptures has been written by Bogart (1979). In the last 10 years two very different descendants of the 19th-century efforts to produce sculpted likenesses (thus bypassing the creative artist) have been spawned. The

first of these is modern holography techniques, which permit the generation of 3-D images using laser beams and, more recently, white light sources.

The second technology, often called "solid photography," requires that the human model pose in front of eight cameras shooting simultaneously from different angles. Linear patterns of light are projected onto the subject's face and all three-dimensional information is coded by the cameras. The coded films are then read by an optical scanner which converts the code into digital information which is processed by a computer to produce an accurate surface map of the person or object. This map is then translated into a series of cutting instructions which are passed to two cutting instruments.

In the system operated by Dynell Electronics Corporation of Melville, New York, instructions are first passed to a "coarse replicator" which rough-hews the shape of the human head in paralene wax (high melting point) in 90° sections. After about 30 min, the rudimentary carving is completed and is passed to the "fine-cut replicator" which is also computer-controlled. This time, instead of a single rotating bit, the tooling consists of 20 rotating blades that finish the work to a very high accuracy in about 40 min of work. Human hands are used only for touch-up of very fine details or for imparting skin-like smoothnesses; witnesses to the procedure are impressed with the results — excellent representations of eyebrows, locks of hair, creases, even moles (Field, 1977). Clearly, the Dynell automated sculpting system is not too distant from the casting robot, conceptually or technologically. If treated as a serious item for further development, it is likely that casting robot technology could be ready in a decade or less starting from the current state-of-the-art.

Laser Machining System Automation

Nonlaser spot welding has been a standard automated industrial technique for many years. Welding robots at Chrysler's Hamtramck assembly plant put uniform spot welds on parts assemblies with positional accuracy exceeding 1.3 mm. Typical operation includes a sequence of 24 welds on four automobile assemblies at once (Tanner, 1979). One of the largest and most fully automated welding lines in the world operates at Volvo's Torslanda plant in Gothenburg, Sweden. The new welding line consists of 27 Unimate robots which replace 67 workers with 7. The installation is fully automated, including loading and unloading stations, intermediate assembly of all automobile body parts, lining, and clamping preparatory to welding. The line does a total of 754 spot welds per assembly, and each Unimate is directed by 2-8K programmable controller computers (Mullins, 1977). Kawasaki Unimate robots have been applied to arc welding of motorcycle frames and automobile rear axle housings (Seko and Toda, 1974).

Accuracy in arc welding is more difficult to achieve than in spot welding, but apparently much progress has been made in this area.

Nonlaser machining is also highly automated. The generalized machining center can perform a number of functions in typical operation including milling, drilling, boring, facing, spotting, counterboring, threading, and tapping, all in a single workpiece setup and on many different surfaces of the workpiece (Gettleman, 1979). A numerical-control machine operated by the Giddings and Lewis Machine Tool Company has an automatic tool changer with 40 tools. It machines all sides of a workpiece with one setup. (Setup time is usually 50-90% of total machining time, and a typical part might normally require a dozen setups or more, so this is a substantial savings.) A machined block requiring 174 separate operations can be completed automatically in 43 min; the former method required 4 machines with 3 operators and took 96 min to finish the part. Piggott (personal communication, 1980) estimates that a "typical part" weighing 0.1 kg will require about 20 machining operations. If 10% of all LMF parts must be closely machined after casting, a single Giddings N/C robot could perform all 2,000,000 necessary machining operations in just 0.94 year. Since several such robots could be available in the early LMF, this item is noncritical.

A more sophisticated methodology (Luke, 1972) is used in the Lockheed CAD/CAM system. In this system, the user designs a part of arbitrary shape in three dimensions on an interactive computer-driven TV console. This description is processed to yield a series of machine operations and is then passed to a set of 40 sophisticated N/C machines which make the part "from scratch" out of feedstock supplied at one end. On the average, parts are machined correctly five out of every six tries.

If all LMF parts had already been designed and placed in memory, a shop in space using the Lockheed system could manufacture each of the 1000 different SRS parts. With the addition of pattern recognition software capable of recognizing any part presented to a camera eye, in any physical condition (e.g., rotated, broken, partly melted, partly obscured) (Perkins, 1977), and a simple goal-setting command hierarchy, the Lockheed system might be able to recognize and repair damaged parts presented to it randomly.

The purpose of describing the above nonlaser welding and machining systems is to suggest that laser machining should be equally automatable because the laser may be viewed as another modality for delivering heat or cutting action to a workpiece. Any nonlaser automated welding/machining technology in principle may be modified to accept a laser as its active machining element.

Lasers already have found many automated applications in industry. Computer-driven lasers presently perform automated wire-to-terminal welding on relay plates for electronic switching circuits (Bolin, 1976). There are

automated laser welding lines for manufacturing metal-enclosed gasprotected contacts for telephone switchgear (Schwartz, 1979). A computer-controlled laser welding system at Ford Motor Company allows welding parameters for a number of different automobile underbody designs to be stored in the central memory and retrieved as required for seam welding body-pans (Chang, personal communication, 1978). In the garment industry, the cutting of patterns from single-ply or multilayer stacks of fabrics is easily fully automated and rates of up to 61 m/min have been achieved (Luke, 1972; Yankee, 1979). Flash trimming of carbon resistors has been successfully automated. Automated marking and engraving (with alphanumeric characters) is another application of computer-guided lasers (Yankee, 1979). Numerous other laser applications have already been put into operation (see sec. 5F.4) but are not yet automated. Lasers for many automobile body assembly tasks are impractical today because the component metal pieces to be welded, which are stamped metal sheet, are too inaccurate to permit a close enough fit for laser welding to be feasible — though, according to Schwartz (1979), "this situation may change gradually in the future."

Lunar seed lasers should be able to operate at many different power settings, preferably spanning a broad continuum. Precision machining of liquid- and air-tight valves, laser mirror surfaces, and various other small intricate parts will demand the closest scrutiny of the rate at which energy is delivered to the workpiece. Lasers may also be used for super-accurate ranging and sizing measurements, which require an ultralow power capability as well as sophisticated optics, timing, and data processing systems. Automation of the LMF Laser Machining System will require close computer/mechanical control to perform each of the seven basic machining steps described earlier in section 5F.5.

Some consideration should also be given to the architecture of beam delivery to the workpiece. Laser power may be transmitted directly, in which case the entire laser assembly must be swiveled as various operations are performed. One alternative is to use a system of lightweight movable mirrors to angle laser energy in the desired direction to impact the workpiece. Reflectivities up to 0.86 for aluminum on glass would give an absorbed power density of 14 to 140 W/cm² for a 1-10% efficient 10 kW laser beam with a 1 cm² cross section. This heating may be reduced by at least an order of magnitude by "jiggle" the mirrors along their plane to spread the beam impact spot over a wider area while maintaining precise directional control. Another possible solution is to locate a high power laser in some central location and convey the beam to its destination via large fiber-optic light pipes. There are possible materials closure problems with fiber-optics, and absorbed energy may damage or destroy the glass, but this alternative offers many interesting opportunities and cannot be logically ruled out.

The team recognizes that lasers may not be the optimum technology for an autonomous replicating lunar facility. Their inclusion in the present design is intended as a heuristic device to illustrate, not unequivocally select, a particular option. For example, industrial experts in manufacturing technologies are split over whether lasers or electron beams are generally superior or more versatile, e.g., Schwartz (1979) favors lasers and Yankee (1979) favors e-beams. The MIT study group selected electron-beam cutting over lasers because "lasers are less efficient and require more maintenance and repair than EB guns" (Miller and Smith, 1979), a conclusion not adequately documented in their final report.

Nor is it absolutely clear that conventional machine tools such as mills, lathes, or drills are unsuitable for use in space. The problem most often cited in this context is that the tool bit and workpiece may vacuum weld during machining. However, cold welding is known to occur only between identical metals or between those with very similar crystallographic characteristics (such as aluminum and magnesium). Steel, for instance, will not vacuum weld to aluminum. Neither will any metal part cold weld to cast basalt.

Further, ceramic cutting tools have recently been developed which have increased the cutting speeds of mills and lathes dramatically. When tungsten carbides were introduced in 1929, cutting speeds quadrupled to 100 to 200 m/min. Since the 1950s, ceramic and other cemented oxide (alumina) and refractory tool materials such as nitrides and borides have been successfully employed in achieving cutting rates of 300 m/min and higher (Ansley, 1968). Ceramic tools will not cold weld to anything.

A more critical problem would seem to be the seizing of internal machine components, rather than vacuum welding between tool and workpiece. This difficulty could perhaps be surmounted by bathing enclosed machinery in lubricants, a light oxygen atmosphere trapped by airtight seals, or by using basalts or ceramics to construct or merely protectively coat internal machine moving parts.

Automation of Other Systems

The remaining subsystems within the parts fabrication sector must also be automated for full LMF autonomous operation. These subsystems include:

(1) Kilns and metallurgical furnaces: The extraterrestrial fiberglass production system using solar energy, designed by Ho and Sobon (1979), is designed to be automated. This system includes melting and drawing operations. According to the authors, "the systems will be automated, but minimum manpower will be required for maintenance. For the lunar plant, maintenance will be required at the beginning of each lunar day to begin the drawing process."

(2) Basalt threads: The system of Ho and Sobon will be automated. Also, a series of eleven specific steps which a

manufacturing robot such as a Unimate must perform in order to completely automate the thread-drawing procedure is given in appendix 4D.

(3) Wire wrapping: An automatic insulation wire-wrapping machine has been described in some detail by Miller and Smith (1979).

(4) Sheet metal and cutting operations: Miller and Smith (1979) discuss in some detail aluminum ribbon and sheet operations. Vacuum vapor deposition as a fabrication technique is also described in Johnson and Holbrow (1977). These will be at least partially automated.

(5) Refractory and cement production: Ansley (1968) has described a concrete batching plant equipped with electronic controls permitting the selection of some 1500 different formulas and which give twice the output of manually operated plants. Batches are prepared by inserting a punched card into a reader to specify the formula to be used, and the system does the rest automatically if adequate materials have been supplied.

(6) Ball mills and magnetic purification: These are standard automated technologies, assumed available in space processing models provided by O'Neill (1976), Phinney et al. (1977), and others.

5F.7 Sector Mass and Power Estimates

In lieu of a complicated breakdown of fabricator sector component subsystems with detailed analysis of each, table 5.18 illustrates a more practical approach. This information was assembled from various sources and gives typical masses and power requirements for parts fabrication facilities in previous studies.

The nominal annual output of the original lunar seed is 100 tons/year. Using the most extreme machine productivity values given in table 5.18, fabrication sector mass may range from 137 kg up to 20,400 kg. A similar comparison with the power requirements values gives a range of 0.3-345 kW for sector energy consumption. The upper ranges of these estimates are probably most appropriate in the replicating lunar factory application.

5F.8 Information and Control Estimates

Even in the absence of a detailed analysis of the necessary control operations, it is obvious that the complete description of all parts will dominate computer memory requirements. Since each typical part has a characteristic

TABLE 5.18.-- COMPARISON OF FABRICATION PLANT
MASSES AND POWER REQUIREMENTS FROM PRE-
VIOUS RELATED STUDIES

Source	Plant mass, kg/kg sec output	Power requirement, W/kg plant
Johnson and Holbrow (1977) annealing, trimming, pressing plate silica glass plant	8.3×10^5	2.2
Ho and Sobon (1979) fiberglass threads/ rods plant	3.8×10^5	16.9
Johnson and Holbrow (1977) bulk process- ing and heavy industry estimate for human workers	4.3×10^4	2
O'Neill, Driggers, and O'Leary (1980) estimated range for machine shop bulk fabrication systems	3.6×10^5 - 3.6×10^6	---
Miller and Smith (1979) MIT Study on Space Manufacturing Facility	6.4×10^6	12
Vajk et al. (1979)	3.6×10^5	---

surface area of 10^{-3} m², then if the surface of each is mapped to 1 mm² resolution per pixel, each part will require 1000 pixels for complete coverage. Each pixel must identify three position coordinates, materials used, machining operations to be performed, etc. If 100 bits/pixel is adequate, then roughly 10^5 bits/part are required in memory for a total of 10^{11} bits of storage for all 1,000,000 parts in the original lunar seed. This crude estimate is intended as a combined total for description and operation of the system.

5F.9 References

- Acharekar, M. A.: Lasers Finding Growing Use in Industrial Welding. *Welding Engineering*, December 1974, pp. 9-11.
- Ansley, Arthur C.: *Manufacturing Methods and Processes*. Chilton Book Company, Philadelphia, 1968. Revised and enlarged edition.
- Barash, M. M.: Optional Planning of Computerized Manufacturing Systems. Proc. 3rd NSF/RANN Grantee's Conference on Production Research and Industrial Automation, Cleveland, October 1975.
- Barash, M. M.: Integrated Machining Systems with Work-piece Handling. In Proc. 6th Intl. Symp. Industrial Robots and Industrial Robot Technology, University of Nottingham, March 1976, International Fluidics Services, Kempston, 1976, pp. H4/29-36.
- Bogart, Michele: In Art the Ends Don't Always Justify Means. *Smithsonian*, vol. 10, June 1979, pp. 104-108, 110-111.
- Bolin, S. R.: Bright Spot for Pulsed Lasers. *Welding Design and Fabrication*, August 1976, pp. 74-77.
- Brereton, Roy G., ed.: Discussion Meeting on Gossamer Spacecraft (Ultralightweight Spacecraft): Final Report, JPL Publication 80-26, Jet Propulsion Laboratory, Pasadena, California, 15 May 1980. 174 pp. NASA CR-163275.
- Campbell, Ivor E.; and Sherwood, Edwin M., eds.: *High-Temperature Materials and Technology*. Wiley, New York, 1967.
- Cook, Nathan H.: Computer-Managed Parts Manufacture. *Scientific American*, vol. 232, February 1975, pp. 22-29.
- Field, Roger: Computerized Cameras, Knives Sculpt Quickly. *Science Digest*, vol. 81, April 1977, pp. 77-78.
- Freitas, Robert A., Jr.: Baseline Reference Design for a Growing, Self-Replicable Lunar Manufacturing Facility. Summer Study Document, 8 July 1980. 56 pp.
- Gentleman, K. M.: *Fundamentals of NC/CAM*. Modem Machine Shop 1979 NC/CAM Guidebook, 1979.
- Gitzen, W. H.: Alumina Ceramics. U.S. Air Force Rept., AFML-TR-66-1621, 1966.
- Grant, T. J.: The Need for On-Board Repair. In Project Daedalus – The Final Report on the BIS Starship Study, Anthony R. Martin, ed., British Interplanetary Society, 1978 (Journal of the British Interplanetary Society, Supplement, 1978), pp. S172-S179.
- Ho, Darwin; and Sobon, Leon E.: Extraterrestrial Fiberglass Production Using Solar Energy. In *Space Resources and Space Settlements*, John Billingham, William Gilbreath, Brian O'Leary, and B. Gossett, eds. NASA SP-428, 1979, pp. 225-232.
- Honda, Fujio, ed.: Methodology for Unmanned Metal Working Factory. Project Committee of Unmanned Manufacturing System Design, Bulletin of Mechanical Engineering Laboratory No. 13, Tokyo, 1974.
- Honda, F.; Kimura, M.; Ozaki, S.; Tamagori, K.; Ohmi, T.; and Yoshikawa, H.: Flexible Manufacturing System Complex Provided with Laser – A National R&D Program of Japan. Proc. IFAC Symposium on Information-Control Problems in Manufacturing Technology, September 1979, pp. 7-11.
- IAF Conference: Robots To Produce Robots At Fujitsu. From an English-language newspaper handout at the Conference, Tokyo, 21-28 September 1980. See also *The Washington Post*, 13 October 1980, p. A-38.
- Johnson, P. D.: Behavior of Refractory Oxides and Metals, Alone and in Combination, *in Vacuo* at High Temperatures. *J. Am. Ceram. Soc.*, vol. 33, no. 5, 1950, pp. 168-171.
- Johnson, Richard D.; and Holbrow, Charles, eds.: *Space Settlements: A Design Study*, NASA SP-413, 1977.
- Kaiser Aluminum Company: New Kaiser Super Refractories, May 1962.
- Kopecky, Lubomir; and Voland, Jan: The Cast Basalt Industry. In *Geological Problems in Lunar Research*, Harold Whipple, ed., Annals of the New York Academy of Science, vol. 123, July 16, 1965, pp. 1086-1105.
- Luke, Hugh D.: *Automation for Productivity*. John Wiley and Sons, New York, 1972.
- Miller, Rene H.; and Smith, David B. S.: Extraterrestrial Processing and Manufacturing of Large Space Systems, vols. 1-3, NASA CR-161293, September 1979.
- Moegling, Frank: Robot-Controlled Mold-Making System. *Robotics Today*, Fall 1980, pp. 30-31.

Subsystem control hardware is likely to use vastly less computer capacity than this. The entire Sundstrand integrated parts manufacturing line is managed by an IBM 360/30 central computer with microcomputers driving each robot station. While some tricks might be employed to reduced redundancy (such as "chunking" large similar areas), more convoluted surfaces will require extra description. It is likely that the main driver will be the requirements for parts description.

- Mullins, Peter J.: Robot Welding of Car Bodies. *Automotive Industries*, 1 February 1977. Reprinted in *Industrial Robots: Volume 2 – Applications*, William R. Tanner, ed., Society of Manufacturing Engineers, Dearborn, Michigan, 1979, pp. 151-152.
- Nagler, H.: Feasibility, Applicability, and Cost Effectiveness of LBW of Navy Ships, Structural Components and Assemblies. *Contr. N00600-76-C-1370*, vols. 1-2, 22 December 1976.
- Nitzan, David; and Rosen, Charles A.: Programmable Industrial Automation. *IEEE Trans. Comp.*, vol. C-25, December 1976, pp. 1259-1270.
- Ohmi, T.; Kanai, M.; and Honda, F.: Research and Development of a Flexible Manufacturing System Complex Provided with Laser – A Japanese National Project. Paper presented at the Workshop on Flexible Manufacturing Systems, 11-14 September 1978, Peoria, Illinois.
- O'Neill, Gerard K.: Engineering a Space Manufacturing Center. *Astronautics and Aeronautics*, vol. 14, October 1976, pp. 20-28, 36.
- O'Neill, Gerard K.; Driggers, Gerald; and O'Leary, Brian: New Routes to Manufacturing in Space. *Astronautics and Aeronautics*, vol. 18, October 1980, pp. 46-51.
- Perkins, W. A.: Model-Based Vision System for Scenes Containing Multiple Parts. General Motors Research Laboratories Publication, GMR-2386, June 1977.
- Phinney, W. C.; Criswell, D. R.; Drexler, E.; and Garmirian, J.: Lunar Resources and Their Utilization. In *Space-Based Manufacturing from Extraterrestrial Materials*, Gerard K. O'Neill, ed., AIAA Progress in Astronautics and Aeronautics Series, vol. 57, AIAA, New York, 1977, pp. 97-123.
- Robson, T. D.: High-Alumina Cements and Concretes. John Wiley and Sons, Inc., New York, 1962.
- Schwartz, M. M.: Metals Joining Manual. McGraw-Hill Book Co., New York, 1979.
- Seko, K.; and Toda, H.: Development and Application Report in the Arc Welding and Assembly Operation by the High-Performance Robot. In Proc. 4th Intl. Symp. Industrial Robots, Tokyo, Japan, November 1974, Japan Industrial Robot Assn., Tokyo, 1974, pp. 487-596.
- Spotts, M. F.: Design Engineering Projects. Prentice-Hall, Inc., New Jersey, 1968.
- Subramanian, R. V.; Austin, H. F.; Raff, R. A. V.; Sheldon, G.; Dailey, R. T.; and Wullenwaber, D.: Use of Basalt Rock for the Production of Mineral Fiber. Pacific Northwest Commission Annual Report, Contract #NR-3001, College of Engineering, Washington State University, Pullman, Washington, June 1975. 79 pp.
- Subramanian, R. V.; and Kuang-Huah, Shu: Interfacial Bonding in Basalt Fiber-Polymer Composites. 34th Annual Technical Conference, Reinforced Plastics/Composites Institute, Section 17C, 1979, pp. 1-10.
- Tanner, W. R., ed.: *Industrial Robots: Volume 2 – Applications*, Society of Manufacturing Engineers, Dearborn, Michigan, 1979. 287 pp.
- Vajk, J. P.; Engel, J. H.; and Shettler, J. A.: Habitat and Logistic Support Requirements for the Initiation of a Space Manufacturing Enterprise. In *Space Resources and Space Settlements*, John Billingham, William Gilbreath, Brian O'Leary, and B. Gossett, eds., NASA SP-428, 1979, pp. 61-83.
- Yankee, Herbert W.: Manufacturing Processes. Prentice-Hall, Englewood Cliffs, New Jersey, 1979. 765 pp.
- Zachary, Wm. B.: A Feasibility Study on the Fabrication of Integrated Circuits and Other Electronic Components in Space Utilizing Lunar Materials. Paper presented at the 5th Princeton/AIAA/SSI Conference on Space Manufacturing, 18-21 May 1981, Princeton, NJ.

APPENDIX 5G

LMF ASSEMBLY SECTOR

5G.1 Assembly Sector Components and Technology Assessment

After raw lunar soil has been processed by the chemical processing sector into metallic and nonmetallic elements, and the parts fabrication sector has used these substances to manufacture all parts needed for LMF construction activities (growth, replication, or production), it is the job of the assembly sector to accept individual completed parts and fit them together to make working machines and automated subsystems themselves capable of adding to the rate of construction activities. A number of basic functions are required to perform sophisticated assembly operations. These are outlined in the assembly sector operations flowchart in figure 5.18. Each functional subsystem is discussed briefly below.

Parts Input

Parts produced by the fabrication sector are delivered either to inventory or directly to the assembly sector via mobile Automated Transport Vehicle (ATV) which runs on wheels or guide tracks. Parts are also retrieved from inventory by the ATVs. All retrieved or delivered parts are placed in segregated bins as input to the automated assembly system.

Parts Recognition/Transport/Presentation (RTP) System

The Recognition/Transport/Presentation (RTP) system is responsible for selecting the correct parts from the input bins, transporting them to within the reach of assembly robots, and presenting them in a fashion most convenient for use by the assembly robots. This will require a manipulator arm, vision sensing, probably tactile sensing, and advanced "bin-picking" software.

Early research concentrated on the identification and handling of simple blocks. For instance, at Hitachi Central Research Laboratory prismatic blocks moving on a conveyor belt were viewed, one at a time, with a television camera and their position and orientation determined by special software. Each block was then tracked, picked up with a suction-cup end-effector, and stacked in orderly fashion under the control of a minicomputer (Yoda et al.,

1970). In another early experiment performed at Stanford University, a TV camera with color filters and a manipulator arm was developed that could look at the four multi-colored blocks of an "instant Insanity" puzzle, compute the correct solution to the puzzle, and then physically stack the blocks to demonstrate the solution (Feldman et al., 1974).

At the University of Nottingham, the identity, position, and orientation of flat workpieces were determined one at a time as they passed under a down-looking TV camera mounted in a vertical turret much like microscope lens objectives. A manipulator then rotated into a position coaxial with the workpiece and acquired it (Heginbotham et al., 1972). More recently, software developed by General Motors Laboratories can identify overlapping parts laid out on a flat surface. The computer analyzes each part, calculates geometric properties, then creates line drawing models of each object in the scene and memorizes them. Subsequently, objects coming down the conveyor belt which resemble any of the memorized parts in shape — even if only small sections of a part can be seen or the lighting is poor — will be identified correctly by the system (Perkins, 1977).

In a recent series of experiments performed at SRI International, workpieces transported by an overhead conveyor were visually tracked. The SRI Vision Module TV camera views a free-swinging hanging casting through a mirror fixed on a table at 45°. An LSI-11 microprocessor servos the table in the x-y plane to track the swinging part. If a part is swinging over a 20 cm arc at about 0.5 Hz, the tracking accuracy is better than 1 cm continuously (Nitzan, 1979; Nitzan et al., 1979; Rosen, 1979). A moderate research and development program could produce an arm capable of tracking and grabbing a swinging part.

At Osaka University a machine vision system consisting of a television camera coupled to a minicomputer can recognize a variety of industrial parts (such as gasoline engine components) by comparing visual input from unknown parts with stored descriptions of known parts. The system can be quickly trained to recognize arbitrary new objects, with the software generating new internal parts models automatically using cues provided by the operator. The present system can recognize 20-30 complex engine parts as fast as 30 sec/part, and new objects can be

learned in 7 min (Yachida and Tsuji, 1975). Another system developed at SRI International can determine the identity, position, and orientation of workpieces placed randomly on a table or moving conveyor belt by electro-optical vision sensing, then direct a Unimate industrial robot arm to pick up the workpiece and deliver it to the desired destination (Agin and Duda, 1975).

Contact sensing may also be used in parts recognition. Takeda (1974) built a touch sensing device consisting of two parallel fingers each with an 8 X 10 needle array free to move in and out normal to the fingers and a potentiometer to measure the distance between the fingers. As the fingers close, the needles contact an object's surface contour in a sequence that describes the shape of the object. Software was developed to recognize simple objects such as a cone.

Of direct relevance to the lunar self-replicating factory RTP system is the "bin-picking" research conducted at SRI International. This involves the recognition and removal of parts from bins where they are stored by a robot manipulator under computer control. Three classes of "bins" may be distinguished: (1) workpieces highly organized spatially and separated, (2) workpieces partially organized spatially and unseparated, and (3) workpieces in completely random spatial organization. Simple machine vision techniques appear adequate for bin picking of the first kind, essentially state-of-the-art. Semiorganized parts bins (second class) can be handled by state-of-the-art techniques, except that picking must be separated into two stages. First, a few parts are removed from the bin and placed separately on a vision table. Second, standard identification and manipulation techniques are employed to pick up and deliver each part to the proper destination. Parts bins of the third class, jumbled or random pieces, require "a high level of picture processing and interpretive capability" (Rosen, 1979). The vision system has to cope with poor contrast, partial views of parts, an infinite number of stable states, variable incident and reflected lighting, shadows, geometric transformations of the image due to variable distance from camera lens to part, etc., a formidable problem in scene analysis. Some innovations have been made at General Motors in this area (Perkins, 1977), but researchers believe that progress using this technique alone will be slow, and that practical implementation will require considerably faster and less expensive computational facilities than are presently available (Rosen, 1979).

At SRI an end-effector with four electromagnets and a contact sensor has been built to pick up four separate castings from the top of a jumbled pile of castings in a bin. A Unimate transports the four castings to a backlit table and separates them. Then a vision subsystem determines stable states, position, and orientation, permitting the Unimate gripper to pick up each casting individually and transfer it to its proper destination (Nitzan et al., 1979).

Although clearly more work needs to be done, a great deal of progress already has been made. It is possible to

imagine a 5-10 year R&D effort which could produce the kind of RTP system required for the LMF assembly sector. Considerably more effort will be required to achieve the level of sophistication implied by Marvin Minsky's reaction to a discussion of current bin-picking and conveyor belt picking technology: "On this question of the variety of parts on assembly lines, it seems to me that assembly lines are silly and when we have good hand-eye robots, they will usually throw the part across the factory to the machine who wants it and that machine will catch it" (Rosen, 1979). The RTP system for the self-replicating LMF does not require this extreme level of robot agility.

Parts Assembly Robots

Once the correct parts have been identified, acquired, and properly presented, assembly robots must put them together. These assemblies — electric motors, gearboxes, etc. — are not yet working machines but rather only major working components of such machines. Thus it may be said that assembly robots assemble simple parts into much more complex "parts."

There has been a certain amount of basic research on aspects of programmable assembly. At MIT in 1972 a program called COPY could look at a simple structure built of children's building blocks, then use a manipulator to physically build a mirror image of the structure to prove its "understanding" of the block shapes and orientations. It would do this by withdrawing the blocks it needed from a collection of objects in its field of view, randomly spread out on a table (Winston, 1972). In Japan, a Hitachi robot called HIVIP could perform a similar task by looking at a simple engineering drawing of the structure rather than at the physical structure itself (Ejiri et al., 1971). In Edinburgh the FREDDY robot system could be presented with a heap of parts comprising a simple but disassembled model. Using its TV cameras and a manipulator, the system sorted the pieces, identified them correctly, then assembled the model. Assembly was by force and touch feedback, using a vise to hold partial assemblies, and parts recognition was accomplished by training (Ambler et al., 1975).

Research has also begun on the problems involved in fitting parts together or "parts mating." For instance, Inoue (1971) programmed a manipulator to insert a peg into a hole using force sensing at the manipulator joints. A more sophisticated version was later built by Goto at Hitachi Central Research laboratory. This version consisted of a compliant wrist with strain gauge sensors to control the insertion of a 1.2-cm polished cylinder into a vertical hole with a 7 to 20 μm clearance in less than 3 sec (Goto et al., 1974).

Besides fitting, assembly operations also include fastening. The most common methods include spot welding, riveting, arc welding, bolting, nailing, stapling, and gluing, all of which have been automated to some degree.

Numerical-control (N/C) riveting machines have replaced human riveters in the production of jetliner wings at Boeing Aerospace (Heppenheimer, 1977). At Westinghouse Electric Corporation a four-joint programmable manipulator under minicomputer control performs arc welding along curved trajectories (Abraham and Shum, 1975). According to information gleaned from Ansley (1968) and Clarke (1968), the Gemini spacecraft required 0.15 m/kg of seam welds and 6.9 spot welds/kg. Thus, for a 100-ton LMF seed equal to the Gemini capsule in its welding requirements, 15,000 m of seam welding would be required. This should take about a month of continuous work for a dedicated 5-10 kW laser welder (see appendix 5F). Another alternative is to make positive use of vacuum welding. Surfaces of parts to be fastened would be cleaned, then pressed gently together, causing a cold weld if they are made of the same or similar metallic material. Cast basalt end-effectors will probably be required for handling in this case.

At a high level of sophistication, assembly of certain well-defined machines from basic parts has been studied. Abraham and Beres (1976) at Westinghouse have described a product line analysis in which assembly line automation sequences were considered for constructing ten candidate assemblies, including a continuous operation relay (300 assembly steps), low voltage bushings (5 parts), W-2 low voltage switches (35 parts), fuse assembly (16 steps), and a small motor rotor assembly (16 steps). The tasks and implementation list for a sample motor rotor assembly is shown in table 5.19. This research has evolved into the Westinghouse APAS System, which uses state-of-the-art industrial robots and can automatically assemble complete electric motors of eight different classes representing 450 different motor styles discovered in a broad survey of all motors (van Cleave, 1977).

Other major industry and laboratory accomplishments include the following:

- Typewriter assemblies — At IBM Research Laboratories a program has been under way to use a multi-degree-of-freedom manipulator with a computer-controlled system for assembling small but complex parts. A high-level programming language for mechanical assembly was developed and used to acquire and assemble irregular typewriter parts (Will and Grossman, 1975).
- Water pump assembly — At Stanford University a manipulator called the "Stanford Arm" was programmed to assemble a water pump consisting of a total of 9 parts (base, gasket, top, and six screws). Joint forces were determined indirectly from measurements of drive motor currents. The software compensated for gravity and inertial forces, and included force feedback to locate holes for inserting two pins used to align the gasket (Bolles and Paul, 1973).

TABLE 5.19.— ASSEMBLY TASKS FOR A ONE-ROBOT CONFIGURATION, TO ASSEMBLE SMALL MOTOR ROTORS

Sequential tasks	Task implementation methods
1. Heat core in oven	New vertical in-line oven
2. Place shaft in hot core	Dedicated assembly unit
3. Quench cool	Water spray
4. Transfer subassembly to in-line conveyor	Pick and place device #1
5. Stake shaft	Automatic stake machine
6. Test subassembly	Automatic test device
7. (Optional — remove reject subassembly)	Computer-controlled robot #1
8. Retrieve switch from vision table	
9. Place switch on shaft	
10. Retrieve top sleeve	
11. Place top sleeve on shaft	
12. Press top sleeve and switch	Dedicated assembly units
13. Assemble bottom sleeve and press	
14. Assemble rubber washers	
15. Transfer subassembly to conveyor	Pick and place device #2
16. Assemble nylon washers	Dedicated assembly unit

- Compressor cover assembly — An assembly station using computer vision, various other sensors, and a robot arm with a force-controlled gripper and an *x-y* table has been developed to place and fasten the cover on an air compressor assembly (see fig. 5.43). There are 10 parts in the assembly operation, although one "part" is a preassembled compressor housing (McGhie and Hill, 1978).
- Motor and gearbox assemblies — Kawasaki Laboratories has demonstrated that complex motor and gearbox assemblies can be put together with precision

feedback sensors and appropriate manipulator grippers and fixtures. Kawasaki uses vibratory motion to jiggle parts with suitable bevels and tapers into place during assembly which automatically compensates for minor misalignments or tolerance variations (Thompson, 1978).

Automobile alternator assembly — A programmable robot assembly station built at the Charles Stark Draper Laboratory can assemble a commercial automobile alternator which consists of 17 individual parts, in a total of 162 sec using 6 tools (Nevins and Whitney, 1978). Simple changes such as using multiple head screwdrivers and assembling several units at once should bring the assembly time down to 60 sec/unit (Thompson, 1978). Figure 5.44 shows the functional components and flow pattern of the Draper machine. The Japanese have made similar advances. In fact, one such robot has been successfully assembling automotive alternators on a production basis in a standard factory environment for more than 3 years (Thompson, 1978).

- Gasoline engine assembly -- Kawasaki's most impressive undertaking is the development of a pilot line for the automated assembly of small gasoline engines (Seko and Toda, 1974). Under control of one minicomputer, the assembly proceeds sequentially through five work stations, each including two small Kawasaki Unimates, a table, special jigs and tools, parts feeders, and special end-effectors. Controlled by the minicomputer but working independently, each robot performs a sequence of previously taught assembly operations including parts acquisition, parts mating, and, if necessary, parts fastening operations. No sensors were used for manipulative control and, consequently, there is heavy reliance on expensive jigging for orientation of workpieces. By the mid-1970s, the system was slow and not cost effective, but significant improvements were already being planned (Nitzan and Rosen, 1976).
- Expert system assembler — Some work has been done by Hart (1975) in developing a computer-based consultant able to "talk someone through" the assembly

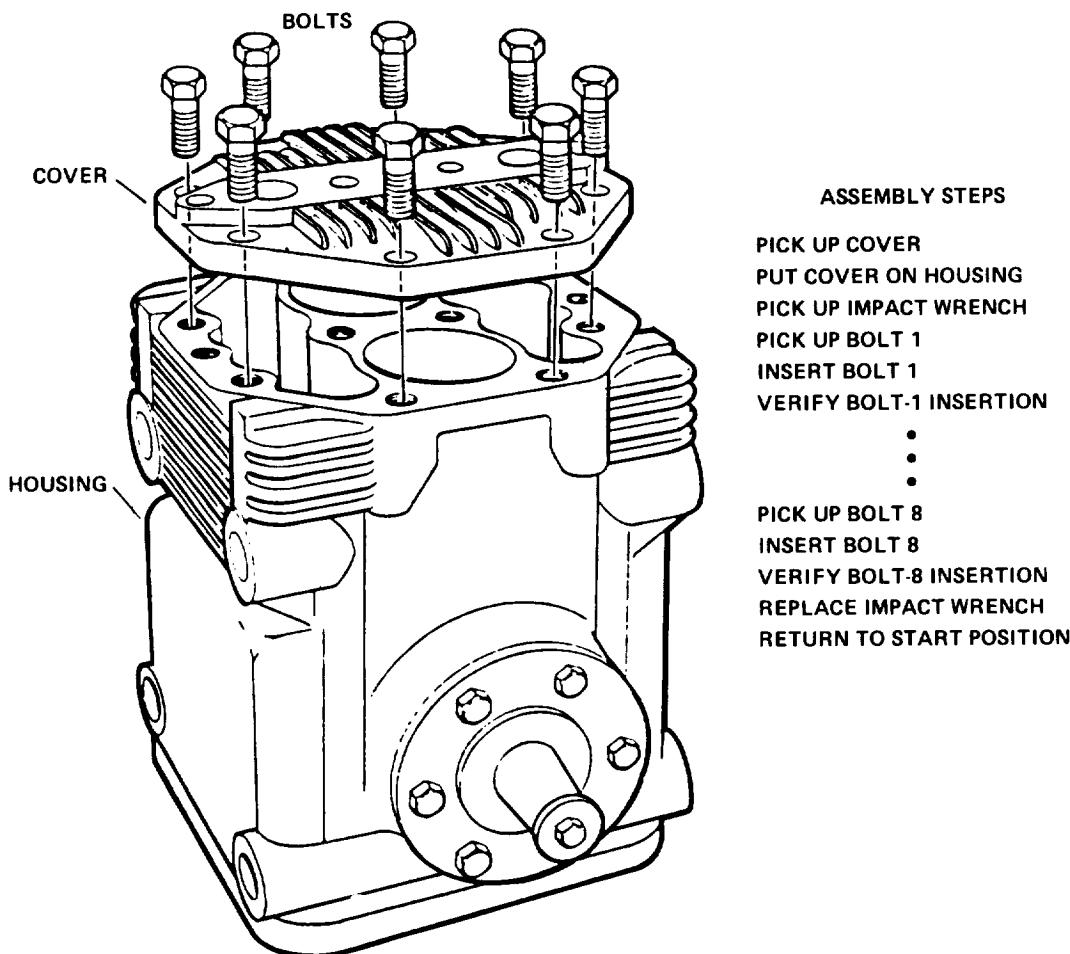


Figure 5.43.— Exploded view of SRI compressor cover assembly. (Rosen et al., 1978.)

of a complicated air-compressor assembly. In principle, the same kind of system could be used to "talk a robot," such as a repair robot with many different functions or a rescue robot, through the same assembly steps.

Clearly, a great deal of progress has been made, but much more remains to be made in all areas before an LMF-capable universal assembly system could be designed. Nitzan, private communication, 1980) estimates such a system might become available commercially by the end of the present century at the present rate of development. The amazing progress of the Japanese in developing "unmanned manufacturing" systems confirms this estimate, and suggests that by the end of the present decade a serious effort to design a universal assembly system of the type required for the lunar SRS might be successful.

If the original LMF seed has about 10^6 parts which must be assembled within a replication time $T = 1$ year, then parts must be assembled at an average rate of 31 sec/

part. If subassembly assembly is included with successive ranks of ten (i.e., 10 parts make a subassembly, then 10 subassemblies make a more complex subassembly, etc.), then 1.111111×10^6 assembly operations are required which is only 28 sec/part. This is about typical for assembly operations requiring 100% verification at each step, using state-of-the-art techniques. The Draper robot described earlier assembles 17 parts in 162 sec, or 9.5 sec/part, and the improvement to 60 sec for the whole alternator assembly task would decrease this to 3.5 sec/part, an order of magnitude less than the mean continuous rate required for successful LMF operation.

Assembly Inspection Robots

After parts have been assembled by assembly robots with 100% verification at each step, the final assembly must be inspected as a final check to ensure it has been correctly built from the correct parts. According to Rosen (1979),

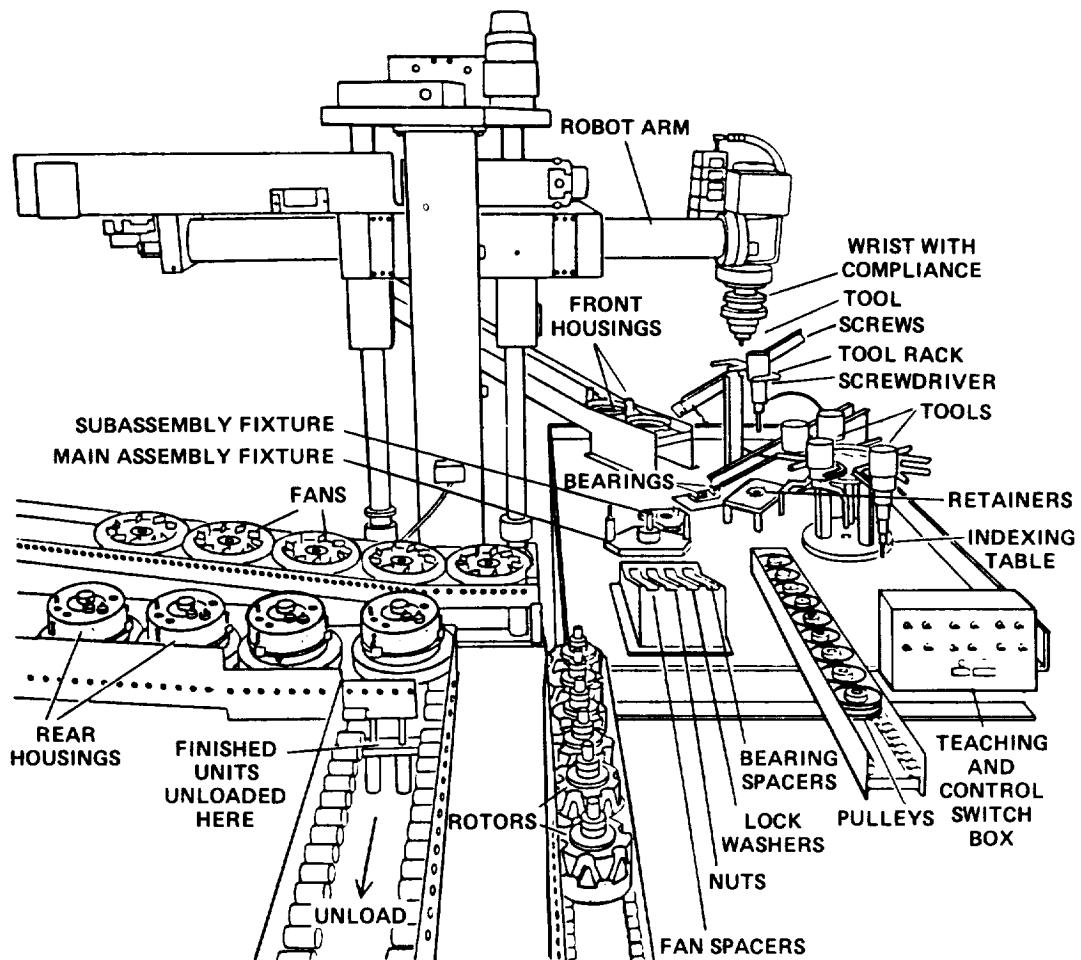


Figure 5.44.—Functional components of the Draper automobile alternator assembly robot. (Nevins and Whitney, 1978.)

machine vision for inspection may be divided into two broad classes: (1) inspection requiring highly quantitative measurement, and (2) inspection that is primarily qualitative but frequently includes semiquantitative measures.

In the quantitative inspection class, machine vision may be used to inspect stationary and moving objects for proper size, angles, perforations, etc. Also, tool wear measurements may be made. The qualitative inspection class includes label reading, sorting based on shape, integrity, and completeness of the workpiece (burrs, broken parts, screws loose or missing, pits, cracks, warping, printed circuit miswiring), cosmetic, and surface finishes. Each type of defect demands the development of specialized software which makes use of a library of subroutines, each affecting the extraction and measurement of a key feature. In due course, this library will be large and be able to accommodate many common defects found in practice. Simple vision routines utilizing two-dimensional binary information can handle a large class of defects. However, three-dimensional information, including color and gray-scale, will ultimately be important for more difficult cases (Rosen, 1979).

With the SRI-developed vision module, a number of inspection tasks have been directed by computer. For example, washing machine water pumps were inspected to verify that the handle of each pump was present and to determine in which of two possible positions it was. A group of electrical lamp bases was inspected to verify that each base had two contact grommets and that these were properly located on the base. Round and rectangular electrical conduit boxes were inspected as they passed on a moving conveyor, the camera looking for defects such as missing knockouts, missing tabs, and box deformation (Nitzan, 1979).

An inspection system developed by Auto-Place, Inc. is called Opto-Sense. In one version, a robot brings the workpiece into the field of vision. Coherent laser light is programmed by reflection off small adjustable mirrors to pass through a series of holes and slots in the part. If all "good part" conditions are met, the laser light is received by the detector and the part is passed. In addition to looking at the presence or absence of holes and object shape, the laser system can also check for hole size and location, burrs or flash on parts, and many other conditions (Kirsch, 1976). Range-imaging by lasers is well suited for the task of inspecting the completeness of subassemblies (Nitzan et al., 1977).

An inspection system designed for an autonomous lunar factory would need an internal laser source, a three-dimensional scanning pattern, at least two detectors for simple triangulation/ranging, a vision system for assembly recognition and position/orientation determination, and a large library of parts and assemblies specifications so that the inspection system can determine how far the object

under scrutiny deviates from nominal and a valid accept/reject/repair decision may be made.

Electronics Assembly Robots

Electronics components, including resistors, capacitors, inductors, discrete semiconductor components (diodes, thyristors), and microelectronic "chips" (microprocessors, RAMs, ROMs, CCDs) are produced by the Electronics Fabrication System in the fabrication sector. Aluminum wire, spun basalt insulation, and aluminum base plates are provided from the bulk or parts fabrication system described in appendix 5F. After these parts are properly presented to the electronics assembly robots, these robots must assemble the components into major working electronics systems such as power supplies, camera systems, mini/microcomputer CPUs, computer I/O units, bulk memory devices, solar cell panels, etc. Electronics assembly appears to require a technology considerably beyond the state-of-the-art.

Present techniques for automated electronics assembly extend mainly to automatic circuit board handling. For instance, Zagar Inc. uses an automatic PCB drilling machine, and Digital Systems Inc. has an N/C automatic drilling machine with four speeds for drilling four stacks of boards simultaneously (Ansley, 1968). A circuit-board assembly line at Motorola allows automatic insertion of discrete components into circuit boards — the plug-in modular 25-machine conveyor line applied 30,000 electrical connections per hour to printed circuit modules used in Motorola Quasar television sets (Luke, 1972). Using four specialized assembly machines developed for Zenith, a single operator can apply more than half a million electrical contacts to more than 25,000 PCBs in one 8-hr shift (Luke, 1972).

Probably one of the most advanced electronics assembly systems currently available is the Olivetti/OSAI SIGMA-series robots (Thompson, 1978). The minicomputer-controlled SIGMA/MTG two-arm model has eight degrees of freedom (total) and a positioning accuracy of 0.15 mm. In PCB assembly, boards are selected individually from a feeding device by a robot hand, then positioned in a holding fixture. This method frees both hands to begin loading integrated circuit (IC) chips into the boards. The robot hands can wiggle the ICs to make them fit if necessary. ICs are given a cursory inspection before insertion, and bad ones are rejected. Assembly rates of 12,500 IC/hr are normally achieved (50 IC/PCB and 250 PCB/hr) for each robot arm pair, 2-3 per human operator. The two arms are programmed to operate asynchronously and have built-in collision avoidance sensors. In other operations, different SIGMA-model robots assemble typewriter parts such as ribbon cartridges, typewriter key cap assemblies, and mechanical key linkages.

The SIGHT-1 computer vision system developed by General Motors' Delco Electronics Division locates and calculates the position of transistor chips during processing for use in car and truck high-energy ignition systems. It also checks each chip for structural integrity and rejects all defectives (Shapiro, 1978). The simple program logic for the IC chip inspection is shown in figure 5.45.

A most serious gap in current technology is in the area of inspection. There are few if any systems for automatic circuit verification — at present, inspection is limited to external integrity and structural irregularities or requires a human presence. At present, neither IC nor PCB performance checking is sufficiently autonomous for purposes of SRS.

Bin Packing for Warehouse Shipment

Bin packing (or crate loading for shipment) is a straightforward problem in robotics provided the parts and crate presentation difficulties have already been solved. SRI International has done a lot of work in this area. For example, using feedback from a proximity sensor and a triaxial force sensor in its "hand," a Unimate robot was

able to pick up individual preassembled water pumps from approximately known positions and pack them neatly in a tote-box. In another experiment boxes were placed randomly on a moving conveyor belt; the SRI vision system determined the position and orientation of each box, and permitted a Unimate robot arm to pack castings into each box regardless of how fast the conveyor was moving (Rosen et al., 1978). At Hitachi Central Research Laboratory, Goto (1972) built a robot "hand" with two fingers, each with 14 outer contact sensors and four inner pressure-sensitive conductive rubber sensors that are able to pick up blocks located randomly on a table and pack them tightly onto a pallet.

A related and interesting accomplishment is the stenciling of moving boxes. In an experiment at SRI International, boxes were placed randomly on a moving conveyor and their position and orientation determined by a vision system. The visual information was used by a Unimate robot to place a stencil on the upper right corner of each box, spray the stencil with ink, then remove the stencil, thus leaving a permanent marking on each box (Rosen et al., 1976). An immediate extension of this technique would be to use the vision module to recognize a particular kind of

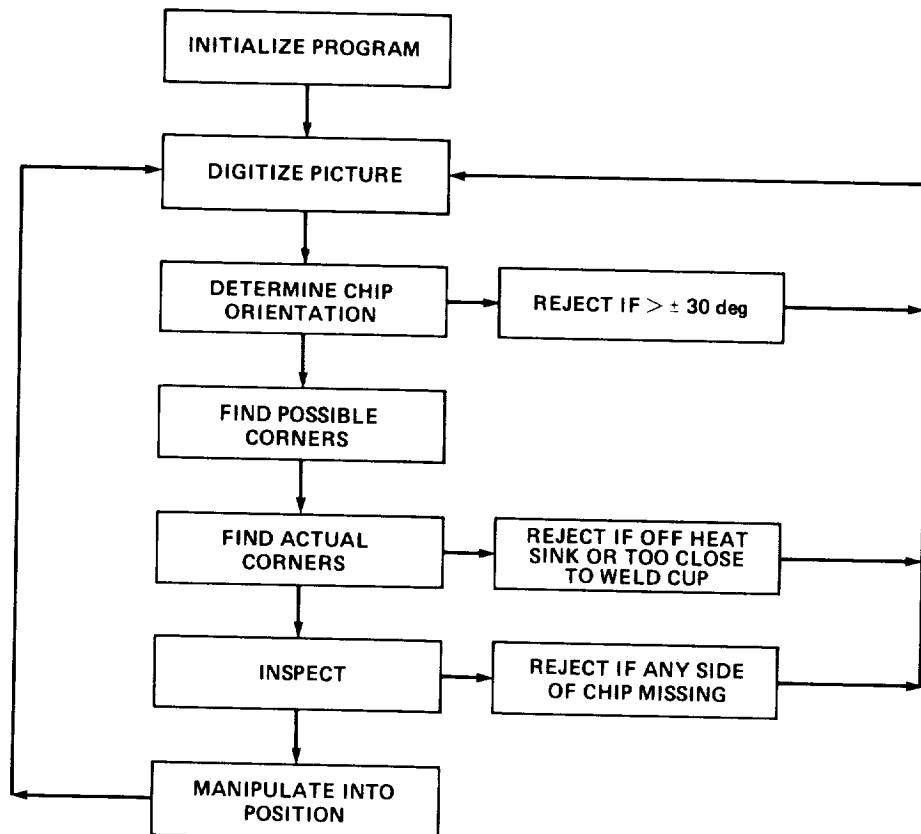


Figure 5.45.— Program logic for the GM/Delco IC "chip" inspection system.

box coming down the conveyor line, and then choose one of many possible stencils which was the "name" of that kind of box. Then the stenciling could be further extended to objects in the boxes, say, parts, in which case the end result would be a robot capable of marking individual objects with something akin to a "universal product code" that warehouse or assembly robots could readily identify and recognize.

Automated Transport Vehicles

Automated Transport Vehicles (ATVs), or "parts carts," are responsible for physically moving parts and subassemblies between sectors, between robot assembly stations, and in and out of warehouses in various locations throughout the LMF. Mobile carriers of the sophistication required for the lunar seed do not exist, but should be capable of development within a decade given the present strong interest in developing totally automated factories on Earth.

Luke (1972) describes a tow-cart system designed by SI Handling Systems, Inc., for use in manufacturing plants. These "switch-carts" serve as mobile workbenches for assembly, testing and inspection, and for carrying finished products to storage, shipping areas, or to other work areas. Carts can be unloaded manually or automatically, or loaded, then "reprogrammed" for other destinations. However, these carts are passive machines -- they cannot load or unload themselves and they have no feedback to monitor their own condition (have they just tipped over, lost their load, had a load shift dangerously, etc.?) They have no means of remote communication with a centralized source of control, and all destination programming is performed manually. The ideal system would include vision and touch sensors, a loading/unloading crane, vestibular or "balance" sensors, an onboard microcomputer controller, and a radio link to the outside. This link could be used by the ATV to periodically report its status, location, and any malfunctions, and it could be used by the central factory computer to inform the ATV of traffic conditions ahead, new routes, and derailed or damaged machines ahead to avoid or to assist.

A major step forward was the now legendary "Shakey" robot, an SRI project during 1968-1972 (Raphael et al., 1971). Shakey was, in essence, a prototype mobile robot cart equipped with a TV camera, rangefinder, and radio link to a central computer. The system could be given, and would successfully execute, such simple tasks as finding a box of a certain size, shape, and color, and pushing it to a designated position. The robot could form and execute simple plans for navigating rooms, doorways, and floors littered with the large blocks. Shakey was programmed to recover from certain unforeseen circumstances, cope with obstacles, store (learn) generalized versions of plans it produced for later use, and to execute preliminary actions and

pursuance of principal goals. (In one instance, Shakey figured out that by moving a ramp a few feet it could climb up onto a platform where the box it needed to move was resting.) The robot also carried out a number of manipulative functions in cooperation with a Unimate robot arm -- Shakey had no manipulators of its own.

Work of a similar nature is now in progress in French laboratories. For example, the mobile robot HILARE is a modular, triangular, and computer-controlled mobile cart equipped with three wheels (two of them motor-driven), an onboard microcomputer, a sophisticated sensor bank (vision, infrared, ultrasonic sonar/proximity, and telemetry laser), and in the future a manipulator arm will be added (Prajoux, 1980). HILARE's control systems include "expert modules" for object identification, navigation, exploration, itinerary planning, and sensory planning.

The Japanese have also made significant progress in this area. One design is an amazing driverless "intelligent car" that can drive on normal roads at speeds up to 30 km/hr, automatically avoiding stationary obstacles or stopping if necessary (Tsugawa et al., 1979). Other Japanese mobile robot systems under development can find pathways around people walking in a hallway (Tsukiyama and Shirai, 1979), and can compute the relative velocities and distances of cars in real time to permit a robot car to be able to operate successfully in normal traffic (Sato, 1979).

Automated Warehouse Robots

Workpieces and other objects delivered to LMF warehouse facilities for storage must be automatically stowed away properly, and later expeditiously retrieved, by the warehouse robots. Numerous advanced and successful automated warehouse systems have already been installed in various commercial operations. A typical system in use at Rohr Corporation efficiently utilizes space and employs computer-controlled stacker cranes to store and retrieve standardized pallets (Anderson, 1972). The computer keeps records on the entire inventory present at any given time as well as the status of all parts ingoing and outgoing.

Similar techniques were used in the semiautomated "pigeonhole" storage systems for sheet metal and electric motors (in the 3/4 to 30 hp range) first operated by Reliance Steel and Aluminum Company decades ago. Each compartment contained one motor or up to 2250 kg of flat precut aluminum, magnesium, or high-finish stainless or galvanized steel stored on pallets. Retrieval time was about 1 min for the motors and about 6 min for the entire contents of a sheet metal compartment (Foster, 1963; Luke, 1972).

The technology in this area appears not to be especially difficult, although a "custom" system obviously must be designed for the peculiarities of lunar operations.

Mobile Assembly and Repair Robots

A Mobile Assembly and Repair Robot (MARR) must take complex preassembled parts (motors, cameras, microcomputers, robot arms, pumps) and perhaps a limited number of simple parts (bolts, washers, gears, wires, or springs) and assemble complete working LMF machines (mining robots, materials processing machines, warehouse robots, new MARRs). A MARR requires mobility, because it easily permits complex assembly of large interconnected systems and allows finished machines to be assembled *in situ* wherever needed in any LMF sector (Hollis, 1977). A MARR needs full mobility independent of specialized tracks or roadways, a wide range of sophisticated sensors (including stereo vision, IR and UV, radar and microwave, and various contact, contour, and texture sensing capabilities) mounted on flexible booms perhaps 4 m long. MARRs also require at least one "cherry picker" crane, a minimum of two heavy-duty manipulator arms, two light-duty manipulator arms with precision end-effectors, and a wide selection of tools (e.g., screwdrivers, rivet guns, shears, soldering gun, and wrenches). A radio link and onboard computer-controller are also essential.

MARRs have an omnibus mission illustrated by the diversity of the following partial list of tasks:

- Receive assembled subassemblies via automated transport vehicles
- Assemble subassemblies into working LMF machines *in situ* during growth phase(s)
- 100% verification of each final assembly step, with functional checkout as well as structural verification
- Debugging, dry-running, final checkout, and certification of operational readiness of each final assembly
- Repair by diagnostics, followed by staged disassembly if necessary to locate and correct the fault (Cliff, 1981; see appendix 5H)
- Assemble new LMF seeds during replication phase(s)
- Assemble useful products during production phase(s)

According to van Cleve (1977), when General Motors began to consider the design of automated assembly systems for automobiles "the assembly of vehicles was rejected as being too complex for the time being so studies are confined to subassemblies." This area is identified as a major potential technology driver – insufficient research has been conducted on the development of systems for complete automated final assembly of working machines from sub-assemblies in an industrial production setting.

For instance, at General Motors Research Laboratories the most progress made to date is an experimental system to mount wheels on automobiles (Olsztyń, 1973). The location of the studs on the hubs and the stud holes on the

wheels were determined using a TV camera coupled to a computer, and then a special manipulator mounted the wheel on the hub and engaged the studs in the appropriate holes. According to Rosen and Nitzan (1977), "although this experiment demonstrated the feasibility of a useful task, further development is needed to make this system cost-effective." The prospects for semiautonomous assembly robots have recently been favorably reviewed by Leonard (1980).

In Japan, much recent work has dealt with the design and construction of robot "hands" of very high dexterity of the sort which might be needed for fine precision work during delicate final assembly and other related tasks. Takese (1979) has developed a two-arm manipulator able to do tasks requiring cooperation between the arms – such as turning a crank, boring a hole with a carpenter's brace and bit, sawing wood, driving nails with a hammer, and several other chores. Okada (1979), also of the Electrotechnical Laboratory in Tokyo, has devised a three-fingered robot hand of incredible dexterity. Each finger has three joints. The hand of Okada's robot can tighten nuts on a threaded shaft, shift a cylindrical bar from side to side while holding it vertically, slowly twirl a small baton, and rotate a ball while holding it. Further research will extend into more complex movements such as tying a knot, fastening buttons, and using chopsticks.

Although some of the needed technologies for final assembly are slowly becoming available, many are not. Further, no attempt has yet been made to produce a final assembly robot, let alone a truly universal final assembly robot such as the MARRs required for the LMF. Such is a leap beyond even the ambitious Japanese MUM program mentioned in appendix 5F – even MUM envisions a minimum continuing human presence within the factory.

Conceptually, final assembly seems not intractable – a typical machine can be broken down into perhaps a few dozen basic subassemblies. But little research has been done so potential difficulties remain largely unknown. Major problem areas may include verification and debugging, sub-assembly presentation and recognition, actual subassembly interconnection or complex surfaces mating, and heavy lifting; today flexible robot arms capable of lifting much more than their own weight quickly, accurately, and dexterously do not exist.

The MARR system is a major R&D area which must be explored further before LMF design or deployment may practically be attempted.

5G.2 Assembly and LMF Computer Control

As with other sectors, LMF assembly is controlled by a computer which directs the entire factory. The assembly sector minicomputer, on the other hand, directs the many microcomputers which control its various assembly robots.

transport robots, and warehouse robots. The entire manufacturing system is thus controlled by a hierarchy of distributed computers, and can simultaneously manufacture subsets of groups of different products after fast, simple retraining exercises either programmed by an "intelligent" central computer or remotely by human beings. Plant layout and production scheduling are optimized to permit maximum machine utilization and speed of manufacturing, and to minimize energy consumption, inventories, and wastage (Merchant, 1975).

Merchant (1973) suggests that a fully automatic factory capable of producing and assembling machined parts will consist of modular manufacturing subsystems, each controlled by a hierarchy of micro- and minicomputers interfaced with a larger central computer. The modular subsystems must perform seven specific manufacturing functions:

(1) *Product design* by an advanced "expert system" software package or by humans, remotely or interactively, using a computer design system that stores data on models, computes optimal designs for different options, displays results for approval, and allows efficient process iteration.

(2) *Production planning*, an optimized plan for the manufacturing processes generated by a computer on the basis of product-design outputs, scheduling, and line-balance algorithms, and varying conditions of ore-feedstock deliveries, available robot resources, product mix, and priorities. Planning includes routing, timing, work stations, and operating steps and conditions.

(3) *Parts forming* at work stations, each controlled by a small computer able to load and unload workpieces, make parts and employ adaptive control (in-process operation sensing and corrective feedback), and incorporate diagnostic devices such as tool-wear and tool-breakage sensors.

(4) *Materials handling* by different computer-controlled devices such as lifts, warehouse stacking cranes, carts, conveyors, and industrial robots with or without sensors that handle (store, retrieve, find, acquire, transport, load, unload) parts, tools, fixtures, and other materials throughout the factory.

(5) *Assembly of parts* and subassemblies at computer-controlled work stations, each of which may include a table, jigs, industrial robots with or without sensors, and other devices.

(6) *Inspection of parts*, subassemblies, and assemblies by computer-controlled sensor systems during and at the end of the manufacturing process.

(7) *Organization of production information*, a large overseeing computer system that stores, processes, and interprets all manufacturing data including orders; inventories of materials, tools, parts, and products; manufacturing planning and monitoring; plant maintenance; and other factory activities (Nitzan and Rosen, 1976).

Such a completely computer-integrated factory does not yet exist, though various major components of this kind of system have been constructed and are in use in industry in

the United States, Europe, and Japan. The most ambitious plan to reach Merchant's level of full automation is the Japanese MUM program which aims at "unmanned manufacturing" (computer-controlled operations, man-controlled maintenance) in the 1980-1985 time frame and "complete automatic manufacturing" (computer-controlled operations and maintenance) by 2000-2005 (Honda, 1974).

According to advanced planning notes, the most advanced and expensive MUM system would be "metabolic," "capable of being expanded," and "capable of self-diagnosis and self-reproduction.... With a built-in microcomputer, it is a self-diagnosis and self-reproduction system which can inspect functional deteriorations or abnormal conditions and exchange machine elements for identical ones. It is a hierarchy-information system with built-in microcomputer, middle computer, and central control computer. It can alleviate the burden on the central computer, and is capable of rapid disposal in case the computer fails. It is also capable of expansion" (Honda, 1974). Plans to open an automated robot-making factory at Fujitsu in accordance with the MUM philosophy are proceeding smoothly (see appendix 5F).

5G.3 Sector Mass and Power Estimates

A set of mass and power estimates for assembly systems was obtained from several sources and is displayed in table 5.20. Taking the extremes in each range, and given the known required throughput rate to replicate the original LMF seed in 1 year, we find that mass of assembly sector machinery lies between 83-1100 kg and the power consumption between 0.083-19 kW. If the warehouse robots and their fixed plant have a mass of about 1% of the stored goods (parts for an entire 100-ton seed) and a power requirement of about 10 W/kg, their mass is about 1 ton and their power draw about 10 kW.

The automated transport vehicles may have to carry the entire seed mass as often as ten times during the course of a year's growth, replication, or production. This is a hauling rate of 3.2×10^{-2} kg/sec or 0.32 parts/sec. If the average trip for an ATV is 100 m (initial seed diam), with a mean velocity of 1 km/hr (taking account of downtime for repairs, reprogramming, on- and off-loading, rescues, etc.), then the ATV trip time is 360 sec (6 min) and the average load is 11.5 kg/trip or 115 "typical parts"/trip. While a properly designed hauler should be capable of bearing at least its own weight in freight, ATVs require special equipment for manipulation rather than hauling. A conservative estimate for the ATV fleet is 100-1000 kg. If a typical vehicle power consumption is 20 (J/m)/kg (Freitas, 1980), the power requirement for the fleet is 0.56 to 5.6 kW total.

As for MARRs, the "warden" robots in the Project Daedalus BIS starship study (Martin, 1978) served a similar function and were allocated to the main vessel in the amount of 10^{-7} robots/kg-year serviced. To service a

TABLE 5.20.— MASS AND POWER ESTIMATES FOR ASSEMBLY SYSTEMS FROM VARIOUS SOURCES

Source	Plant mass, kg/kg per sec output	Plant power, W/kg plant
Johnson and Holbrow (1977) — Bulk processing and heavy industry estimate for human workers	4.3×10^4	2
Criswell (1980) — for “Cold Macro Assembly”	3.6×10^5	1
PUMA (1980) arm and controller computer, assuming 88 kg mass, 1500 W power, speed 1 part/30 sec assembly, part mass 0.1 kg/part	2.6×10^4	17

100-ton LMF Seed for a century would require one “warden” of mass 1 ton and a power draw of 10 W/kg. Conservatively assigning one MARR each to chemical processing sector, parts and electronics fabrication sectors,

and assembly sector results in a total mass of 4 tons and draws 40 kW of power for the fleet of four MARRs. The main seed computer has a mass of 2200 kg, with 22.2×10^{-2} kg computer/kg serviced as in Martin (1978). With 17 W/kg as for the PUMA robot arm controller computer (Spalding, personal communication, 1980), seed computer power requirements are 37 kW.

5G.4 Information and Control Estimates

The team assumed that the assembly of a typical part may be described by 10^4 bits (about one page of printed text), an extremely conservative estimate judging from the instructions printed in Ford Truck (1960) and Chilton (1971), and especially if the seed has only 1000 different kinds of parts. Thus $(10^4 \text{ bits/part})(10^6 \text{ parts/seed}) = 10^{10}$ bits to permit the assembly sector to assemble the entire initial seed. To operate the sector may require an order less capacity than that needed for complete self-description, about 10^9 bits. Applying similar calculations to other sector subsystems gives the estimates tabulated in table 5.1 — ATVs lie between mining and paving robots in complexity, and warehoused parts, each labeled by 100 bits, require a total of 10^8 bits for identification, and perhaps an order of magnitude less for the computer controller that operates the warehouse and its robots.

5G.5 References

- Abraham, Richard G.; and Beres, James F.: Cost-Effective Programmable Assembly Systems. Paper presented at the 1st North American Industrial Robot Conference, 26–28 October 1976. Reprinted in William R. Tanner, ed., *Industrial Robots*, Volume 2: Applications, Society of Manufacturing Engineers, Dearborn, Michigan, 1979, pp. 213–236.
- Abraham, Richard G.; and Shum, L. Y.: Robot Arc Welder with Contouring Teach Mode. In Proc. 5th International Symposium on Industrial Robots, IIT Research Institute, Chicago, Illinois, September 1975, Society of Manufacturing Engineers, Dearborn, Mich., 1975, pp. 239–258.
- Agin, Gerald J.; and Duda, Richard O.: SRI Vision Research for Advanced Industrial Automation. In 2nd USA-Japan Computer Conference, Session 5-4-5, 1975, pp. 113–117. Proceedings, Aug. 26–28, 1975. American Federation of Information Processing Societies, Montvale, N.J., 1975.
- Ambler, A. P.; Barrow, H. G.; Brown, C. M.; Bonstall, R. M.; Popplestone, R. J.: A Versatile System for Computer-Controlled Assembly. *Artificial Intelligence*, vol. 6, Summer 1975, pp. 129–156.
- Anderson, R. H.: Programmable Automation: The Future of Computers in Manufacturing. *Datamation*, vol. 18, December 1972, pp. 46–52.
- Ansley, Arthur C.: *Manufacturing Methods and Processes*. Chilton Book Company, Philadelphia, 1968. Revised and enlarged edition.
- Bolles, R. C.; and Paul, R.: The Use of Sensory Feedback in a Programmable Assembly System. Computer Science Department, Stanford Univ., Stanford, California, October 1973. Stan-CS-73-396, AD-772064.
- Chilton's Auto Repair Manual, 1964–1971: Chilton Book Company, Philadelphia, 1971.
- Clarke, Arthur C.: *The Promise of Space*. Harper and Row Publ., New York, 1968.
- Cliff, Rodger A.: An Hierarchical System Architecture for Automated Design, Fabrication, and Repair. Paper presented at the 5th Princeton/AIAA/SSI Conference on Space Manufacturing, 18–21 May 1981, Princeton, NJ.
- Criswell, David R.: Extraterrestrial Materials Processing and Construction. NSR 09-051-001 Mod. 24, Final Report, 31 January 1980.

- Ejiri, M. *et al.*: An Intelligent Robot with Cognition and Decision-Making Ability. Proc. 2nd Intl. Joint Conf. on Artificial Intelligence, London, Sept. 1971, British Computer Society, London, 1971, pp. 350-358.
- Feldman, J. *et al.*: The Use of Vision and Manipulation to Solve the Instant Insanity Puzzle. Proc. 2nd Intl. Joint Conf. on Artificial Intelligence, London, Sept. 1971. British Computer Society, London, 1971, pp. 359-364.
- Ford Truck Shop Manual: Ford Motor Company, 1960.
- Foster, David B.: Modern Automation. Pitman, London, 1963.
- Freitas, Robert A., Jr.: A Self-Reproducing Interstellar Probe. J. British Interplanet. Soc., vol. 33, July 1980, pp. 251-264.
- Goto, T.: Compact Packaging by Robot with Tactile Sensors. Proc. 2nd Intl. Symp. Industrial Robots, Chicago, 1972. IIT Research Institute, Chicago, Illinois, 1972.
- Goto, T.; Inoyama, T.; and Takeyasu, K.: Precise Insert Operation by Tactile Controlled Robot HI-T-HAND Expert-2. Proc. 4th Intl. Symp. Industrial Robots, Tokyo, November 1974, pp. 209-218.
- Hart, Peter E.: Progress on a Computer Based Consultant. SRI International Publication 711, January 1975.
- Heginbotham, W. B.; Kitchin, P. W.; and Pugh, A.: Visual Feedback Applied to Programmable Assembly Machines. Proc. 2nd Int. Symp. Industrial Robots, Chicago, 1972. IIT Research Institute, Chicago, Illinois, 1972, pp. 77-88.
- Heppenheimer, T. A.: Colonies in Space. Stackpole Books, PA, 1977.
- Hollis, Ralph: NEWT: A Mobile, Cognitive Robot. Byte, vol. 2, June 1977, pp. 30-45.
- Honda, Fujio, ed.: Methodology for Unmanned Metal Working Factory. Project Committee of Unmanned Manufacturing System Design, Bulletin of Mechanical Engineering Laboratory No. 13, Tokyo, 1974.
- Inoue, H.: Computer Controlled Bilateral Manipulator. Bull. Japanese Soc. Mech. Eng., vol. 14, March 1971, pp. 199-207.
- Johnson, Richard D.; and Holbrow, Charles, eds.: Space Settlements: A Design Study, NASA SP-413, 1977. 185 pp.
- Kirsch, Jerry: Progression of Intelligence in Limited Sequence Robots. Paper presented at the 1st North American Industrial Robot Conference, 26-28 October 1976.
- Leonard, Raymond S.: Automated Construction of Photovoltaic Power Plants. Paper prepared as in-house document, Bechtel National, Inc., San Francisco, California, 1980. 36 pp.
- Luke, Hugh D.: Automation for Productivity. Wiley, New York, 1972.
- Martin, A. R., ed.: Project Daedalus – The Final Report on the BIS Starship Study. British Interplanetary Soc., 1978. (J. British Interplanetary Soc., Supplement 1978.)
- McGhie, Dennis; and Hill, John W.: Vision-Controlled assembly Station. Paper delivered at Robots III Conference, Chicago, Illinois, 7-9 November 1978.
- Merchant, M. Eugene: The Future of CAM Systems. National Computer Conference and Exposition, Anaheim, 1975. AFIPS Conference Proceedings, vol. 44, 1975, pp. 793-799.
- Merchant, M. E.: The Future of Batch Manufacture. Phil. Trans. Royal Soc. London, vol. 275A, 1973, pp. 357-372.
- Nevins, James L.; and Whitney, Daniel E.: Computer-Controlled Assembly. Scientific American, vol. 238, February 1978, pp. 62-74.
- Nitzan, David: Robotic Automation at SRI. Proc. of IEEE Midwest Conference. MIDCON/79, Chicago, Illinois, 6-8 November 1979. Western Periodicals, Hollywood, California, 1979, Paper 5.1.
- Nitzan, D.; Brain, A. E.; and Duda, R. O.: The Measurement and Use of Registered Reflectance and Range Data in Scene Analysis. Proc. IEEE, vol. 65, February 1977, pp. 206-220.
- Nitzan, David; and Rosen, Charles A.: Programmable Industrial Automation. IEEE Trans. Computers, vol. C-25, December 1976, pp. 1259-1270.
- Nitzan, D.; Rosen, C.; Agin, G.; Bolles, R.; Gleason, G.; Hill, J.; McGhie, D.; Prajoux, R.; Park, W.; and Sword, A.: Machine Intelligence Research Applied to Industrial Automation. 9th Report, SRI International, August 1979.
- Okada, Tokuji: A Versatile End-Effector with Flexible Fingers. Robotics Age, vol. 1, Winter 1979, pp. 31, 3-39.
- Olszty, J. T., *et al.*: An Application of Computer Vision to a Simulated Assembly Task. Proc. 1st International Joint Conference on Pattern Recognition, Washington, D.C., 1973, pp. 505-513.
- Perkins, W. A.: Multilevel Vision Recognition System. 3rd International Joint Conference on Pattern Recognition, Coronado, Calif., 8-11, 1976. New York, IEEE, 1976, pp. 739-744.
- Perkins, W. A.: Model-Based Vision System for Scenes Containing Multiple Parts. General Motors Research Laboratories Publication GMR-2386, June 1977a.
- Perkins, W. A.: A Model-Based Vision System for Industrial Parts. General Motors Research Laboratories Publication GMR-2410, June 1977b.
- Prajoux, Roland: Robotics Research in France. Robotics Age, vol. 2, Spring 1980, pp. 16-26.
- Raphael, B. *et al.*: Research and Applications – Artificial Intelligence. NASA CR-131991, 1971.
- Rosen, Charles A.: Machine Vision and Robotics: Industrial Requirements. In Computer Vision and Sensor-Based Robots, George G. Dodd, Lothar Rossol, eds., Plenum Publ. Co., 1979, pp. 3-20, 20-22 (discussion).

- Sub.
fer.* ; A.; and Nitzan, David: Use of Sensors in
ble Automation, vol. 10, December 1977,
- In* .; Agin, C.; Andeen, G.; and Berger, J.: Machine
ice Research Applied to Industrial Automation.
port, SRI International, November 1976.
827/8, NSF/RA-761655.
- J.; Nitzan, D.; Agin, G.; Bavarsky, A.; Gleason, G.;
, J.; McGhie, D.; and Park, W.: Machine Intelligence
research Applied to Industrial Automation. 8th Report,
SRI International, August 1978.
- ato*, T.: Automotive Stereo Vision Using Deconvolution
Technique. From 6th Intl. Joint Conf. on Artificial
Intelligence, Tokyo, Japan, 1979. Stanford Univ.,
Computer Science Dept., Stanford, Calif., 1979.
- Seko, K.; and Toda, H.: Development and Application
Report in the Arc Welding and Assembly Operation by
the High Performance Robot. In Proc. 4th Intl. Symp.
Industrial Robots, Tokyo, Japan, November 1974,
pp. 487-596. Japan Industrial Robot Assn., Tokyo,
1974.
- Shapiro, Sydney F.: Digital Technology Enables Robots to
See. Computer Design, January 1978. Reprinted in
William R. Tanner, ed., Industrial Robots, Volume 2:
Applications, Society of Manufacturing Engineers,
Dearborn, Michigan, 1979, pp. 271-276.
- Takeda, S.: Study of Artificial Tactile Sensors for Shape
Recognition — Algorithm for Tactile Data Input. Proc.
4th Intl. Symp. Industrial Robots, Tokyo, Japan,
November 1974, Japan Industrial Robot Assn., Tokyo,
1974, pp. 199-208.
- Takese, Kunikatsu: Force Control of a Multi-Jointed Robot
Arm. Robotics Age, vol. 1, Winter 1979, pp. 30, 32-36.
- Thompson, Terrence: I See, Said the Robot. Assembly
Engineering, Hitchcock Publ. Co., 1978. Reprinted in
William R. Tanner, ed., Industrial Robots, Volume 2:
Applications, Society of Manufacturing Engineers,
Dearborn, Michigan, 1979, pp. 265-270.
- Tsugawa, S. *et al.*: An Automobile with Artificial Intelli-
gence. Paper delivered at 6th Intl. Joint Conf. on Artifi-
cial Intelligence, Tokyo, Japan, 1979. International
Joint Conference on Artificial Intelligence, 1979,
pp. 893-895.
- Tsukiyama, T.; and Shirai, Y.: Detection of the Movements
of Men for Autonomous Vehicles. From 6th Intl. Joint
Conf. on Artificial Intelligence, Tokyo, 1979. Interna-
tional Joint Conference on Artificial Intelligence, 1979.
- Van Cleave, David A.: One Big Step for Assembly in the
Sky. Iron Age, vol. 28, November 1977. Reprinted in
William R. Tanner, ed., Industrial Robots, Volume 2:
Applications, Society of Manufacturing Engineers,
Dearborn, Michigan, 1979, pp. 209-212.
- Will, P. M.; and Grossman, D. D.: An Experimental System
for Computer Controlled Mechanical Assembly. IEEE
Transactions on Computers, vol. C-24, September 1975,
pp. 879-888.
- Winston, P. H.: The MIT Robot. In Machine Intelligence,
B. Meltzer, D. Michie, eds., vol. 7, Edinburgh Univ.
Press, 1972.
- Yachida, M.; and Tsuji, S.: A Machine Vision for Complex
Industrial Parts with Learning Capacity. Proc. 4th Intl.
Joint Conference on Artificial Intelligence, Tbilisi, Sept.
1975. Artificial Intelligence Laboratory, Publications
Dept., MIT, Cambridge, Mass., 1975.
- Yoda, H.; Ikeda, S.; Ejiri, M.: New Attempt of Selecting
Objects Using a Hand-Eye System. Hitachi Review,
vol. 22, no. 9, 1973, pp. 362-365.

APPENDIX 5H

HIERARCHICAL SYSTEM ARCHITECTURE FOR AUTOMATED DESIGN, FABRICATION, AND REPAIR

The desire to create the self-replicating telefactor (Bekey and Naugle, 1980; Heer, unpublished draft notes of the Pajaro Dunes Goal Setting Workshop, 1980) leads to a number of interesting systems design problems. Early theoretical work by von Neumann (1966) showed that self-replicating machines could in principle be built. Laing (1975, 1977) has further elaborated this theme in a novel approach to the problem. Practical considerations in the creation of self-replicating machines have been treated by von Tiesenhausen and Darbro (1980). Freitas (1980) and Valdes and Freitas (1980) have dealt with the application of self-replicating machines to the exploration of deep space.

This appendix presents an architecture for a system which can perform automated design, fabrication, and repair of complex systems. This methodology should be a useful component of any self-replicating system.

SH.1 System Level Architecture

This section describes the architecture of a hierarchical fabrication system which starts with raw materials and outputs finished products. At the system level, each layer or “rank” of the hierarchy looks just like any other rank; however, the internal details of the various ranks may be entirely different. The present approach was inspired by Miller (1978). Figure 5.46 shows the basic system architecture which consists, from left to right, of rank upon rank of fabricators. Adjacent ranks of fabricators are separated by a transportation and communication subsystem. Ultimately, final products issue from the system at the extreme right.

For generality, the transportation and communication subsystems for each rank are shown to be disjoint. Indeed, it is evident that the subsystem T_1 which handles raw materials, such as ores, will no doubt differ in detail from the

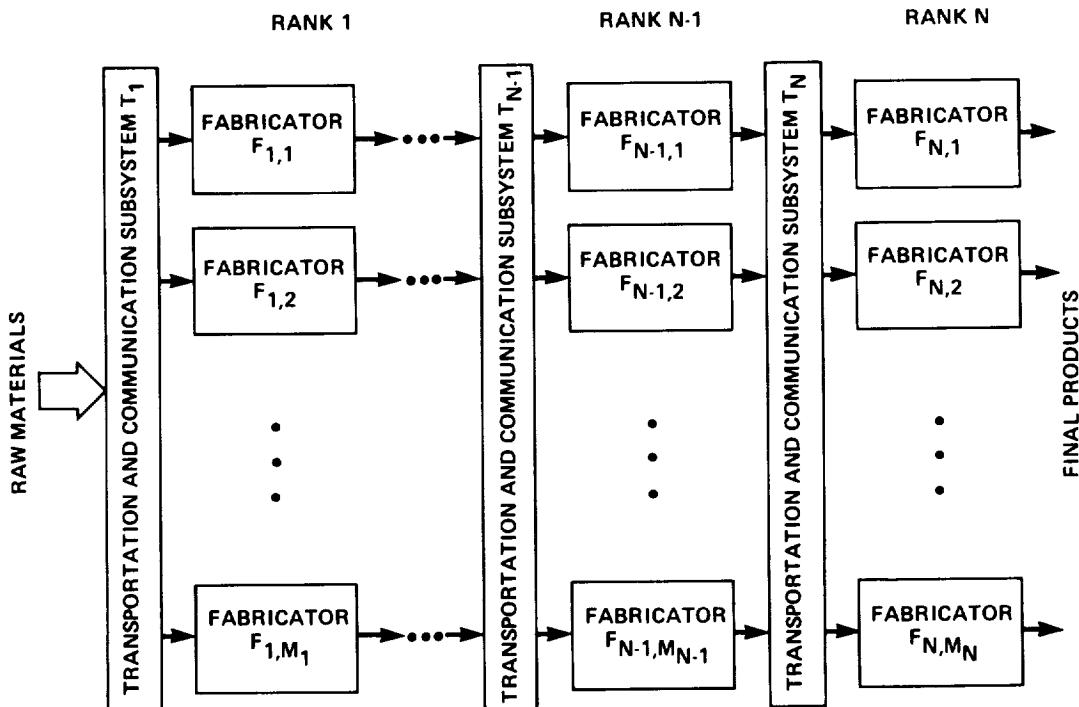


Figure 5.46.—Basic hierarchical system architecture.

subsystem T_L which handles electronic parts, such as electronic circuits. Furthermore, they will both differ from the system T_N which handles major subassemblies, such as complete power plants or complete computer systems. However, at the system level, each transportation and communication subsystem performs the same function: T_L handles the transfer of information and material between the fabricators at rank L and those at rank $L+1$ as shown in figure 5.47.

Although the internal details of the transportation and communication subsystem, T_L , need not concern us here, we shall consider how they appear logically to their adjacent ranks of fabricators. To the fabricators they look like a

random-access, nonblocking switching network for information (e.g., our telephone system), and like a network of roads and delivery trucks for products. In other words, each message or product is dispatched from a fabricator with a unique address (e.g., telephone number or street address) attached to it. It is the function of T_L to see that the messages and products reach their specified destinations.

In figure 5.47, it will be seen that a fabricator at rank $L+1$ can request (via the status links) information on the types of product and their availability from each of the fabricators at rank L . Then the fabricator at rank $L+1$ transmits orders to the rank L fabricators for those products it requires as inputs to its process. The fabricators at

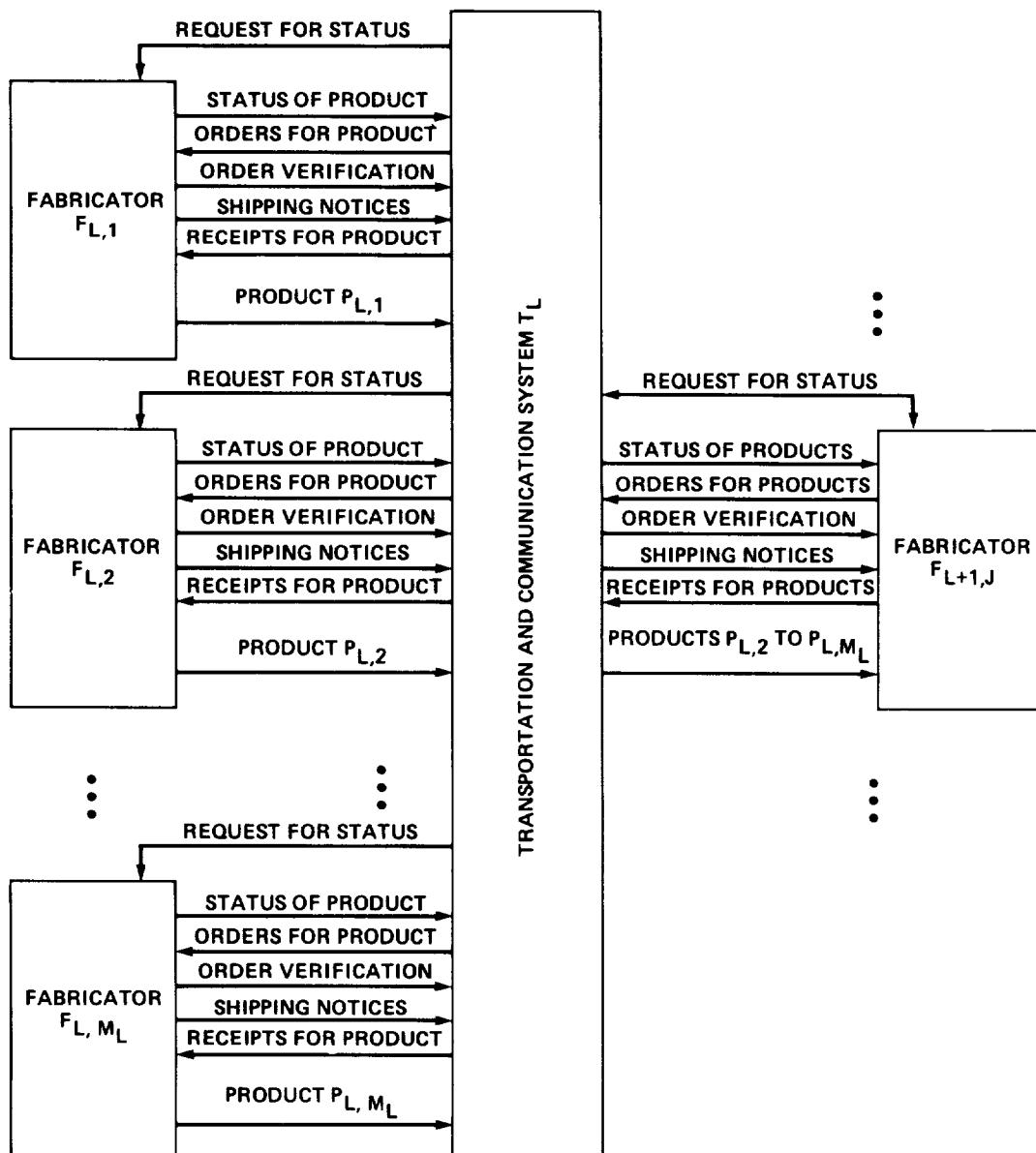


Figure 5.47.—Interrank interface.

rank L respond with order verifications which include expected time of shipment (to aid scheduling at rank $L+1$). Physical transportation of product through T_L is expected to be slow compared to information transfer; therefore, provision has been made for the transmission of shipping notices. Although this is logically redundant it can aid error recovery if the physical transport of materials goes awry. When the product arrives at its rank $L+1$ destination a receipt is returned to the rank L shipper to complete the transaction.

It may appear that the complexity of this interrank interface is not absolutely necessary. While strictly speaking this may be so, it is intended to aid in error recovery and to facilitate adaptive multipurpose behavior throughout the entire system. Indeed the multiple hierarchical feedback loops are borrowed both from living organisms and from human industrial economies (Miller, 1978).

Although some modest amount of diversity is expected in the transportation and communication subsystems, the fabricators are expected to be extremely diverse. They will range from ore smelters, to rolling mills, to high precision lithography for integrated circuits, to final assembly of complex products. The next section of this appendix will show, however, that at the system level all fabricators are essentially similar.

5H.2 Fabricator Morphology

Although great diversity of fabricators is envisioned, relatively few subsystem level primitives are required regardless of the rank L of the fabricator and these primitives are common to all ranks L .

A basic (Morph I) fabricator node appears as figure 5.48. It consists of six subdivisions which correspond to the production oriented parts of a business. The arrows are intended to denote material flow. Associated with each arrow, but not shown, are bi-directional information exchanges analogous to those associated with intra-fabricator communication in figure 5.47. The intra-fabricator communications can in many instances be simpler because of tighter coupling between the communicating entities. Feedback and ability to manage error recovery should be preserved, however.

Material enters a rank $L+1$ fabricator node through a receiver which places orders for those rank L products

which have been requested by the kitter. The receiver handles the shipping protocol for interface with transportation and communication subsystem T_L .

The kitter is driven by the parts list for whatever product the fabricator is to produce. The kitter orders the parts (through the receiver) and provides complete kits to the producer.

The producer converts the kitted parts with which it is provided into the output product of the fabricator node. This conversion could be a chemical process such as ore smelting or a mechanical process such as milling or PCB assembly.

After the product is produced by the producer, it is tested by the tester. In the basic Morph I, fabricator node product which does not pass the test is simply expelled as waste. Product which passes the test is sent to the stocker which accumulates it for shipment.

The shipper handles the protocol with the rank $L+1$ transportation and communication subsystem. It transmits status information, receives orders, and ships product.

Communication between the transportation and communication subsystems of adjacent ranks can be provided by degenerate fabricator nodes as shown in figure 5.49. Such a fabricator node might consist, as a minimum, of only a receiver and a shipper. This is sufficient to interface the protocols of the two different transportation and communication systems. In the simplest case it only provides a buffer. In a slightly more complex case, perhaps a repackaging or aggregation of product is performed. Kitters and/or stockers could be added, to perform these functions. Note that a Morph IIB fabricator node provides a material flow from higher to lower ranks. This capability is useful in dealing with recyclable scrap.

Figure 5.50 shows how a fabricator node could deal with product which fails its test, but which can profitably be recycled at some lower rank. The additional stocker and shipper can send the failed product, through a series of Morph IIB nodes, to an appropriate rank for recycling. This could be applied, for instance, to a milled part which is out of tolerance and can be recycled at less cost than producing an equivalent amount of material from raw ore.

If, on the other hand, the production process is reversible (such as putting a number of printed circuit cards into a card cage) then a more advantageous approach is shown (fig. 5.51). The disassembler performs the inverse of the

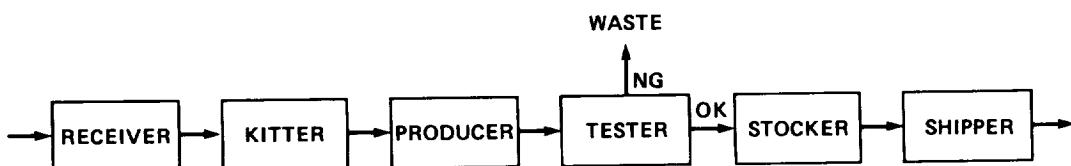


Figure 5.48.—Morph I fabricator node.

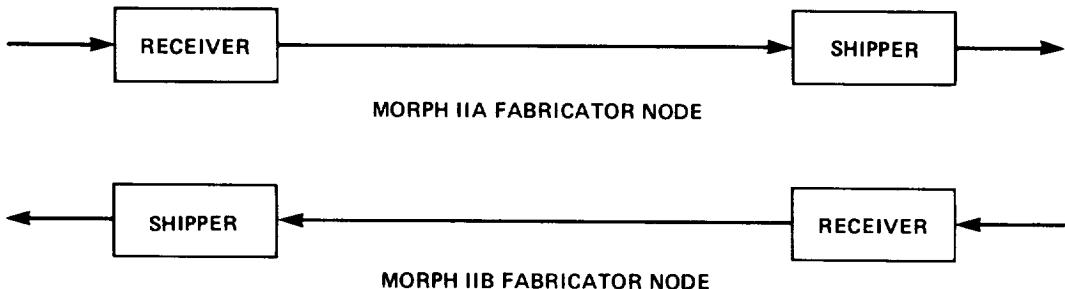


Figure 5.49.—Morph II fabricator node.

production process. The subassemblies are then returned whence they came to be re-tested as subassemblies. Also shown in the lower right hand corner of the figure are a receiver and stocker for returned (potentially faulty) product from the next higher rank.

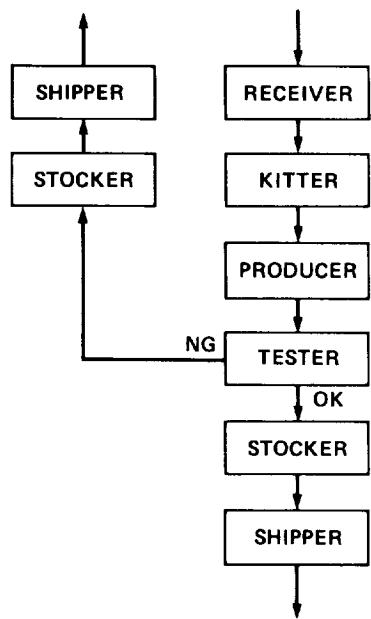


Figure 5.50.—Morph III fabricator node.

One could conceive of doing incoming inspection at the receiver; however, this does not seem to be the best approach. It seems more effective to associate production and test together in the same node (and thus the same rank) since they use much the same information. Furthermore, it seems wasteful to test product both after production and after shipment. This is especially true because the test capability would have to be duplicated at several receivers. Accordingly, for this architecture, test is uniquely associated with production, and suspect product is sent back to its producer for re-test.

We now consider a further embellishment of the fabricator node. Rather than entirely disassembling a faulty product, it may be profitable to rework it. Rework is in general much harder than simple disassembly because it can entail fairly complex diagnosis and repair. We deal here with a simple form of rework, namely, part exchange, as shown in figure 5.52.

In a Morph IV fabricator node, product which fails test (again a cage full of printed circuit cards is a good example) has its parts replaced one at a time. After each replacement, the product is tested again. If it passes it goes to the stocker and the faulty part is recycled or becomes waste. If the product fails again, a different part is exchanged, and so on until the culprit is identified. The decision of which part to exchange can be made on the basis of diagnostic tests, or it can be made at random.

5H.3 Automated Repair

The above has described a system level architecture of an automated fabrication facility. That architecture incorporated a test function after each production function as a way to catch errors as close to the source as possible and to prevent wasted effort. The architecture included paths for recycling product which failed its test. This also was to promote economy. A side effect of this architecture (which was designed only for efficient fabrication) is that it can also perform automated repair. Depending on circumstances this repair can be effected by selective disassembly or selective rework.

First consider selective disassembly for a suspect computer system that needs repair. For the purposes of example it is assumed that the final product is a computer system consisting of four racks of card cages and that each card cage contains 32 printed circuit boards. The PCBs are populated by 11 integrated circuit (IC) types. In order to simplify the discussion, inter-rack cabling, the card cages themselves (with back planes), and power supplies are not considered. Clearly, these can be accommodated in the same way as the components explicitly considered. It is further assumed that plugging ICs into PCBs, plugging PCBs

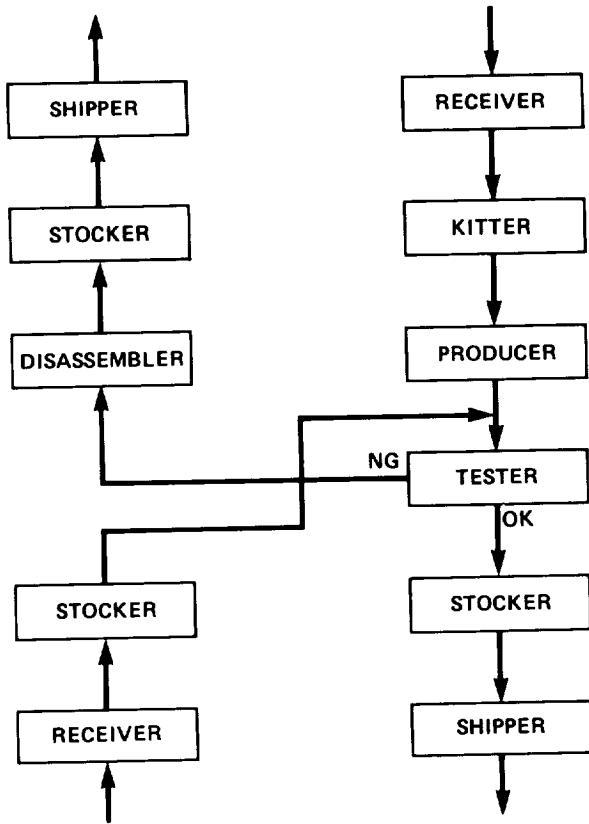


Figure 5.51.—Morph IV fabricator node.

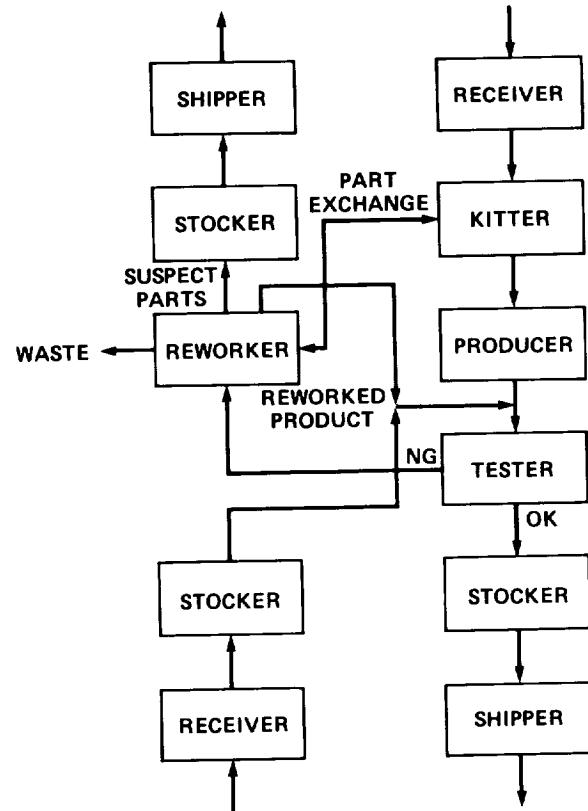


Figure 5.52.—Morph V fabricator node.

into card cages, placing card cages into racks (with interconnection), and cabling the racks together are reversible processes. Furthermore, since the Morph IV fabricator node is the simplest which can accept suspect product from a higher rank source, it is assumed that Morph IV fabricator nodes are used to perform the aforementioned assembly operations. Since IC manufacture is inherently nonreversible (one does not get useful subassemblies by disassembling a finished chip), ICs are assumed produced by a somewhat degenerate Morph IV fabricator node which simply scraps product which fails test rather than disassembling it. Finally, it is also assumed that the entire production facility is idle, there being no current need for additional finished computers.

The repair process begins when the suspect computer is presented to the fabricator node which produced it. This is node $F_{5,1}$ of figure 5.53. In this figure only those parts of the fabrication network which actually participate in the repair are shown. At node $F_{5,1}$ the computer is tested and determined to indeed be faulty. It is then disassembled into four component racks, each of which is sent to the node which produced it. At these nodes the individual racks are tested and in the example the third one is found to be

faulty. Then this rack is disassembled into card cages, etc.

The selective disassembly proceeds until finally at node $F_{1,3}$ a faulty IC is found. With say, 100 ICs per card, 32 cards per cage, 10 cages per rack, and 4 racks per computer, this process has in rather straightforward manner isolated the *one* out of 128,000 ICs that was faulty. Node $F_{1,3}$ provides another IC of that type and scraps the faulty one.

Meanwhile, node $F_{5,1}$, having tested a computer which failed, has placed orders with nodes $F_{4,1}$ through $F_{4,4}$ for a set of racks from which to fabricate a replacement computer. Nodes $F_{4,1}$, $F_{4,2}$, and $F_{4,4}$ return their racks to node $F_{5,1}$. Node $F_{4,3}$, however, places orders for a set of card cages from which to fabricate a replacement for the third rack. Cages 1, and 3 through 10, are returned to node $F_{4,3}$ after testing, while cage 2 is disassembled. The PCBs from cage 2 are pulled out and sent to their respective fabrication nodes. Boards 1 through 12 and 14 through 32 are returned to node $F_{3,2}$. Meanwhile, board 13 is disassembled and the ICs are tested at nodes $F_{1,1}$ through $F_{1,11}$. (It is assumed here that of the roughly 100 ICs per card there are only eleven different types.)

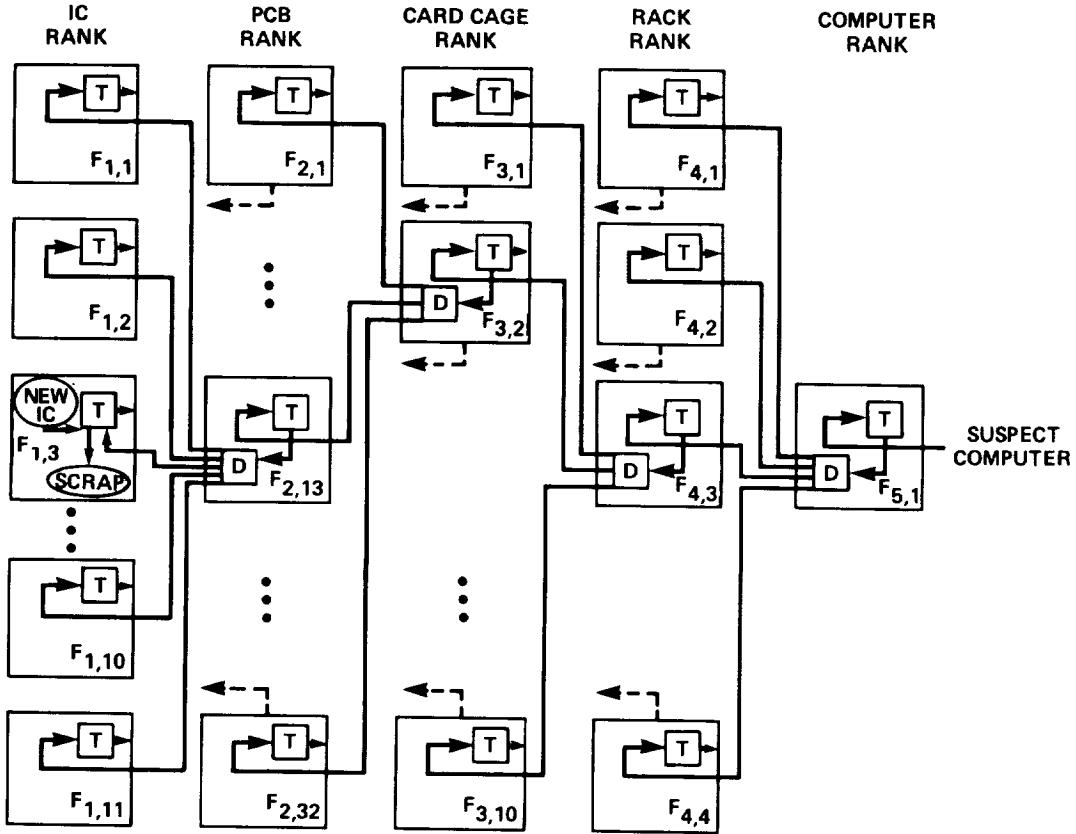


Figure 5.53.— Selective disassembly of failed system.

Now final reassembly can commence as shown in figure 5.54. The heavy lines in the figure trace the path of the replacement IC back into a repaired computer. When the replacement IC reaches node $F_{2,13}$, PCB 13 is reassembled. When the PCB rack reaches node $F_{3,2}$, card cage 2 is reassembled. When this card cage reaches node $F_{4,3}$, rack 3 is reassembled. And, finally, when rack 3 reaches node $F_{5,1}$, the original computer reappears with the one faulty IC replaced.

If, on the other hand, the production facility had been in use when the suspect computer was presented to node $F_{5,1}$, then the original computer would not re-emerge at node $F_{5,1}$. Instead, one more new computer would be produced. The subassemblies obtained by the selective disassembly of the failed computer would be incorporated into many different new computers. However, the failed computer plus one new IC will have resulted in the construction of one new computer to replace the failed one.

In either case, the replacement of the faulty computer is automatic — a consequence of the fabrication system architecture. No additions to the architecture were required to obtain the repair function. Similar arguments hold true if Morph V fabricator nodes are used, but selective rework, rather than selective disassembly, occurs.

SH.4 Automated Design

The system architecture described above readily lends itself to top-down modular design techniques. Using this discipline, at each level of detail, a designer (usually a human) receives a specification for the product he is to design. He then consults a catalog of available lower level products and selects those to be incorporated into his design. If he needs a lower level product which is not available, he generates specifications for that lower level product. The designer also generates assembly instructions for his product and a parts list. The assembly instructions go to the production department and the parts list to the procurement department. The original specification is used by the test department to verify that the product is what was originally requested.

Figure 5.55 shows how such a design function is added to a fabricator node. The designer in this node functions in a capacity analogous to that of a human designer. Automated design is a function which requires a fairly intelligent machine. Indeed, this is a topic of active current interest in the machine intelligence community.

Although the figure does not show the details of the interface, it is intended that the designer of rank L query

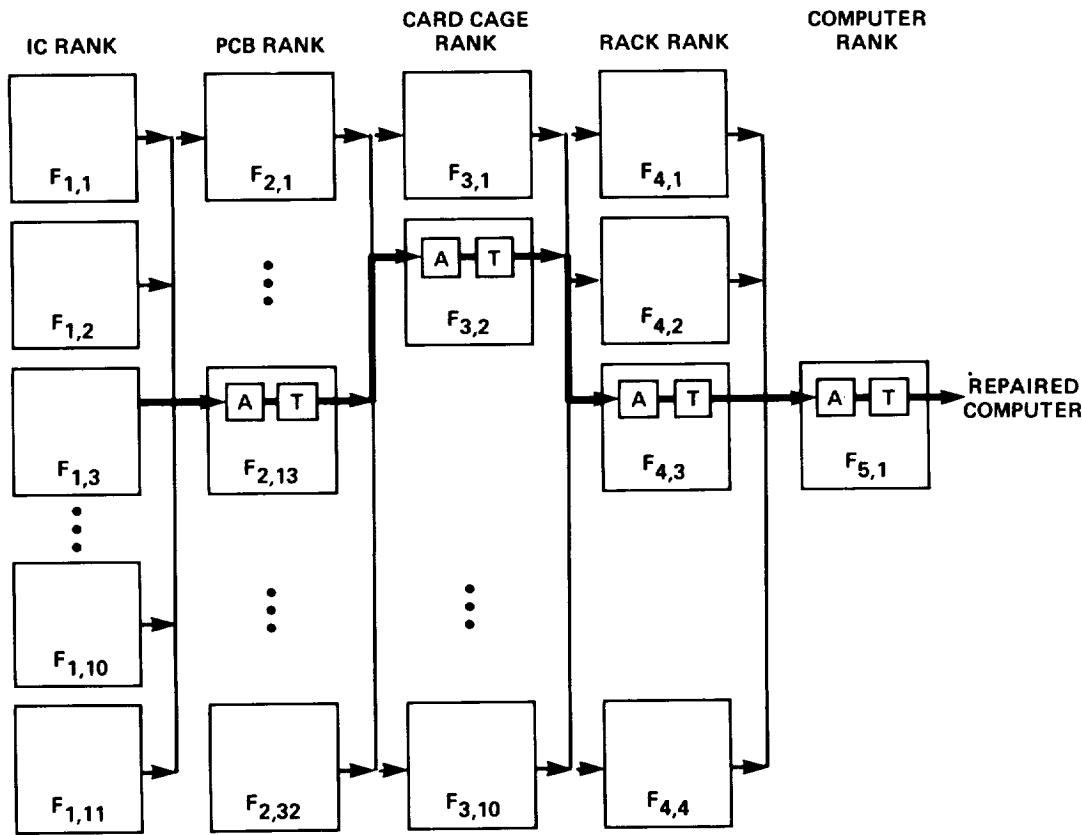


Figure 5.54.— Reassembly of repaired system.

the nodes of rank $L-1$ to ascertain what rank $L-1$ products are available, in much the same way as the kitter does. If the designer finds it needs a product which is not available, it chooses a free node at rank $L-1$ and sends to it the specification for the required product. If there are no free nodes, then the system as configured is not large enough to produce the desired final product. As the system architecture presently stands, outside intervention is required at this point.

5H.5 Utility of Node Morph Architecture

The node morphs enumerated in this paper are not meant to be exhaustive. Compound node morphs may be made by combining elements of two or more morphs. Degenerate morphs from which some functions have been deleted can also be useful. Those morphs enumerated are the ones found useful in the exposition of the system architecture. Additional stocker functions may be desirable in an actual physical system. Potentially useful locations are between the kitter and producer and between the producer and the tester. Also, multiple stockers may be required following a disassembler to handle the number and diversity of

components. Redundant transportation and communication subsystems at each rank would make the overall system more robust, as would redundant fabricator nodes. This could be readily incorporated into a physical system.

This appendix has described the architecture of an automated system which has the following interesting properties:

(1) If it is presented with a final product specification (within its capabilities) it will do the detailed design (all the way down to raw materials if necessary) and then manufacture that product.

(2) If the system is presented with a faulty final product, it will repair it.

Interesting extensions of this architecture would be the ability to add additional fabricator nodes when required and the ability to add entire additional ranks when needed. This is presently under consideration (Cliff, 1981).

The similarities between the system described herein and the industrial complex of a developed nation are fairly obvious and indeed intentional. There are some significant differences, however. The automated system presented here is much more regular: The interrank transportation and communication systems are disjoint, one from another,

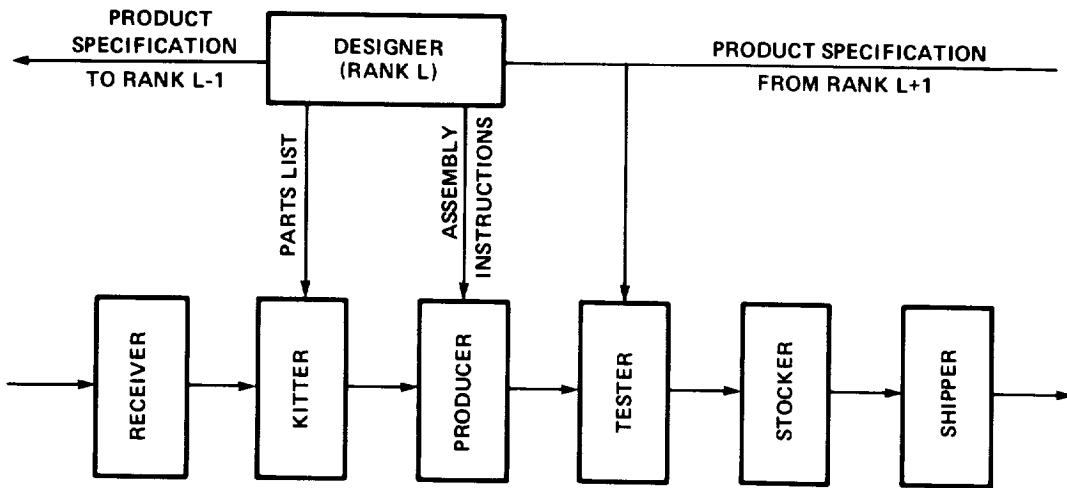


Figure 5.55.—Morph VI fabricator node.

and any one fabricator node makes only one product. The result is that the graph of the system is a lattice, rather than a random network. Furthermore, one could expand the number of nodes in each rank in such a way that the graph of the system becomes a tree. This should facilitate mathematical analysis of the system.

The idea (borrowed from top-down structured programming) is to produce a system which is at once extremely large and complex, but is still both understandable by humans and rigorously analyzable mathematically. This will be especially useful when growth by addition of nodes and ranks is studied.

SH.6 References

- Bekey, Ivan; and Naugle, John E.: Just Over the Horizon in Space. *Astronautics and Aeronautics*, vol. 18, May 1980, pp. 64-76.
- Cliff, Rodger, A.: An Hierarchical System Architecture for Automated Design, Fabrication, and Repair. Paper presented at the 5th Princeton/AIAA/SSI Conference on Space Manufacturing, 18-21 May 1981, Princeton, NJ.
- Freitas, Robert A., Jr.: A Self-Reproducing Interstellar Probe. *J. Brit. Interplan. Society*, vol. 33, July 1980, pp. 251-264.
- Laing, Richard: Some Alternative Reproductive Strategies in Artificial Molecular Machines. *J. Theoret. Biol.*, vol. 54, Oct. 1975, pp. 63-84.

- Laing, Richard: Automation Models of Reproduction by Self-Inspection. *J. Theoret. Biol.*, vol. 66, June 7, 1977, pp. 437-456.
- Miller, James G.: *Living Systems*. McGraw-Hill Book Company, N.Y., 1978.
- Valdes, Francisco; and Freitas, Robert A., Jr.: Comparison of Reproducing and Nonreproducing Starprobe Strategies for Galactic Exploration. *J. of the Brit. Interplan. Soc.* vol. 33, November 1980, pp. 402-408.
- Von Neumann, John: *Theory of Self-Reproducing Automata*, A. W. Burks, ed., Univ. of Illinois Press, Urbana, Illinois, 1966.
- Von Tiesenhausen, Georg; and Darbro, Wesley A.: Self-Replicating Systems — A Systems Engineering Approach. NASA TM-78304, July 1980.

APPENDIX 5I

LMF SOLAR CANOPY POWER SUPPLY

The solar canopy provides electrical power for the entire lunar factory complex described in section 5.3.4. The canopy consists of many sections of automated (active or passive) solar energy collection devices. Mobile robots begin erecting the canopy after a useful fraction of the LMF base platform has been laid down and the central computer system installed in a depression near the hub. The canopy is just a simple framework of lightweight vertical metal/basalt struts snapped into universal connectors bolted into the heavy basalt foundation of the LMF. Horizontal wires or thin crossbeams support the solar panel mechanisms. The solar canopy is designed to be broken into relatively small sections for ease of assembly, installation, maintenance, and repair.

5I.1 LMF and Solar Canopy Geometry

One of the major constraints on LMF shape is the necessity for solar energy collection. The LMF may be visualized geometrically as a very broad, squat cylinder with some net density d_L (kg/m^3), mass (exclusive of platform) M , radius R , and height H . All factory energy is gathered using a “rooftop” surface area approximately the same size as the underlying foundation platform, so the fundamental constraint on factory size may be expressed by the condition $M/\pi R^2 = d_L H \leq MP_s/P$, where P_s is the usable energy delivered to the LMF by its solar collectors (roughly $150 \text{ W}/\text{m}^2$ for high quality photovoltaic devices at 45° angle of incidence) and P is the total power required by the initial lunar seed (about 1.7 MW; see sec. 5.3.4-5).

For a factory mass $M = 10^5 \text{ kg}$, $R \geq 60 \text{ m}$, the figure used elsewhere in this report. Estimates from O'Neill et al. (1980) that solar power systems (SPS) in the 100 kW range can be assembled for 8 kg/kW suggest a total mass for canopy collector panels (1.7 MW) of 13,600 kg, although this figure was derived from space-based SEPS and SPS design studies. A mass of 22,000 kg was adopted for the canopy, which includes transformers, diodes, cabling, and other necessary support devices. Since $d_L H = 8.8 \text{ kg}/\text{m}^2$, the LMF in fact will be quite “roomy” inside – a “typical” population of 1-ton factory machines would be separated by an average distance of $2[10^3/(\pi d_L H)]^{1/2} = 12 \text{ m}$.

Another major factor in determining basic factory configuration is the degree of isolation desired from the external surroundings. There appear to be few compelling

reasons for solid massive walls enclosing a fully automated lunar manufacturing facility. Inclement weather, cleanliness, provision of human-habitable volume, protection from the dangers of seismic activity, noise/pollution abatement, and theft prevention are the usual reasons for heavy walls on Earth, yet these factors should have little if any impact upon factory construction in space or on the Moon. Further, rigid solid walls hinder growth and might delay reconfiguration as the LMF expands in size. The cleanliness problem in an open factory is expected to be minor, as mobile robots are designed either for external or internal operation but not for both (though in special circumstances MARR machines can be towed to external sites by mining robots).

The simplest solar canopy configuration is a web-like metal structure overlaid with flat solar panel assemblies. These cells are suspended from a series of crossbeams spaced at regular intervals along chords of the circular LMF. These crossbeams may be as thin as wires if adequately supported by strategically placed vertical columns. Calculations of stress reveal that a 1 mm radius aluminum rod (typically $10^8 \text{ N}/\text{m}^2$ tensile or compressive strength) should be strong enough to support a 22-ton canopy structure with a loading safety factor of about 5. Support posts are 1-cm diam aluminum/basalt columns placed at intervals of 10 m across the factory floor and anchored with universal connectors and several braces and struts for stabilization. These posts have an overload factor of more than 100, hence should be able to sustain low-speed accidental impacts by mobile robots without buckling. The total mass for the entire framework is well under 1 ton.

Ideally, all lunar operations should be conducted continuously with only scheduled maintenance shutdowns. However, continuous operation is possible only if continuous power is also available. A number of options for power storage during the lunar night have been considered in the context of a lunar base (Criswell, 1979; Vajk et al., 1979) in the literature. Possibilities include nuclear plants, volume heat capacity storage, chemical storage (batteries, fuel cells, exothermic reactants), capacitor banks, gravitational energy storage, pressurized gas, flywheels, and SPS transmission from orbit to lunar surface collection stations. These many promising alternatives, however, were not explored in depth. Without such an option, the baseline LMF must be placed on standby during the lunar night with one working year requiring two calendar years.

SI.2 Solar Canopy vs Lunar Igloo Designs

In the solar canopy LMF design the entire automated factory complex is erected on a fused basalt platform resting on the lunar surface. Above the factory floor is a relatively flimsy framework of solar energy collectors which provide system power.

The "lunar igloo" is an alternative in which geodesic domes of 120 m diam are constructed over each seed factory. Additional factory growth is accommodated by adjacent domes of similar size built with a network of connecting tunnels. Each dome is covered with at least 2–5 m of lunar topsoil which may be sufficient to permit

the retention of an internal 0.3-atm oxygen atmosphere. This configuration might be handy in preventing accidental vacuum welding and could simplify servicing and troubleshooting by humans during system failures. Light could be admitted to the underground LMF via a converging reflective geometry (Hyson, personal communication, 1980).

Since these models represent fundamentally different design concepts (see fig. 5.56), the team compared the two directly on a number of significant factors enumerated in table 5.21. The conclusion was that the canopy model is possibly superior in the present fully automated self-replicating LMF application, but that the igloo model is not precluded in other scenarios.

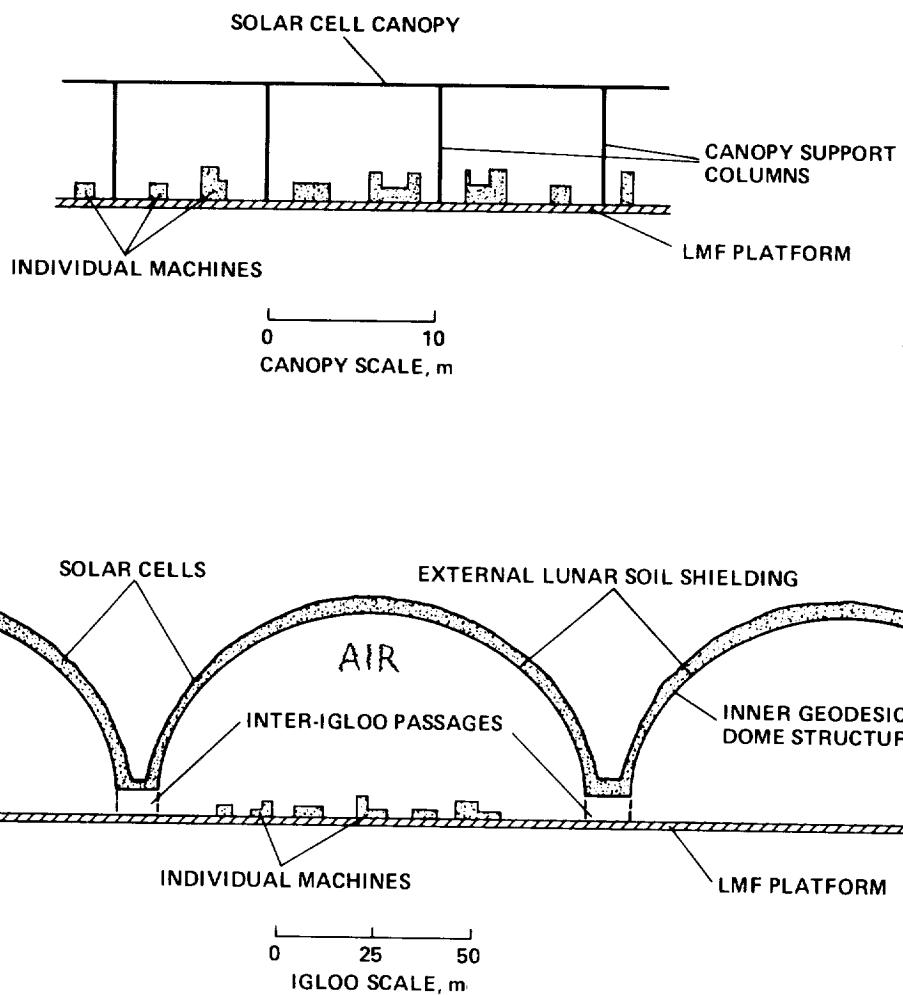


Figure 5.56.— Schematic of Solar Canopy and Lunar Igloo models of self-replicating or growing LMF.

TABLE 5.21.— COMPARISON OF IMPORTANT FACTORS FOR SOLAR CANOPY AND LUNAR IGLOO MODELS OF SELF-REPLICATING OR GROWING LMF

Some important factors	Solar canopy	Lunar igloo
1. Maintain useful atmosphere?	no	yes
2. Maintain useful vacuum?	yes	yes
3. Prevent solar cell degradation?	no	no
4. Prevent external optics degradation?	no	no
5. Prevent internal optics degradation?	no	yes
6. System temperatures easily controlled?	no	yes
7. Low mass foundation structure?	yes	yes
8. Low mass total structure?	yes	no
9. System construction mechanically easy?	yes	less easy
10. Easy maintenance of system integrity?	yes	less easy
11. Internal lighting easily available?	yes	no
12. Human repairman accessible?	yes	yes
13. Human repairman habitable?	no	yes
14. Easy horizontal mass flow?	yes	yes
15. Simplicity of overall system design?	yes	less simple
16. Easy to expand LMF system size/mass?	yes	no
17. Waste heat easily rejected?	yes	no
18. Terrestrial manufacturing processes easily transferred?	less easy	yes

5I.3 References

Criswell, David R.: The Initial Lunar Supply Base. In Space Resources and Space Settlements, John Billingham, William Gilbreath, Brian O'Leary, and Beulah Gossett, eds., NASA SP-428, 1979, pp. 207-224.

O'Neill, G. K.; Driggers, G.; and O'Leary, B.: New Routes to Manufacturing in Space. Astronautics and Aeronautics, vol. 18, October 1980, pp. 46-51.

Vajk, J. P.; Engel, J. H.; and Shettler, J. A.: Habitat and Logistic Support Requirements for the Initiation of a Space Manufacturing Enterprise. In Space Resources and Space Settlements, John Billingham, William Gilbreath, Brian O'Leary, and Beulah Gossett, eds., NASA SP-428, 1979, pp. 61-83.

APPENDIX 5J

COMPLEXITY AND LEVEL OF DETAIL IN ROBOT PROGRAMMING

Programming one robot manipulator to assemble another robot manipulator is not a trivial task. The finest level of detail involves controlling the individual steps of the various stepping motors which in turn control the independent degrees of freedom of the manipulator.

In order to gain some insight into robot programming, the Replicating Systems Concepts Team visited Dr. Charles H. Spalding at the research laboratories of Unimation, Inc., in Mountain View, California. Dr. Spalding demonstrated the operation of the PUMA 500 robot manipulator for the team. This manipulator system consists of a five-degree-of-freedom electrically servocontrolled arm combined with an electronics package containing a DEC LSI-11 control computer, individual microcomputer systems for each degree of freedom, and drivers for the servo motors.

In a system such as the PUMA with separate microcomputers for each degree of freedom the individual microcomputers must receive commands specifying either the required rate of motion for their respective degrees of freedom, or the desired position of that degree of freedom, or both. In the PUMA, the individual microcomputers are controlled by a larger, more powerful microcomputer (a DEC LSI-11, a member of the PDP-11 family). The LSI-11 can direct the end effector of the robot manipulator to trace out a number of different predetermined paths in three-dimensional space. In the present configuration (depending on the complexity of the selected paths), on the order of 1000 programmed motion steps can be accommodated. The PUMA robot has about 500 distinguishable "parts," about 50 in the wrist assembly alone.

The next order of sophistication in robot control is at the level of elementary assembly operations. The command "put a washer on the bolt" requires the performance of subtasks such as:

- (1) Move the end effector to the washer supply.
- (2) Grasp a washer.
- (3) Move the end effector to the end of the bolt.
- (4) Orient the washer so it is perpendicular to the bolt.
- (5) Translate the washer so the axis of the bolt passes through the center of the hole in the washer.
- (6) Translate the washer along the axis of the bolt.
- (7) Release the washer.
- (8) Retract the end effector.

Still more sophisticated operations include the joining of subassemblies. To join two subassemblies each one must be brought into the proper relative position and several

washers, nuts, connectors, etc., must be installed. It is not clear, without further study, just how much of this hierarchy of operations could be controlled by the LSI-11 that has become an industry standard. However, the team has no doubt that a suitably powerful computer can be constructed in a module not exceeding 1 m³ in volume, which would also serve as a base for an advanced robot manipulator. Spaulding estimated that 5 years of adequate funding and manpower support could probably produce a robot manipulator system capable of assembling a duplicate of itself from prefabricated parts.

The team discovered no fundamental difficulties with the software, although the programming task will be extremely challenging. In current industrial robotics applications, each manipulator has a very limited number of tasks to perform. To use one manipulator to perform as many tasks, each of the complexity required by the SRS demonstration, the state-of-the-art in robot programming should be advanced considerably.

A top-level description of the steps required to produce a robot manipulator system complete with control computer and required electronics support (see fig. 5.29) might include the following sequence:

- (1) Assemble base frame and bolt it to floor.
- (2) Install card cages in frame.
- (3) Install cables between card cages.
- (4) Insert printed circuit cards into card cages.
- (5) Assemble manipulator waist joint support to base frame.
- (6) Install waist joint.
- (7) Install manipulator trunk (vertical member).
- (8) Install shoulder joint.
- (9) Install upper arm.
- (10) Install elbow joint.
- (11) Install forearm.
- (12) Install wrist joint.
- (13) Install end effector.
- (14) Install television camera mast.
- (15) Install television camera.
- (16) Connect television camera to electronics in base.
- (17) Connect manipulator to electronics in base.
- (18) Connect AC power and turn on computer.
- (19) Transfer construction software.
- (20) Boot up the new computer.

Having been replicated as thus detailed, the new robot is on its own.

APPENDIX 5K

ISSUES AND CONCEPTS FOR FURTHER CONSIDERATION

During the present study the Replicating Systems Concepts Team considered numerous concepts relating to the problems of self-replicating systems (SRS). The following is a partial list of various notions, ideas, suggestions, and research directions which came to the team's attention but which could not be adequately explored in the time available.

5K.1 Definitions

(1) Reproduction — What is a good, precise definition of "self-reproduction" or "self-replication"? What exactly is a "self-replicating system"? Does replication include any assembly of a physical copy of self? A copy of patterns? Is full assembly from molecular or atomic primitives required? Shall minimal reproduction be defined in terms of basic functions, bits of information processed, or some other measure? Is there some irreducible minimum necessary for "reproduction"? Most regard simple autocatalysis or Ashby's falling dominoes as not representative of "true" replication. However, perhaps a New Guinea islander would regard the cafeteria tray line (with seemingly equal justification) as "not real" when the source of human reproduction — viewing our environment as "too well-ordered to be believable."

(2) Growth — Exactly what is the distinction between growth and reproduction? What is the difference between these concepts and the notion of "self-organization"? What about "self-assembly"? These are common terms in regular use, and need to be more precisely characterized.

(3) Repair — What is the difference between self-repair and self-reproduction? Ordinarily replication involves duplication of the whole system, whereas repair involves replacement of only some subset of it. But at what point does "repair" become "reproduction"? Is machine self-repair or self-reproduction more difficult from a technical standpoint, and why? (Self-repair may require an analytical intelligence, whereas much of reproduction can be accomplished by "rote.")

(4) Telefactor, teleoperator, intelligent tools, autonomous, etc. — precise definitions are needed. Is there a clear dividing line between biological reproductive systems and advanced self-replicating robot systems?

5K.2 Evolutionary Development

(1) Which theoretical models would be easiest to cast into physical engineering form: the von Neumann kinematic model, the Laing self-inspection approach, the Thatcher methodology, or some other alternative? Under what conditions would each be desirable from a pragmatic engineering standpoint? The Laing approach, for instance, may prove superior to the von Neumann kinematic model in the case of extremely large, complex self-reproducing systems where the universe of components is so vast that self-inspection becomes essential to maintain order or where rapid evolution is desired.

(2) Specific "unit growth" and "unit replication" models of SRS were considered in detail during the present study. Under what conditions is one or the other optimum? Are there any fundamental differences between the two in terms of performance, stability, reliability, or other relevant factors? What might SRS emphasizing "unit evolution" or "unit repair" be like?

(3) Can SRS be designed to have few or no precision parts? Can milling and turning operations be eliminated? What substitutes might be found for the usual precision components such as ball bearings, tool bits, metering instruments, micron-feature computer chips, etc.? It is possible to imagine Stirling engines, solar mirrors, electromagnets, and mechanical gear-trains using only native lunar basalt, iron, and gases with no chemical processing — but are complete (but simple) SRS possible using just two or three non-chemically recovered elements/minerals? Could SRS be patterned after terrestrial biological protein synthesis, in which the factory is made up of perhaps two dozen fundamental "building blocks" (similar in function to amino acids) assembled in virtually limitless combinations?

(4) To what extent is intelligence a prerequisite for reproduction? (Amoebas appear to replicate well with almost no intelligence at all.) Does increasing intelligence make more efficient the processes of biological, and potentially machine, replication? Is there a law of diminishing returns, or does more intelligence always produce a superior reproductive entity?

(5) What forms of machine intelligence might possibly be required for a fully autonomous SRS, that are not now being adequately pursued by artificial intelligence researchers? A few possibilities include learning, memory structure, advanced task planning, adaptivity, association, creativity, intuition and "hunch" formation, hypothesis generation, self-awareness, survival motives, sophisticated database reasoning, symbolic meaning of knowledge, autonomous problem solving, and insight. Similarly, the state-of-the-art in robotics and automation from the viewpoint of SRS development needs to be examined.

(6) What is the least complex biological self-replicating system? How does it work? Can similar processes and analogies be drawn upon for use in the development of self-replicating machine technology? What is the minimum critical mass for a stable ecosystem? For a machine economy with closure?

(7) What is the possibility of semisentient workpieces? This concept is sometimes referred to as "distributed robotics." Perhaps each workpiece in an assembly process could be imbued with some small measure of machine intelligence using advanced microelectronic circuitry. Parts could then assist in their own assembly and subsequent installation and maintenance.

(8) Can computers be programmed to write their own self-assembly software? Perhaps an "artificial intelligence expert system" is required?

(9) What can be said about the possibility of machine "natural" or "participatory" evolution? How fast might machines "evolve" under intelligent direction? Is there any role for the concept of "sex" in machine replicating systems?

(10) Competing machines of different types, loyalties, or functions may interact destructively. For example, machines could disassemble others and cannibalize the parts. This might be viewed as adaptive or aggressive, if the disabled machine is willing; ecological if the stricken device is already dysfunctional and of no further use, etc. Or, competing machines could inject neighbors with senility software to accelerate deterioration as a prelude to subsequent cannibalism; "Frankenstein programs" in which the infected machine returns to its point of origin and adversely affects its creators; "hidden defect programs" which cause output of defective product so that the affected machine will be retired early; or "virus programs" which cause the host machine to begin producing output as directed by the invader to the exclusion of all else.

SK.3 Cost Effectiveness

(1) What are the proper tradeoffs among production, growth, and reproduction? Should these proceed serially or

simultaneously? Should the LMF be permitted to grow indefinitely, or should useful production be siphoned off from the start? How big is big enough? What are the trade-offs between "litter size" and number of generations in terms of efficiency and cost effectiveness? How long a replication time or doubling time is economically acceptable and feasible? Are there "diseconomies of scale" that make a small seed factory difficult to achieve? Should whole systems, or just their components, be replicated? At what point should factory components specialize in particular functions? Should these components be permitted to replicate at different rates within the expanding factory complex? What is the optimum mix of special-purpose and general-purpose robots? What are the other relevant factors involved?

(2) How and under exactly what conditions can a replicating system "exponentiate"? What should be exponentiated – economic value, number of items, quality, or complexity of product? What are the fundamental limitations and most significant factors? What are the important considerations of reliability, mean lifespan, replication time, unit and system costs? How does component reliability relate to replicating system lifespan? Multiple redundancy increases the mean time to failure but concurrently increases system complexity, which might lead to higher costs and added difficulty in overall design and coordination. How can error propagation in SRS be quantified and analyzed mathematically? Should evolutionary biological notions such as "mutation" and "survival of the fittest" be made a part of SRS designs?

(3) How can closure be defined, studied, and achieved? What are the different aspects of closure? How can closure be demonstrated? Is less than full closure acceptable in some applications? What might be the principles of "closure engineering"? To what extent should/can/must reproducing machines be energetically and materially self-sufficient? How many "vitamin parts" can be imported from Earth and still retain economic viability? Can artificial deposits of special materials be created on other worlds for the convenience of SRS machines (e.g., crash a comet into the Moon)?

(4) What sorts of useful output might self-replicating robot systems produce? Would there be an emphasis on services or products? Would terrestrial or extraterrestrial consumption dominate?

5K.4 Man and Machine

(1) What is the most appropriate mix of manned and automated functions in complex, self-replicating machine systems? Does this optimum mix vary from mission to mission, or can certain general categories be established?

For manned functions, what is the most efficient mix of physical and mental labors?

(2) What is the cost tradeoff between man and machine? Is, say, a fully automated lunar factory cheaper to design, deploy, and operate than one which is fully manned, or remotely teleoperated? Is a lunar base populated by humans cheaper than a “colony” of replicating machines? Is the oft-heard assertion that “in a factory with automation, productivity is inversely proportional to the number of human workers involved” true? What should be the ratio of biomass/machine mass in SRS factories?

(3) Is it possible that very highly advanced machines could evolve to the point where humans could no longer understand what their machines were doing? Would “their” interests begin to diverge from ours? Would they replace us in the biosphere, or create their own and not displace us? Would they keep us happy, feeding us the information we request while spending most of their time on higher-order operations “beyond our understanding”?

